Greater Amberjack Count

Lead PIs

Sean P. Powers[1] and Mark A. Albins[1]

Design/Modeling/Analysis Lead Co-PIs:

John Hoenig[2], Grace Chiu[2], and Lynne Stokes[3]

Regional Lead Co-PIs:

Texas: Jay R. Rooker[4]

Louisiana: Michael A. Dance[5]

Mississippi/Alabama: Sean P. Powers[1]

Florida/Georgia: Theodore S. Switzer[6]

South/North Carolina: Jeffrey A. Buckel[7]

Cross-regional Lead Co-PIs:

Kevin M. Boswell[8], Steven A. Murawski[9], Eric A. Saillant[10], J. Marcus Drymon[11], Matthew J. Catalano[12], Nicole M. Phillips[13], Stephen R. Midway[5], Steven B. Scyphers[1], Gregory W Stunz[15] and R. J. David Wells[4]

NOAA Collaborators:

Nate Bacheler[16] and Matthew Campbell[16]

Affiliations

1. University of South Alabama, Stokes School of Marine and Environmental Sciences, Dauphin Island Sea Lab
2. College of William and Mary, Virginia Institute of Marine Science
3. Southern Methodist University
4. Texas A&M University at Galveston, Marine Biology
5. Louisiana State University, Oceanography and Coastal Science
6. Florida Fish and Wildlife Conservation Commission
7. North Carolina State University, Applied Ecology
8. Florida International University, Biological Sciences
9. University of South Florida, Biological Oceanography
10. University of Southern Mississippi, Coastal Sciences
11. Mississippi State University, Coastal Research and Extension
12. Auburn University, School of Fisheries
13. University of Southern Mississippi, School of Biological, Environmental, and Earth Sciences
14. Florida International University, Institute of Environment
15. Texas A&M University at Corpus Christi, Marine Biology
16. NOAA Fisheries, Southeast Fisheries Science Center

Table of Contents

1. Executive Summary 3
2. Global review of Greater Amberjack life history 9
3. Population genomics 45
4. Environmental DNA 83
5. Exploitation 112
6. Movement ecology 118
7. Age and Growth 150
8. Acoustics Overview 177
9. Gear Calibrations 199
    9.1. Model-based calibration of survey gears 199
    9.2. Single to spherical camera calibration 245
10. Abundance Estimation 261
    10.1. Overview of Survey-Methods Approach 261
    10.2. Overview of Bayesian Approach 267
11. Regional Estimates 309
    11.1. Western Gulf of America 309
    11.2. Alabama-Mississippi 341
    11.3. Western Florida 380
    11.4. South Atlantic 461
12. Public Outreach 520
13. Co-produced science-angler involvement 532
14. Appendices 546
    14.1. Appendix A: Post-release Mortality 546
    14.2. Appendix B: Diet 549
    14.3. Appendix C: Otolith microchemistry 556
    14.4. Appendix D: Greater Amberjack Recreational Survey Instrument 566

1. Executive Summary

Sean P. Powers

Management of many reef fish species in the southeastern United States remains highly controversial with stakeholders questioning the results of federal stock assessments, which serve as the basis for management advice. In an effort to better inform resource agencies and build stakeholder confidence, the U.S. Congress identified funding to support projects whose goals included filling critical knowledge gaps as well as using the latest technology to estimate stock abundance independent of the current age structured stock assessment approach. Two previous projects focused on Red Snapper – one in the Gulf of Mexico (led by Greg Stunz, TAMUCC) and one in the South Atlantic U.S. (led by William Patterson, University of Florida). A third project focused on Greater Amberjack in the Gulf and South Atlantic (this project) was begun in 2021.

In the Southeastern U.S. there are two stocks of Greater Amberjack: a Gulf of America stock and a South Atlantic stock. According to the most recent stock assessments (2021 SEDAR 70), the Gulf of America stock is overfished and was experiencing overfishing. However, recent catch data (2023) indicated the stock was no longer experiencing overfishing. The South Atlantic stock was not considered overfished (2020 stock assessment) and not subject to overfishing based on 2022 catch data. Summary stock assessment information can be found on Stock SMART (https://apps-st.fisheries.noaa.gov/stocksmart?app=homepage).

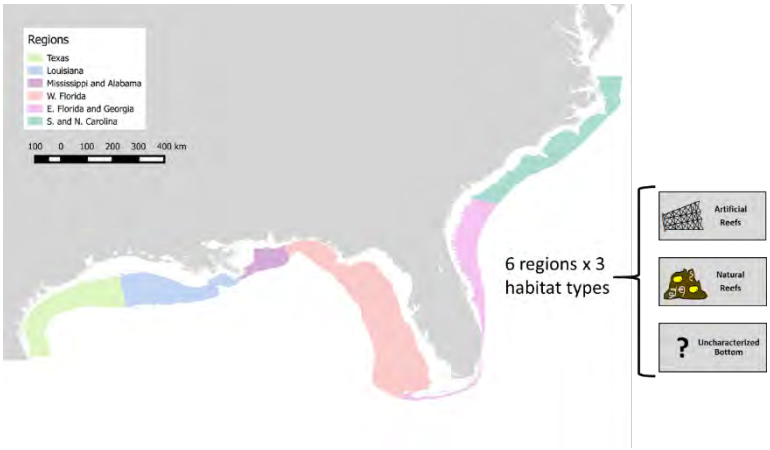


Figure 1.1. Map of project subregions. Sampling area delimited by 20 m and 150 m isobaths.

The current project proposed seven objectives that are discussed in this report: (1) synthesize bottom habitat observations and construct a spatially explicit habitat model throughout the study area (SA and GoM), (2) synthesize existing fisheries data and stakeholder knowledge regarding regional, habitat-specific abundance of GAJ across the study area, (3) design and conduct a comprehensive study to estimate, within a reasonable margin of uncertainty, the regional, habitat-specific estimates of abundance of GAJ using a combination of video observations (towed, ROV mounted, and stationary) and hydroacoustics, (4) determine movement behavior and connectivity of GAJ at local and regional scales using a combination of acoustic telemetry, conventional tagging and genetic markers, (5) assess the efficacy of environmental DNA (eDNA) methods in conjunction with visual and acoustic approaches to determine presence/absence and/or relative abundance of GAJ as well as closely related species, (6) update biological information for GAJ across the study region (age-length) and (7) engage with stakeholder groups and collaborate with existing outreach teams to facilitate stakeholder input and to communicate the results of the project.

Results for each objective are included in the sections of this report and briefly highlighted in this executive summary. Chapter 2 (life history) addresses elements of objective 1, while other elements of this objective are covered in the regional estimate chapters (11.1 through 11.4). It is important to note that the estimates derived in this report need to be viewed in the context of the limitations of the current habitat-based maps. While some supplemental funds were provided to facilitate habitat-based estimates in the South Atlantic, this supplemental effort was not designed to result in comprehensive maps of the region. In fact, higher CVs in many of the regions are, to a large degree, due to a lack of sufficient high-resolution bottom habitat maps (e.g., stratum C in Alabama/Mississippi region and Uncharacterized bottom in the Western Gulf). These are regions known to have abundant but scattered natural reef habitat; however, sufficient mapping has not been conducted to quantify the habitat extent with high certainty.

Regarding objective 4 (connectivity and movement), Chapter 3 details the results of the population genetics, Chapter 5 reports the results of conventional tagging including exploitation estimates, and Chapter 6 synthesizes the acoustic telemetry component of the project. Genetic markers population in this study did not reveal divergence of the Gulf and Atlantic stocks or other subdivision within US waters. The lack of genetic divergence found in this study was surprising but could indicate recent/ongoing mixing of previously divergent units, a hypothesis that could be examined further using higher density genome scans and analysis of identity by descent. In contrast to the pop genetics results, we found that exchange between regions was low, with some evidence of movement between the eastern and western Gulf, but no movements between the Gulf and Atlantic. Interestingly, there were a couple individuals tagged in the Atlantic that moved both across international boundaries and into basins outside of the study area (i.e. movements to Belize, Bahamas), demonstrating the species is capable of inter-basin movement. Habitat use and residency patterns were region- and habitat-specific, reflecting differences in the composition and distribution of habitat, with lower residency in areas of more continuous available habitat and higher residency at more isolated structures in the Gulf. Clear temporal trends in residency and movement were common to each region with movements peaking during winter/spring and coinciding with reproductive timing. Similarly, regional differences in movement patterns were tied to differing reproductive strategies between the Atlantic and Gulf to optimize temperature regimes for spawning and oceanography for larval transport. Greater amberjack in the Atlantic underwent lengthy migrations to warmer waters of the Florida Keys, while Gulf amberjack moved shorter distances to specific habitats (e.g., natural reefs) during the spawning season. We leveraged both the acoustic telemetry effort as well as a companion high dollar, conventional tagging program to estimate exploitation, natural mortality, and migration. Average annual exploitation rates (South Atlantic: 0.09, eastern Gulf of Mexico: 0.06, and western Gulf of Mexico: 0.06) lower than SEDAR estimates. Note, these include harvest plus dead discards. The *apparent* natural mortality estimate was estimated at 0.80 yr-1 [+/- 0.07]), which is higher than estimates of M for 1+ Greater Amberjack (0.25 to 0.3) used in recent stock assessments. Model-based acoustic tagging mortality was similar to external estimates at 0.23 (+/- 0.03). It is important to note that apparent natural mortality is inclusive of many Greater Amberjack individuals of unknown fate (not recpatured or detected by a hydrophone deployed during the study), including those that emigrated from the study area. Readers should note that many of the fish were tagged with acoustic tags that have a multi-year life span. Hence, as detection within the large collaborative network of researchers that maintain arrays of hydrophones are reported, the fate of many of these fish will become known and hence apparent natural mortality

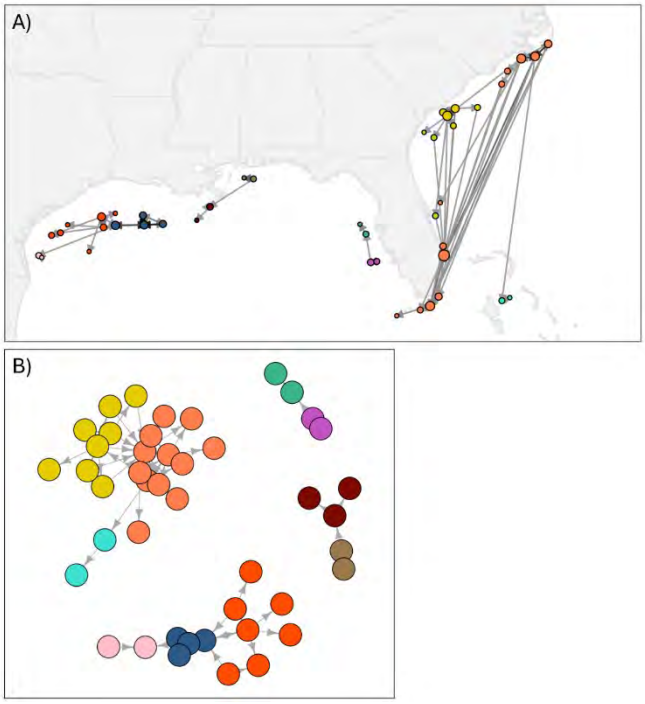


Figure 1.2. Network based on 100 km gridded nodes, plotting with spatial reference the respective nodes and edges including all movements of tagged Greater Amberjack. Degree centrality of each node is indicated by point size, edge weight is depicted by line color. Node colors are indicative of distinct communities identified in the cluster analysis for each network. B) The same network, plotted without spatial reference using Fruchterman-Reingold algorithm.

can be parsed into other categories. The analysis team will continue to revise these estimates as new data become available.

The original RFP released by MS/AL Sea Grant requested the inclusion of eDNA, a topic identified in the AJ scoping process that occurred prior to the project. In response, we designed a study, described in Chapter 4, to assess the efficacy of environmental DNA (eDNA) methods in conjunction with visual and acoustic approaches to determine presence/absence and/or relative abundance of GAJ as well as closely related species. Chapter 4 of this report details the developement and validation of a novel tetraplex ddPCR™ assay able to detect and distinguish between *Seriola dumerili*, *S. fasciata*, *S. rivoliana*, and *S. zonata* from seawater eDNA samples. To ground truth eDNA ddPCR™ data, we compared eDNA ddPCR™ detections and quantifications to abundance data inferred from multiple camera (ROV, S-BRUV, chevron trap) and acoustic (VPS array) gears deployed during a fisheries research gear-calibration cruise in Alabama waters in 2022. Both eDNA and camera gears detected the presence of *S. dumerili* and *S. rivoliana* at both sites and all sampling days, but not *S. fasciata* and *S. zonata*. EDNA concentrations were higher for *S. dumerili* than *S. rivoliana* at both sites for all sampling days, in line with abundance patterns obtained from camera gears. The highest correlation (r = 0.97) was obtained between the measures of eDNA concentrations sampled 'between' other gear deployments and ROV estimates of abundance. This high correlation adds to a growing number of studies demonstrating the potential of eDNA as an indicator of abundance for fisheries stock assessments. Because of budget limitations and the need to develop protocols, the eDNA component of the project was not intended to be implemented in all regions of this study.

The results of objective 6, update biological information for GAJ across the study region, which primarily focused on updated age-length relationships is included in section 7 of this report. The observed growth differences support the above-mentioned stock delineation. Genetic divergence between stocks is not evident but management histories and the exploitation histories of the Gulf and the Atlantic have been quite different, as the Gulf stock has been overfished for more than a decade. The South Atlantic has a less restrictive minimum size limit (SEDAR, 2020a), which may disproportionately remove faster-growing

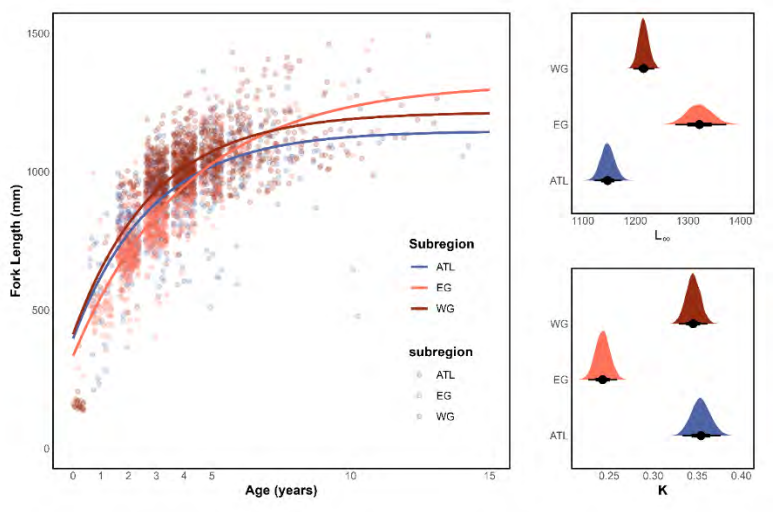


Figure 1.3: Right) Predicted Von Bertalanffy subregion-specific growth curves for greater amberjack (*Seriola dumerili*) predicted fork length as a function of calendar age for all observations (*n* = 2,823). Left) Posterior distributions, along with 95% and 80% credible intervals, for each growth parameter. L_∞ = asymptotic length; *K* = growth coefficient. ATL = Atlantic; EG = Eastern Gulf; WG = Western Gulf.

individuals. This could bias growth estimates toward larger asymptotic lengths without representing the true growth potential of the stock.

As expected, differences in water clarity as well as the shape and extent of many fish habitats (e.g., broad natural terraces of rocks, scattered rock outcroppings, prefabricated, artificial reef structures, vertically complex oil and gas platforms) meant that no one video-based gear could be used in all habitats. Section 8 gives an introduction and overview to the active acoustic gear and protocols developed for the project. The results of the regional abundance estimates are provided within each regional chapter (11.1 through 11.4). Chapter 9 details the results of several calibration experiments and approaches that are designed to examine the comparability of the gear types used (acoustic and video). Although substantive progress was made during the experimental work, a single grand calibration among gears was not developed. Chiu et al. (Section 9.1 and 9.2) detail the work and provide suggestions for further development of the promising calibration approach.

The overall and regional estimates (Chapter 11) were within the scale of the most recent stock assessments of Greater Amberjack. For the South Atlantic, the estimate derived from underwater cameras was 3.08 million with a percent coefficient of variation (CV% = [stdev/mean] * 100) of 35%. The estimate derived using acoustics was similar, 3.09 million, with a higher CV of 41.5%. For scale of comparison, the terminal year (2018) of the latest South Atlantic Greater Amberjack Stock Assessment was 2.64 million. While a CV around the stock assessment estimate is not reported, annual estimates over the last 20 years of the stock assessment (1999-2018) ranged from 1.7 to 3.9 million age 1+ Greater Amberjack. Our point-estimate for the camera-only Gulf of America population size is higher than that from the terminal year (2018) of the latest Gulf of America Greater Amberjack stock assessment (4.34 million in our study vs. 1.39 million from SEDAR 59). However, the CV for the camera-only estimate is quite large, at 46.4 %. As noted above, the higher CV for the Gulf estimate stems from the lack of high-resolution mapping data in the shelf-edge reef complex (stratum C) in the Alabama/Mississippi sub-region and uncertainty in uncharacterized bottom in the Western Gulf. For comparison the last twenty-year (1999-2018) range for SEDAR 59 is 1.39 to 3.50 million. Greater Amberjack population sizes greater than 4.34 million have been estimated in the Gulf, although these estimates were prior to 1987. The acoustic estimate derived by this study for the Gulf of America region was 1.45 million with a CV of 21%.

The PIs recommend the video-based methodology for any comparisons. Among the reasons for this choice is that the length frequency distribution can be derived. These length frequencies indicated differences in size composition between the three major areas (W. Gulf, E. Gulf and SA). The PIs also express caution with any direct comparisons of our results to the stock assessment results. We list the range of population size estimates derived from the stock assessment purely to inform the reader of the scale of our estimates compared to past management assumptions. While there is some overlap – for example, camera-based fishery independent surveys are used in the stock assessment and used in our estimation – the methodology used by the stock assessment is fundamentally different than adopted in our study. Further, our study is designed to be a point estimate in time (2023) and successful management requires estimating rates of exploitation on an annual basis. Comparisons of variance estimates are also not possible between the stock assessment approach and our study. Unlike the stock assessment, the original request for proposal requested that estimates be given by habitat type and region. This approach required multiple combinations of spatial and habitat strata to be incorporated into our design and estimation methods. Hence, CVs reported in this report reflect this suggestion and are influenced by the quality of habitat mapping and whether such habitat maps incorporate uncertainty.

An earlier version of the study's report benefited from the suggestions of a SEDAR sponsored review. That review primarily focused on only one of the objectives of our study - the abundance estimates. It should; be noted that regardless of the estimates derived, the study has fundamentally advanced our understanding of the biology and ecology of this important reef fish species. In writing the final report, the investigators incorporated as many of these suggestions from the review committee as feasible. The estimates and CV detailed in this report are not the same ones examined by the review committee.

Finally, the results of this study will continue to be improved as publications benefit from journal peer-review procedures and are ultimately accepted. We encourage the reader to refer to those contributions as they become available.

Table 1.1. Video-based estimates of GAJ absolute abundance on hard bottom and uncharacterized habitats between 20 and 150m depth. Uncertainty in the Western Gulf estimate is primarily due to variable fish counts and extremely large areas associated with uncharacterized and natural reef habitats. Uncertainty in the Alabama-Mississippi estimate is driven by variable fish counts and high and variable sidescan-mapping-based habitat area estimates in the outer shelf-edge natural reef complex.

Region	Estimate	SD	CV %
Western Gulf	2,548,592	1,733,436	68.0
Alabama-Mississippi	1,072,090	1,017,278	94.9
Western Florida	726,416	164,933	22.7
Total Gulf	***4,347,098***	***2,016,634***	***46.4***
South Atlantic	***3,079,636***	***1,077,180***	***35.0***

Table 1.2. Acoustic-based estimates of GAJ abundance on hard bottom and uncharacterized habitats between 20 and 150m depth. Note that uncertainty in the Alabama-Mississippi estimate is underestimated due to lack of sufficient information on some spatial strata within the sub-region.

Region	Estimate	SD	CV %
Western Gulf	615,788	211,828	34.4
Alabama-Mississippi	285,761	13,701	4.8
Western Florida	550,405	241,405	43.9
Total Gulf	***1,451,954***	***321,458***	***22.1***
South Atlantic	***3,089,458***	***1,281,063***	***41.5***

2. Global review of Greater Amberjack life history

Mischa K. Schultz[1], Michael A. Dance[1], Sean Powers[2], and Stephen R. Midway[1]

[1]Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, United States

[2]Department of Marine Sciences, University of South Alabama, Mobile, AL.

Abstract

The greater amberjack is a socio-economically important fish targeted in commercial, artisanal, and recreational fisheries around the world. Population trends are generally unknown; however, there are indications of localized declines in waterbodies such as the Gulf of Mexico. Despite the importance of the greater amberjack fishery worldwide, there has been relatively little effort to study this species, and as a result, the current literature presents a fragmented picture where the focus of greater amberjack scientific research has been on captive rearing and reproduction for aquaculture production. Knowledge of the ecology and biology of the greater amberjack is incomplete, which limits both stock assessments and the information required to maintain sustainable levels of exploitation. This review paper synthesizes available information relating to the greater amberjack and its global fisheries. Fundamental questions remain unanswered for the major stocks of greater amberjack, such as spawning seasonality, frequency, and estimates of fecundity, survival throughout the life cycle of the species, particularly for larval and juvenile fish post-settlement, and larval dispersion and connectivity patterns with adult stocks.

Introduction

The greater amberjack is a socio-economically important fish included in many commercial fisheries, particularly artisanal fisheries (Pereira et al. 2023), as well as recreational fisheries (Harris et al. 2011). Population trends are generally unknown; however, there are indications of localized declines in areas where significant fisheries exist, such as the Gulf of Mexico (SEDAR 2020). Despite the importance of the greater amberjack fishery worldwide, there has been relatively little effort to study this species, and as a result, the current literature presents an incomplete picture, where the focus of greater amberjack research has been on captive rearing and reproduction for aquaculture production (see Sicuro and Luzzana 2016 for a full review of *Seriola* spp. farming). Incomplete knowledge of the ecology and biology of the greater amberjack (particularly wild populations) limits stock assessments and hinders the ability to maintain sustainable levels of exploitation. The aim of this study was to synthesize available information related to the greater amberjack and its fishery. The review begins with an overview of the species' geographical distribution and genetics, followed by key aspects of its life history, including reproduction, age and growth, and feeding ecology. The current state of the fishery is then examined, and the review concludes by identifying existing knowledge gaps to establish priority research areas.

Geographical distribution

The greater amberjack has a discontinuous distribution occurring in tropical and sub-tropical shelf waters of the western Atlantic Ocean, eastern Atlantic Ocean, and Indo-West Pacific Ocean (Figure 2.1). In the western Atlantic Ocean, greater amberjack are widely distributed from Nova Scotia through the Gulf of Mexico and Caribbean, and into the waters off Central and South America to Brazil. In the

eastern Atlantic Ocean, greater amberjack are found from the Bay of Biscay on the Spanish coast to southern Africa. In the Indo-West Pacific, greater amberjack have been observed from South Africa throughout Australian waters to the South China Sea. The circumtropical distribution of the greater amberjack has been attributed to the dispersal abilities associated with a pelagic life-history (Swart et al. 2015). Recently, novel documentation of greater amberjack has taken place in Russian waters (Luzhniak et al. 2020), the eastern Arabian Sea (Rajesh et al. 2021), eastern Tasmania (Stuart-Smith et al. 2016), and in the south-east Pacific Ocean off the coast of Rapa Nui (Morales et al. 2022), and off the Pitcairn Islands (Duffy et al. 2017). Novel documentation of greater amberjack may be a function of increased effort or range expansion. Studies suggest range expansion could be related to adult dispersal, increasing ocean temperatures, or larval and adult transport facilitated by currents and gyres.

Genetics

Genetic tools are critical for defining geographic boundaries when conducting population assessments of marine fishes (Cadrin 2020). In greater amberjack, genetic investigations began in response to stock declines in the US and Gulf of Mexico during the 1990s (Gold and Richardson 1998), and concurrently with the growing interest in *Seriola* species for aquaculture (Nugroho et al. 2000). Since the early 2000s, several studies have provided insight into stock structure at broad spatial scales, although the degree of connectivity and mixing among populations is not well understood. Within ocean basins, subpopulations of greater amberjack are recognized, with minimal gene flow between geographical areas (Figure 2.1; Gold and Richardson 1998; Šegvić-Bubić et al. 2016). Gold and Richardson (1998) first examined greater amberjack spatial structure using genetic samples from the northern Gulf of Mexico and the US southeast Atlantic coast. A mitochondrial-based genetic assessment revealed differences between the Gulf of Mexico and the US southeast Atlantic, identifying at least two subpopulations (stocks) separated somewhere between the Florida Keys and the central-western Florida coast. This separation was hypothesized to result from reduced larval dispersal from potential spawning grounds in the Florida Keys, or alternatively, from recent or infrequent gene flow between regions. Using genetic data based on 11 loci, Hargrove et al. (2018) identified low differentiation among Gulf of Mexico samples and reproductive adults collected from two spawning areas, identifying that greater amberjack population connectivity is more complex than a two-stock, non-mixing model proposed by Gold and Richardson (1998). Similar to genetic studies, otolith shape analysis has also been used to investigate inter- and intra-basin population structure (Crandall et al. 2013). This work detected further evidence of differences between greater amberjack in the Gulf of Mexico and the US southeast Atlantic; however, substructure within the Gulf of Mexico was not clearly resolved. Other related examples of known genetic separation include greater amberjack originating from Vietnam and Japan (Nugroho et al. 2000) and the South China Sea and Japan (Araki et al. 2018), suggesting there may be at least two subpopulations of greater amberjack in the coastal waters of the Pacific (Figure 2.1): a population in the East China Sea (ECS) and in the South China Sea (SCS), however the degree of mixing across subpopulations in the Pacific is unclear. The subpopulation structure of greater amberjack of the Pacific is likely influenced by mixing during the spawning season or transport of pelagic eggs in the coastal waters of Japan to Vietnam (Nugroho et al. 2000). In the ECS from Taiwan to Japan, there is no significant genetic differentiation among greater amberjack samples, implying greater amberjack in the ECS are composed of one population (Hasegawa et al. 2020). Interestingly, there is no geographical barrier between the ECS and SCS; and greater amberjack tagged in the ECS do not exhibit migration patterns to SCS (Tone et al. 2022), potentially explaining the genetic differentiation between

subpopulations. To better define population structure between the Northern Atlantic and Mediterranean, Šegvić-Bubić et al. (2016) used both microsatellite loci and mitochondrial DNA applied to greater amberjack collected from the Canary Islands, northeastern Atlantic, and in the western and central Mediterranean (Balearic Sea, Tyrrhenian Sea, Ionian Sea, Strait of Sicily and Adriatic Sea). Mitochondrial DNA supported evidence of two phylogroups with a homogenous gene pool distributed across the Mediterranean, and the microsatellites confirmed a clear Mediterranean-Atlantic division with weak geographic heterogeneity between groups. Again, highlighting the complexity of greater amberjack intra-basin mixing patterns. Spawning site fidelity and circulation patterns could create barriers to gene flow that complicate intra-basin subpopulation structure.

Overall, genetic evidence indicates that greater amberjack exhibit population structuring at broad spatial scales, with distinct subpopulations likely separated by large geographic distances or oceanographic barriers between basins. Finer-scale structure within regions such as the Gulf of Mexico or the Mediterranean remains unresolved. These findings highlight the complexity of greater amberjack population dynamics, which may be shaped by adult movement patterns, spawning site fidelity, and variable larval dispersal. Despite progress, key questions remain, particularly regarding the mechanisms driving mixing among genetically distinct groups. Future studies that integrate genetic approaches with oceanographic models and movement patterns will be critical to understanding connectivity and refining stock boundaries for improved management.

Life history

Reproduction

The reproductive cycle of greater amberjack has been described based on the histological analysis of the gonads, examination of gonadal development stages, gonadosomatic index (Marino et al. 1995; Micale et al. 1999; Murie and Parkyn 2008; Harris et al. 2011; Zupa et al. 2017; Murie et al. 2020), along with the determination of sex steroid levels in plasma and gonad biochemical composition (Massari et al. 2004). Greater amberjack are gonochoristic. Histological processing shows that sex differentiation begins at 101 days post hatch (dph) with the formation of the ovarian cavity or testes at a total length (TL) of 14.98 ± 6.20 cm and is complete at 408 dph (41.22 ± 3.83 cm TL; Papadaki et al. 2021). A full review on oogenesis and spermatogenesis is available in (Zupa et al. 2017). Hormone-based sex discrimination (Aoki et al. 2019) and observation of urogenital pore characteristics (Smith et al. 2014) have been suggested as accurate and reliable non-lethal sex determination methods for greater amberjack. In hatchery-reared greater amberjack, no sexual dimorphism is observed, and the sex ratio has been reported as 1:1 (Papadaki et al. 2021). On the other hand, in wild populations the sex ratio can vary from 1:1 (Kozul et al. 2008; Smith et al. 2014) to female skewed in different geographical regions: 1:1.1 (US Atlantic, Harris et al. 2011), 1:1.7 (Gulf of Mexico, Murie and Parkyn 2008) to 1:2.5 (Gulf of Mexico, Thompson et al. 1999). Only one study found a male-dominated sex ratio of 1.1:1 in the Pacific Ocean (Kikkawa and Everson 1984). This pattern suggests that either females are being selectively harvested by the fisheries or that a naturally female-skewed (natal) sex ratio exists in the population (Smith et al. 2018). Although observed differences in sex ratios are often small and may reflect natural variability, some studies have attributed these patterns to differences in sampling methodology (e.g., fishery-dependent vs. fishery-independent data). Fishery-independent surveys may provide a more accurate representation of the true population sex ratio, as they are not influenced by size-selective regulations such as minimum size limits. Regardless of the cause, skewed sex ratios can reduce the

reproductive potential of the stock and may have important implications for population dynamics and management strategies, such as geographical or seasonal closures.

The age or length of a fish at sexual maturation can vary depending on environmental conditions and levels of fishing pressure (Kendall et al. 2014; Véron et al. 2020). According to gonadal analysis, greater amberjack can reach sexual maturity at an early age (≈1 year, Harris et al. 2011) and demonstrate sexual dimorphism in maturity schedules, with female greater amberjack maturing at larger sizes and older ages than males (Marino et al. 1995; Harris et al. 2011). Size at 50% (L50) maturity for females ranges from 733 mm FL (US Atlantic, Harris et al. 2011) to 1120 mm SL (Mediterranean, Marino et al. 1995). Reported data on size at 50% for males is limited but ranges from 644 mm FL (US Atlantic, Harris et al. 2011) to 1090 mm SL (Mediterranean, Marino et al. 1995). The earliest size at maturity recorded for greater amberjack comes from the US Atlantic Ocean: 464 mm fork length (FL) for males, and 514 mm FL for females (Harris et al. 2011). The timing of sexual maturation of greater amberjack is not always consistent among geographical regions (Table 2.1). Estimates of L50 based on physiological maturity indicate that greater amberjack in the US Atlantic likely achieve maturity at smaller sizes (males: L50 = 644 mm FL; females: L50 = 733 mm FL) and younger ages (A50 = 1.3 years) than in the Gulf of Mexico (females: L50 = 900 mm FL) and the Mediterranean. A lower size at maturity found by Sley et al. (2014) for fish sampled in the Gulf of Gabes is likely due to individuals with developing gonads being classified as mature. Across regions, most greater amberjack are sexually mature by age 5 or 6 (Kozul et al. 2008; Murie and Parkyn 2008).

Understanding the relationship between fecundity, spawning duration, and spawning frequency is crucial in fisheries biology for estimating reproductive potential (Murua and Saborido-Rey 2003; Lowerre-Barbieri et al. 2011; Tara Marshall 2016). Greater amberjack are multiple batch spawners, releasing several batches of eggs within a reproductive season and displaying indeterminate patterns of oocyte development (Marino et al. 1995). Because oocytes continue to mature and release during the spawning season (Murua and Saborido-Rey 2003), greater amberjack are known to have indeterminate fecundity. The spawning duration of greater amberjack, estimated using the gonadosomatic index (GSI) and histological analysis of migratory-nucleus (MN) and postovulatory follicles (POFs), can span up to seven months (Harris et al. 2011). Peak spawning duration ranges from 73 to 77 days (Harris et al. 2011; Murie et al. 2020). Spawning frequency, assessed by the incidence of mature females with POFs, indicates that greater amberjack spawn approximately every 4.5 to 5 days during peak periods, with longer intervals (12+ days) as the spawning season extends (Harris et al. 2011; Murie et al. 2020). An individual female may spawn up to 14 times within a season (Harris et al. 2011). In mature greater amberjack, GSI increases with age and size, suggesting that older fish have higher reproductive potential (Marino et al. 1995). Batch fecundity, or the number of eggs per batch, ranges from 0.5 to 6 million oocytes and increases with female size (Harris et al. 2011; Murie et al. 2020). By multiplying batch fecundity by the number of spawning events, potential annual fecundity for greater amberjack is estimated at 18 to 59 million oocytes for individuals 930–1296 mm FL (Harris et al. 2011). Fecundity estimates are not available for all subpopulations.

Despite the global distribution of the greater amberjack, all spawning studies have come from the northern hemisphere (Harris et al. 2011; Sley et al. 2014; Zupa et al. 2017; Murie et al. 2020). Spawning studies estimate spawning seasonality using gonadosomatic index (GSI) peaks and histological analysis to identify hydrated oocytes (Harris et al. 2011; Murie et al. 2020). The spawning season for greater amberjack begins when water temperatures warm up in the early spring months and concludes by

summer. Monthly GSI changes indicate reproductive peaks in different spawning grounds (Corriero et al. 2021). Peak spawning occurs from March – April in the Pacific Ocean (Kikkawa and Everson 1984), March – May in the Gulf of Mexico (Wells and Rooker 2004; Murie and Parkyn 2008; Murie et al. 2020), April – May in the western North Atlantic Ocean (Sedberry et al. 2006; Harris et al. 2011), and June – July in the Mediterranean (Sley et al. 2014; Zupa et al. 2017). Recorded GSI peaks are highest in the Pacific Ocean, potentially associated with regional differences in fecundity (Zupa et al. 2017). Many reproductive studies show that greater amberjack finish spawning by June or July (Sley et al. 2014; Murie et al. 2020), but it should be noted that both ripe greater amberjack and almaco jack have been collected as late as November (Kikkawa and Everson 1984). In mariculture, greater amberjack have been observed in natural spawning until October (Jerez et al. 2006). The protracted occurrence of ripe ovaries may indicate year-round incidental spawning (Kikkawa and Everson 1984). Additional evidence of protracted spawning comes from seasonal collections of juveniles ranging from 250–500 mm in spring, 180–500 mm in summer, and 8–500 mm in fall, suggesting that adult spawning occurs across an extended period from winter through summer (Galbraith et al. 2022), but further sampling of juveniles and reproductive adults throughout the year is needed to confirm this.

Similar to other reef fish (Biggs et al. 2021), greater amberjack form aggregations on spawning grounds during the reproductive season. Greater amberjack are presumed to form both transient and resident aggregations on spawning grounds, where some individuals migrate hundreds of miles to a spawning aggregation and others spawn within their home range (Biggs et al. 2021; Tone et al. 2021). Spawning grounds are associated with abrupt discontinuities in bottom structure, such as at the edge of the continental shelf (Hasegawa et al. 2020) and shelf-edge reef sites (Sedberry et al. 2006), with low salinities and chlorophyll concentrations, and weak current velocities (Aoufi et al. 2023). Spawning grounds are specific to geographical regions and have been observed from fisheries-based observations (Kozul et al. 2008) and inferred from satellite telemetry (Tone et al. 2021). Specific spawning grounds have been identified in the central Mediterranean (Kozul et al. 2008), the outer shelf of the South China Sea (Tone et al. 2021), and in the Caribbean along the continental shelf edge (Graham and Castellanos 2005). Multiple jack species, including greater amberjack, have been documented to aggregate and spawn at the same sites at groupers at numerous locations in the Caribbean (Graham and Castellanos 2005; Heyman and Kjerfve 2008). Greater amberjack in the East China Sea begin their spawning migrations in February, moving southward—potentially to a shared spawning ground in the Kuroshio Current off the east coast of Taiwan (Tone et al. 2021). A similar example of a shared spawning area between subpopulations occurs off the Florida peninsula, where individuals from both the Gulf of Mexico and the US South Atlantic have been observed (Murie and Parkyn 2008; Harris et al. 2011). More recent evidence suggests that the Gulf of Mexico subpopulation may also utilize spawning grounds located within the Gulf itself (Hargrove et al. 2018; Murie et al. 2020). Species distribution models created by Gruss et al. (2018) located potential spawning aggregations of greater amberjack across the entire shelf edge in the Gulf of Mexico. Spawning migrations in greater amberjack likely link subpopulations across their broad geographic range, facilitating gene flow and demographic exchange that support metapopulation stability (Secor 2015). These connections may buffer local declines by allowing individuals from more productive regions to replenish less productive areas, similar to patterns observed in species like Atlantic cod, Pacific herring, and Striped Bass (Robichaud and Rose 2004; Secor et al. 2009; Secor et al. 2020). Understanding spawning dynamics is especially important for amberjack, given their extensive movements between shelf-edge habitats and offshore spawning grounds (Tone et al. 2021).

While individuals form relatively large aggregations in association with spawning, individual spawning often occurs in pairs (Graham and Castellanos 2005). Greater amberjack courtship behavior has been observed following full moons between February and October (Graham and Castellanos 2005) and around sunset (Heyman and Kjerfve 2008). During a spawning event, greater amberjack display extended paired courtship in semi-pelagic habitat along the edges of reef drop-offs (Graham and Castellanos 2005; Heyman and Kjerfve 2008). Courting pairs exhibit behavior where the pursuing fish nuzzles the gonopore of the lead fish, following which the lead fish exhibits a color change where the body appears darker, and the fins are lighter. Courtship coloration is common for other members of Carangidae, such as the horse-eye jack (Caranx latus) and jack crevalle (C. hippos; Heyman and Kjerfve 2008). Females also exhibit rapid vertical swimming and diel vertical migrations during spawning (Tone et al. 2021), suggesting a complex set of cues driving reproduction across space and time.

Early life history

Ontogenetic habitat shifts and movements during early life stages play a key role in shaping connectivity, recruitment, and ultimately the population dynamics of marine fishes. After spawning, embryos of amberjack develop rapidly, with the first appearance approximately 10 hours post-fertilization, and hatching occurring 30–45 hours after spawning (Masuma et al. 1990; Papandroulakis et al. 2005). Hatched larvae measure only ≈2.88 mm total length (TL). The mouth opens 66 hours post-hatching, and swim bladder inflation occurs by 120 hours. Newly transformed amberjack larvae remain in the water column, where they continue developing through tail flexion at ~5 mm TL before transitioning into the post larval stage. The post larva quickly grows into a juvenile from 30 – 54 days (9 – 34 mm), where they may already have developed the characteristic eye band (Figure 2.2; Masuma et al. 1990; Hasegawa et al. 2020). Larval or early juvenile greater amberjack are transported by currents from spawning sites and accumulate in frontal zones until they reach about 30 mm (standard length; Wells and Rooker 2004).

The abundance and distribution of pelagic larvae are influenced by oceanographic and biological factors (Rooker et al. 2013; Mohan et al. 2017). Surface currents, temperature, and salinity are important in determining the distribution of early life stages of greater amberjack (Wells and Rooker 2004; Raya and Sabatés 2015; Hasegawa et al. 2017). Greater amberjack abundance is elevated in thermal fronts and associated eddies, particularly in sea surface temperatures (SST) ranging from 24–28°C, though the optimal range varies by region (Wells and Rooker 2004; Raya and Sabatés 2015; Hasegawa et al. 2017) Young-of-the-year (YOY) greater amberjack are more abundant in offshore areas characterized by higher salinity levels, typically between 33‰ and 38‰ (Wells and Rooker 2004; Sinopoli et al. 2006). The temperature and salinity values associated with increased greater amberjack larval abundance are not surprising, given that in laboratory rearing studies, optimum salinity and temperature conditions were between 32‰ and 35‰ (Chen C et al. 1997) and 26°C (Fernández-Montero et al. 2018). Laboratory studies indicate that shifts in temperature beyond this optimal temperature induce metabolic changes and oxidative stress, which may be detrimental to juvenile survival and growth (Fernández-Montero et al. 2018; Sanchez-Cueto et al. 2023), which may have implications for greater amberjack survival and growth in the wild. For successful recruitment, larval amberjack must be spawned in a time and location where they will encounter suitable settlement habitat as soon as 30 days post-hatch (dph). Frontal zones, known for accumulating larval fish, also serve as convergence points for Sargassum (Zhang et al. 2024). This co-occurrence may create high-density nursery habitats that enhance early survival. If winds

shift unfavorably or current strength weakens, sargassum can disperse or be lost, potentially leading to reduced recruitment success.

Recruitment success for greater amberjack depends on ocean currents, as well as encounter rates with suitable nursery habitat such as sargassum, flotsam, and fish aggregating devices (FADS; Badalamenti et al. 1995; Andaloro and Pipitone 1997; Wells and Rooker 2004; Hasegawa et al. 2017). After approximately 30 dph, greater amberjack will settle into nursery habitat and begin growing rapidly at an estimated ≈2 mm per day (Wells and Rooker 2004), benefiting from an abundant prey base, including sargassum crabs, shrimp, and plankton (Badalamenti et al. 1995). It has been hypothesized that the young of *Seriola* have increased visual sensitivity, which allows larvae to keep visual contact with large objects such as algae or sargassum clumps (Pavlov and Kasumyan 2000). Schooling behavior has not been examined for greater amberjack, but for the sister species, the Japanese yellowtail (*Seriola* quinqueradiata), schooling occurs after the fish reach 13 mm in length (Sakakura and Tsukamoto 1999). The cohort of pelagic juveniles for other *Seriola* species appears to begin schooling early and migrate together until the early juvenile stage, associated with floating sargassum, where most individuals are aged within ≈4 days of each other. Japanese yellowtail juveniles associated with the sargassum school in groups of up to 60 individuals around sargassum (Hasegawa et al. 2017). Likely to avoid predation, solitary individuals and small schools tend to remain close to the drifting seaweed, whereas large schools tend to actively swim around sargassum mats.

After approximately 150 days in the sargassum (Wells and Rooker 2004), greater amberjack juveniles will begin searching for reef habitat. As the sargassum drifts across suitable habitat, amberjack will recruit to natural or artificial structures. During this transitional period, greater amberjack undergo an ontogenetic diet shift where they begin primarily focusing on foraging fish, and potentially crustaceans (Badalamenti et al. 1995; Pipitone and Andaloro 1995; Kim et al. 2023). Greater amberjack are voracious eaters, supporting their relatively fast growth from settlement until maturity. The observation of an ontogenetic habitat switch is supported by the small sizes of greater amberjack collections associated with drifting sargassum (30–210 mm standard length), as well as a shift in diet that occurs around > 200 mm standard length (Pipitone and Andaloro 1995; Wells and Rooker 2004). The risks of predation during reef settlement may be reduced by the growth juvenile greater amberjack experience while associated with sargassum, but mortality is typically high for juvenile marine fishes, and juvenile losses remain unknown for this species. The long-distance habitat transition from offshore pelagic environments to reef-associated habitats carries inherent risks, particularly the potential for unsuccessful transport, which can significantly impact recruitment. Interannual variability in oceanographic conditions and fluctuations in sargassum availability may further contribute to differences in year-class strength. Although conclusions cannot be made, the influence of early life history on recruitment variability can only be hypothesized for greater amberjack due to the lack of data on larval abundance indices and mortality rates.

At approximately 300 mm FL, greater amberjack have already recruited to rocky hard bottom or artificial habitat (Figure 2.2; Wells and Rooker 2004). Sub-adult greater amberjack (250–350 mm fork length) may initially settle on smaller artificial reef structures such as concrete pyramids and domes before transitioning to larger offshore habitats (Dance et al. 2011). Similar patterns of transitional habitat use have been described for other reef fish species, such as groupers (Arena et al. 2007) and red snapper (Gallaway et al. 2009). Other artificial structures that provide additional refuge for young-of-year greater amberjack as they transition from pelagic to benthic include nearshore petroleum platforms or fish

aggregating devices (FADs; Sinopoli et al. 2006; Munnelly et al. 2020). Notably, FADS located farther offshore appear to retain juvenile amberjack for longer periods, whereas FADs near coastal reefs facilitate early dispersal (Sinopoli et al. 2006). The location of transitional habitats, such as FADs or artificial reefs, may influence juvenile distribution differently. Offshore FADs located far from suitable reef habitat may act as ecological traps—attracting juveniles that ultimately fail to settle—whereas nearshore FADs may facilitate successful dispersal and transition to benthic habitats.

Age and growth

Age determination in greater amberjack has been done with scales (Burch 1979; Kozul et al. 2008) and fin rays (Murie and Parkyn 2008), but is typically done using sagittal otoliths (Manooch and Potts 1997; Harris et al. 2011). Pectoral fin rays show distinct opaque zones alternating with translucent zones and have been used to establish comparative ageing criteria to otoliths (Murie and Parkyn 2008). Greater amberjack sagittal otoliths are thin and small, causing difficulties when ageing with sections of otoliths (Manooch and Potts 1997; Harris et al. 2011). Sectioned greater amberjack otoliths show a distinct core, first annulus, and second annulus that are clearly separated from each other (Murie and Parkyn 2008). The uncertainty in estimating the age of greater amberjack by otolith reading is linked to disagreement in the formation of the first annulus (Thompson et al. 1999).

Age determination is accompanied by various sources of error, and any age interpretations need to be validated. The goal of age validation is to validate the frequency of formation of a growth increment. Marginal increment analysis has been used to validate annulus formation for greater amberjack otoliths (Thompson et al. 1999; Murie et al. 2020); however, this method has failed in the past due to the inability to identify the condition of the otolith edge (Thompson et al. 1999). Because of the difficulties with identifying the condition of the otolith edge, mark-recapture of chemically tagged fish has been used to validate the periodicity of growth increment formation (Thompson et al. 1999). The marginal increment analyses and mark-recapture methods suggest a single annulus is deposited each year for adult greater amberjack (Manooch and Potts 1997; Thompson et al. 1999; Harris et al. 2011). Although there seems to be an agreement on annual periodicity of increment formation, differences in the exact timing of increment formation have been reported. Increment formation has been documented from November – March (Thompson et al. 1999), March – May (Manooch and Potts 1997), and June (Harris et al. 2011). Differences in time at increment formation could affect the assignment of calendar age for some specimens (Harris et al. 2011). Therefore, it is possible that specimens with the same birth year, even if they are captured on the same date, could be assigned to different age classes depending on the readability of the otolith (Rodriguez-Marin et al. 2022). These sources of uncertainty highlight the importance of accounting for variation in increment formation and otolith interpretation when assigning age classes, particularly for age-based assessments and growth analyses.

Growth patterns of adult life stages of greater amberjack have been studied using growth parameters derived from the von Bertalanffy growth function (VBGF, Murie and Parkyn 2008; Harris et al. 2011; Murie et al. 2020). The VBGF relates length to age (derived from otoliths) with three parameters: $L\infty$ (asymptotic mean length), K (growth coefficient), and t0 (the theoretical time at which fish length equals zero; Von Bertalanffy 1938). A comparison of published von Bertalanffy growth parameters by greater amberjack age and growth studies shows differences in L∞ (1109 – 1648 mm FL), K (0.175 to 0.230), and t0 (-0.59 to – -1.01; Table 2.2), potentially indicating spatial variability in growth patterns. Greater amberjack growth increases rapidly for the first 3 – 4 years of life. After age 9 – 10, growth slows

until reaching a maximum recorded age of 17 (Manooch and Potts 1997). Growth rates differ between males and females, with females reaching larger sizes than males (Murie and Parkyn 2008; Harris et al. 2011). While this phenomenon has been attributed to differences in longevity (Thompson et al. 1999; Murie and Parkyn 2008), it should be noted that Harris et al. (2011) found no difference in longevity between males and females in the US Atlantic. Limitations of these studies include low sample sizes of smaller individuals (<250 mm FL) and reliance on fishery-dependent data. While previous studies provide estimates of growth parameters, there have been no investigations into spatial or temporal variability in the growth of greater amberjack. Additionally, most age and growth data originate from the Gulf of Mexico and the US Atlantic, which may limit their applicability across all subpopulations. Updated and region-specific age and growth information is needed throughout the species' range to better inform conservation and management strategies.

Juvenile growth patterns of greater amberjack are less well studied than adult growth patterns. Wells and Rooker (2004) estimated daily growth rates of YOY greater amberjack (30 – 210 mm SL) using a linear growth equation, suggesting an overall mean growth rate of 1.45 mm/d. Specific growth rate was found to increase at both higher temperature and salinity, with individuals collected in warmer (29.8°C), more saline, offshore waters growing faster than individuals collected in cooler (28.7°C), nearshore waters. Cohort-specific growth patterns indicated that later cohorts had faster growth, likely due to warmer temperatures in the later season. If later cohorts have faster growth, the timing of spawning could influence recruitment potential for greater amberjack in different subpopulations.

Natural Mortality

Growth estimates indicate that greater amberjack grow rapidly throughout early life stages (Wells and Rooker 2004), which is often accompanied by high mortality rates up to 50% day-1 (Houde 2002). In the Gulf of Mexico, Wells and Rooker (2004) developed a single catch curve for YOY greater amberjack, estimating a daily mortality coefficient (Z) of 0.45% between 40 and 139 days of age. While this study provides an important baseline for early life mortality, more comprehensive research across the range of the greater amberjack is needed to evaluate mortality rates in larval and juvenile stages. Understanding how mortality varies spatially and temporally in response to environmental conditions is essential, as fluctuations in early life survival can significantly influence juvenile abundance and, ultimately, recruitment to the adult stock.

Life history-based estimates using catch curves have produced a wide range of total mortality (Z) values for greater amberjack, with estimates ranging from 0.41 in the Mediterranean, 0.60–0.65 in the US South Atlantic, and 0.68–0.71 in the Gulf of Mexico (Manooch and Potts 1997; Kozul et al. 2008). Catch curve-derived mortality estimates can be biased if data are collected from a single fishery, from regions experiencing low fishing pressure, or across different years that reflect variability in cohort strength. Patterns of recruitment to the fishery also vary regionally: in the Mediterranean, greater amberjack were fully recruited by age 4; in the Gulf of Mexico, also by age 4; whereas in the US South Atlantic, full recruitment occurred later, by age 8 (Harris et al. 2011). These regional differences in recruitment timing influence the interpretation of catch curves, as delayed recruitment can result in lower estimated mortality and obscure underlying population dynamics if not accounted for in model structure or sampling design.

Natural mortality (M) also varies across regions. In the US stock assessments, the Gulf of Mexico assumes an age-constant M of 0.28 $year^{-1}$, whereas the US South Atlantic assumes a slightly lower

constant rate of 0.25 year^{-1} (SEDAR 2020). Age-varying estimates of natural mortality provide additional insight, typically showing a decline in M with age, reflecting higher vulnerability during early life stages and increased survival among older fish. In the Gulf of Mexico, age-specific M values range from 0.750 at age 0 to 0.258 at age 10, while in the US South Atlantic, estimates are higher overall, ranging from 0.82 at age 1 to 0.29 at age 10. These regional differences in natural mortality likely reflect variation in life history traits such as growth, habitat use, and predation risk, and underscore the importance of using spatially explicit and biologically informed parameters when modelling population dynamics or developing region-specific management strategies.

Feeding behavior and trophic ecology

Studies on the trophic ecology of greater amberjack have been primarily restricted to stomach content analyses (Badalamenti et al. 1995; Andaloro and Pipitone 1997; Sley et al. 2016) with few studies using alternative approaches such as stable isotope analysis (Valls et al. 2012; Dillon et al. 2021). Greater amberjack are opportunistic feeders consuming a wide range of prey species but are generally classified as a piscivore (Badalamenti et al. 1995; Andaloro and Pipitone 1997; Sley et al. 2016; Mammel et al. 2024). Greater amberjack are considered roving predators that make feeding excursions over wrecks and reefs (Manooch and Haimovici 1983), and are active feeders throughout the year with differences in feeding intensity throughout the seasons (Mammel et al. 2023). Variation in feeding intensity is likely a response to energy demands for reproduction and spawning (Tone et al. 2021; Mammel et al. 2023).

Fisherman have reported greater amberjack forage primarily during the day (Andaloro and Pipitone 1997), a pattern corroborated in laboratory studies showing increased feeding under daylight conditions (Chen et al. 2007). Diurnal foraging behavior has also been observed in the closely related Japanese amberjack, which feed during the day and return to coastal waters at night (Kasai et al. 2001). Crepuscular feeding patterns, feeding during dawn and dusk, have been documented in pelagic predators like greater amberjack, particularly on hard-bottom habitats (Campanella et al. 2019). Recently, Tone et al. (2022) hypothesized a difference in greater amberjack foraging behaviour between reefs and offshore areas. Around topographic features, greater amberjack showed shallow foraging behaviour in both the day and nighttime and in offshore areas showed deep diving behaviour during the day Tone et al. (2022). Similar rhythms are commonly observed in other reef-associated species (Cardoso et al. 2020; dos Santos Silva Amaral et al. 2023). Interestingly, a video was captured of a juvenile greater amberjack displaying ram suspension feeding behavior (Sanderson et al. 1996) and scanning electron microscopy revealed denticles on the branchial surfaces of greater amberjack that could play a role in particle entrapment. No further records of ram suspension feeding exist in the literature.

Greater amberjack consume a wide range of prey, including fish, cephalopods, and crustaceans, in variable proportions throughout seasons, habitat types, and geographical regions (Badalamenti et al. 1995; Andaloro and Pipitone 1997; Sley et al. 2016). The diet composition of greater amberjack is composed of both pelagic and benthic species (Manooch and Haimovici 1983). Pelagic species of importance in the diet of the greater amberjack include Trichiuridae (Trichiurus spp.; Mammel et al. 2024), Carangidae (Trachurus spp; Scomber scombrus; Andaloro and Pipitone 1997), Clupeidae (Clupeidae spp.; Sley et al. 2016; Mammel et al. 2024), Engraulidae (Engraulis japonicus; Kim et al. 2023), whereas benthic species of importance include Serranidae (Centropristis phiadelphica; Foss 2016), Labridae (Foss 2016), Paralichthyidae (Paralichthys leigthostigma; Foss 2016), Sparidae (Sley et al. 2016),

and Leiognathidae (Leiognathus spp.; Mammel et al. 2024). While the greatest percentage of the greater amberjack diet is composed of fish, they do occasionally feed on cephalopods and crustaceans including Penaeidae (Penaeus spp.; Mammel et al. 2024), Cephalopoda (Loligo vulgaris and Sepiola sp.; Andaloro and Pipitone 1997), and Squillidae (Squilla mantis; Andaloro and Pipitone 1997).

The differences in pelagic and benthic diets may be attributed to habitat association (Foss 2016; Mammel et al. 2023) and imply that greater amberjack may use multiple foraging strategies between reefs and offshore areas (Tone et al. 2022). Individual greater amberjack found in natural habitats feed on a variety of other prey items including molluscs, squids, and crustaceans, whereas greater amberjack found at artificial habitats primarily feed on fish (Foss 2016). Greater amberjack collected off South Florida in the Atlantic Ocean have a higher proportion of crustaceans and cephalopods in their diet than greater amberjack collected in the Mediterranean Sea (Burch 1979; Sley et al. 2016). Another example of dietary differences comes from the Northwestern Hawaiian Islands where greater amberjack feed primarily on octopus and other bottom associated prey (Humphreys and Kramer 1984) but in the main Hawaiian Islands the diet is predominately water column prey, such as Decapterus spp. (Humphreys 1980).

There is some evidence that the composition of benthic and pelagic prey of the greater amberjack shifts throughout the year (Sley et al. 2016; Mammel et al. 2024). Seasonal diet composition appears to be regionally specific; in the Mediterranean fish are representative of 100% of the diet of greater amberjack in the winter with an increase in both molluscs and fish in the spring and summer (Sley et al. 2016), whereas in Taiwanese waters greater amberjack consume mostly pelagic species in spring and summer and demersal species in the winter (Mammel et al. 2024). Because greater amberjack are opportunistic predators, these changes in diet may be due to prey availability (Sley et al. 2016; Mammel et al. 2024).

Greater amberjack exhibit changes in dietary composition at different life stages (Mammel et al. 2024). Juvenile greater amberjack (<200 mm SL) consume plankton and invertebrates such as copepod and decapod larvae (Badalamenti et al. 1995). Once fish reach a size of greater than 200 mm SL they feed mainly on small crustaceans and fish (Badalamenti et al. 1995; Pipitone and Andaloro 1995; Kim et al. 2023). Juvenile amberjack (175 – 499 mm FL) collected in the Korean Strait fed almost exclusively on fish (97.3%) with a seasonal shift from consuming Japanese anchovy (Engraulis japonicus) in the summer to jack mackerel (Trachurus symmetricus) in the autumn (Kim et al. 2023). In contrast, juvenile greater amberjack in the Mediterranean (< 400 mm TL) consume higher proportions of crustaceans and cephalopods than adults (>610 mm TL; Mammel et al. 2024), again suggesting prey availability plays a role in the diet composition of the greater amberjack. After a shift to a piscivorous diet, greater amberjack feed on larger prey throughout ontogeny (Sley et al. 2016). In the Mediterranean, this pattern has been documented where smaller greater amberjack feed on fish in families such as Gobbidae, Sygnatidae, Belonidae, and Dactylopteridae, and adults increase their consumption of larger fish from Caragnidae, Balistidae, and Sparidae (Sley et al. 2016).

Piscivores typically have high frequencies of empty stomachs (Arrington et al. 2002); however, this pattern was not consistent with the greater amberjack throughout its geographical distribution (Table 2.3). The mean frequency of empty stomachs was 30% ranging from 4 – 60% for all greater amberjack diet studies (Table 2.3). Piscivores that consume pieces of their prey have empty stomachs significantly less often than piscivores that engulf their prey (Arrington et al. 2002). Laboratory studies have shown

the greater amberjack intake food items with manducation (Honryo et al. 2021), which may explain the lower frequencies of empty stomachs reported by diet studies.

The ontogenetic diet shift of juvenile greater amberjack observed from gut content analysis (Badalamenti et al. 1995; Pipitone and Andaloro 1995) is confirmed with results from a stable isotope analysis identifying that mean δ15N values are enriched in adults relative to juveniles (Rooker et al. 2006; Mammel et al. 2024), indicating a shift to a higher trophic position with increasing size or age (Rooker et al. 2006). Within reef predator guilds, greater amberjack exhibit the highest niche overlap with species such as Lutjanus spp. (Lutjanidae) and Mycteroperca phenax (Serranidae), suggesting potential competition for prey resources (Dillon et al. 2021). Stable isotope and gut content analysis by Mammel et al. (2024) found that greater amberjack in Taiwanese waters operate as a keystone species that plays a role in the structure and function between lower and higher trophic levels. A similar trend has been observed in the Mediterranean where greater amberjack control the abundance of smaller bodied fish in reef ecosystems (e.g., Labridae, Sparidae; Valls et al. 2012). As a pelagic reef predator, greater amberjack influence prey density and activity in the water column, underscoring the dynamic spatial interactions between predator behavior and habitat structure (Campanella et al. 2019). They also participate in multi-species aggregations with other predators, contributing to cooperative foraging events that enhance predation success (Auster et al. 2013). These complex ecological interactions may influence broader population processes such as individual growth and fecundity, ultimately shaping reef fish community structure.

Habitat use

The greater amberjack is characterized as a reef fish; however, reef association begins after a juvenile stage associated with floating objects such as drifting seaweeds (Figure 2.2; Badalamenti et al. 1995; Wells and Rooker 2004). Post-settlement greater amberjack shift to benthic habitat as they grow, where they frequently concentrate around reefs, rocky outcrops, wrecks, and when present, petroleum platforms (Manooch and Potts 1997; Harris et al. 2011; Benfield et al. 2019; Bacheler et al. 2022). Adult greater amberjack are closely associated with bathymetric features (Manooch and Potts 1997; Harris et al. 2011; Jackson et al. 2018), such as rocky reefs with high relief hardbottom and ledge habitats (Parker and Mays 1998), and are commonly found near the continental shelf in association with seamounts and slope habitat (Berry and Burch 1979; Sedberry et al. 2006; Tone et al. 2021; Bacheler et al. 2022). Small aggregations (10 – 100 individuals) of greater amberjack have also been observed at deep-water (175 – 546 m), high-relief habitats known as deep-water sink holes (Reed et al. 2005). Hard-bottom habitats, like the ones used by greater amberjack, are known to harbor increased prey densities that provide foraging areas for predatory fish (Flavio et al. 2023), potentially explaining the association of greater amberjack with high-relief features (Tone et al. 2022).

Artificial reefs have been widely deployed as a tool to enhance fisheries that hold high fish densities (Paxton et al. 2020). Greater amberjacks are found in high abundances at artificial reefs, likely due to increased structural complexity and vertical relief (Benfield et al. 2019; Garner et al. 2019; Gallaway et al. 2021). One type of artificial structure, petroleum platforms, provide habitat throughout the water column and have a high ratio of structural surface to seafloor surface area (Claisse et al. 2014). In the Gulf of Mexico, greater amberjack have been recorded in high numbers on petroleum platforms (Benfield et al. 2019; Bolser et al. 2020). Gallaway et al. (2021) recently estimated that approximately 45% of the greater amberjack stock resides on these platforms.

While the importance of hard bottom habitat to greater amberjack is well-documented (Harris et al. 2011; Jackson et al. 2018; Tone et al. 2021; Bacheler et al. 2022), marine fish habitat may also be defined by environmental conditions such as water temperature, dissolved oxygen, salinity, and light availability (Bacheler et al. 2022; Tracey et al. 2023). Greater amberjack habitat use in relation to environmental factors have been studied using underwater video (Bolser et al. 2020; Bacheler et al. 2022), satellite telemetry (Boyle et al. 2022a; Tone et al. 2022), and fisheries catch data in conjunction with remote sensing (Mammel et al. 2022). Sea surface temperature (°C), water column temperature (°C) and salinity (psu) seem to consistently influence the distribution and abundance of greater amberjack (Bolser et al. 2020; Bacheler et al. 2022; Mammel et al. 2022). Salinity stress can negatively impact fishes through the increased energetic costs of osmoregulation (Boeuf and Payan 2001), so it is not surprising that greater amberjack respond to water column and sea surface salinity (Bolser et al. 2020; Mammel et al. 2022). Higher relative abundance of greater amberjack occurs with increasing water temperature (15 – 30°C; Bacheler et al. 2022; Mammel et al. 2022). Tagged greater amberjack in the western Pacific inhabited a wider range of water temperatures from 9.5 to 28.1°C (mean: 20.7°C; SD: 1.8; Tone et al. 2022). Greater amberjack are caught more frequently when sea surface salinity ranges from 33 – 35 PSU (Mammel et al. 2022), and their vertical distribution is constrained by lower salinity values (10 – 20 psu; Bolser et al. 2020). Fish species distribution and abundance can be explained by chlorophyll concentration; however, only one recent study has attempted to relate chlorophyll concentration to greater amberjack distribution (Mammel et al. 2022) and found that high greater amberjack catch was related to chl-a concentration ranging from 0.5–1.5 mg/m3.

Apparent depth use of greater amberjack has historically been inferred from fishery-dependent and fishery-independent surveys (Bacheler et al. 2022) and more recently, has been directly observed using satellite and acoustic telemetry (Jackson et al. 2018; Boyle et al. 2022b). These studies reveal a wide range of depths occupied by greater amberjack, yet consistently suggest individual variability, a general avoidance of surface waters, and a tendency to occupy depths shallower than 100 m (Jackson et al. 2018; Bacheler et al. 2022). Surface avoidance may be related to decreased salinities in surface layers in some regions (Bolser et al. 2020), while depth use is further influenced by habitat type (Tone et al. 2022), and body size (Jackson et al. 2018). Smaller individuals (<762 mm FL) typically use shallower waters (~30–40 m), whereas larger individuals (>762 mm FL) are more frequently observed at greater depths (>45 m). Around structured topographic features, greater amberjack commonly occupy depths around 35 m but have also been observed using much deeper habitats, ranging from 150 to over 500 m in offshore areas (Tone et al. 2021; Tone et al. 2022).

Future studies should focus on increasing knowledge of greater amberjack habitat utilization, specifically age-size related habitat use and habitat-specific productivity. In areas with artificial habitat, such as the Gulf of Mexico, understanding how activity differs by habitat type may be important in management focusing on creating or conserving sites. Understanding habitat utilization in relation to environmental factors such as dissolved oxygen and temperature may help predict how oxygen minimum zones, hypoxia, and climate change will impact greater amberjack.

Movement ecology

Knowledge of habitat use and movement patterns among marine fish populations are important to determine population dynamics, spatial structuring and connectivity, and to guide fisheries management (Frisk et al. 2013). Until recently, greater amberjack movement patterns have been

estimated using conventional tag-recapture data (McClellan and Cummings 1997), which is reliant on fisheries dependent recoveries. While greater amberjack tag-recapture data has provided important information on straight line displacement from tagging sites and mortality (McClellan and Cummings 1997), specific movement patterns, including depth use, movement between habitat types, and temporal movement patterns remains unknown when using conventional tagging methods. Advancements in tagging technology (satellite and archival tags) have been useful in highlighting marine fish migration routes and movements between habitats or spawning grounds (Feeley et al. 2018).

Summarizing what is known from both conventional tagging and telemetry studies, greater amberjack are considered seasonal migrants but display two movement states: one characterized by high site fidelity and localized behavior, and another involving seasonal or long-distance migrations, particularly related to spawning (McClellan and Cummings 1997; Ingram and Patterson 2001; Tone et al. 2021). This dichotomy is consistent with the concept of partial migration, where only a portion of the population undertakes long-distance or seasonal migrations while others remain resident (Secor 2015). The most comprehensive study of greater amberjack movement behavior comes from the East China Sea, where both vertical and horizontal movements were described during pre-spawning and spawning periods (Tone et al. 2021). Tagged individuals in this study exhibited distinct seasonal movement patterns. Greater amberjack demonstrated high site fidelity to a shallow ridge habitat, remaining resident for much of the year, but undertook directional migrations to offshore areas in spring, coinciding with the spawning season. In winter, greater amberjack in the East China Sea move to potential foraging grounds, while spring migrations to offshore spawning sites may allow individuals to take advantage of prevailing current systems that facilitate larval dispersal. In US waters, seasonal movements are also apparent, with individuals regularly migrating southward from the South Atlantic Bight (North Carolina, US) to southeast Florida in the US Atlantic coast (McClellan and Cummings 1997). In contrast, other individuals, including some observed in the western Atlantic, appear to undertake unpredictable, long-distance movements from the US Atlantic (North Carolina, South Carolina, and Florida) to Caribbean and Bahamas, as well as some movement to Gulf of Mexico (Berry and Burch 1979; McClellan and Cummings 1997). The longest recorded movement from tag-recapture data was a greater amberjack that travelled 2510 km from the point of tagging (Jacksonville, Florida) to the point of recapture (Colombia, South America; Berry and Burch 1979). This extreme movement may reflect a more exploratory or condition-dependent migratory strategy, underscoring the variability in movement behaviors that likely exist within this species. These patterns suggest that movement behavior is closely linked to reproductive timing and resource availability, and may play a key role in shaping connectivity, recruitment dynamics, and regional population structure.

Although the literature suggests greater amberjack are seasonal migrants, there is evidence some individuals may be range residents living year-round in a home range (Baeyaert et al. 2018). For example, greater amberjack tagged in the Florida Keys National Marine Sanctuary remained within the sanctuary concentrated around sea mounds within a 2,000 km2 area (Baeyaert et al. 2018), and greater amberjack are often recaptured within 160 km of tagging the site (McClellan and Cummings 1997). This pattern is particularly evident for smaller individuals (< 711 mm TL), potentially indicating a relationship between movement state and life stage (Ingram and Patterson 2001). Or, perhaps, seasonal migrations occur within a home range, which has been seen in a population of greater amberjack tagged in the Kuroshio off eastern Taiwan where individuals traveled up to 1212 km but were recaptured within 40 km of the tagging site (Tone et al. 2021). Baeyaert et al. (2018) documented one individual residing in a

single reef for multiple years, with high site fidelity but increased roaming towards coastal areas in the spring and summer months. Resident-contingent greater amberjack exhibit fine-scale movements, including circling individual artificial reef structures and transitioning between artificial and adjacent natural reef habitats (Tharp et al. 2024), potentially following prey densities within a home range or core area.

The coexistence of both resident contingents and migratory groups within a single population of a marine reef predator such as greater amberjack may reflect diverse life history strategies shaped by environmental conditions, reproductive investment, or individual condition. Although much remains unknown about greater amberjack migrations and the connectivity among spawning groups, the presence of both range residents and migratory contingents suggests that these divergent movement strategies could enhance population resilience. By spreading reproductive effort across time and space, such behavioral variation may serve as a form of bet-hedging that buffers against environmental variability and stochastic recruitment failure.

Existing telemetry studies on greater amberjack are limited by small sample sizes and regional focus, making it difficult to generalize movement patterns or fully evaluate connectivity across the species' range, particularly in systems with different oceanographic features. Nonetheless, what emerges clearly is that greater amberjack are individualistic marine predators that exhibit a high degree of behavioral plasticity in their habitat use and movement patterns. Understanding the environmental and biological factors that influence migration initiation is critical. Resident individuals may be more susceptible to localized depletion, while migratory contingents may complicate management by linking distant regions through demographic connectivity. Recognizing and accounting for this movement complexity is essential for developing effective spatial and temporal management strategies for the species.

Fisheries, Harvest, and Management

Greater amberjack has gained worldwide market appreciation due to rapid growth and excellent flesh quality (Ibrahim et al. 1989; Nakada 2010). Greater amberjack farming was developed in the late 1970s and remains an important aquaculture fish in Japan, China, and in the Mediterranean (see Sicuro and Luzzana 2016 for a full review of *Seriola* spp. farming). Outside of aquaculture, the greater amberjack is targeted by artisanal, commercial, and recreational fisheries throughout its geographical range. Commercial gears used to harvest greater amberjack include vertical line or handline and longline (SEDAR 2020), mid-water longlines, purse seines, and gillnets (Ulman et al. 2013). Commercial interest in greater amberjack, at least in the US, was apparently driven by consumer acceptance of greater amberjack as an edible fish and replacement of red drum in the blackened fish market (Gold and Richardson 1998). Greater amberjack are caught by gillnet or drift net in artisanal fisheries (Colloca et al. 2004; Lucchetti et al. 2017). Because of its size and fighting capabilities, the greater amberjack is also a popular sport fish targeted by recreational anglers using rod and reel and are occasionally captured by speargun.

Fisheries landing and catch statistics are fundamental in assessing the health of fish stocks (Ovando et al. 2021). The relative lack of data on greater amberjack fishing effort and fishery landings worldwide may be an obstacle in sustainable management of this species in the future. Greater amberjack catch and landings data are available in some areas of its geographical distribution (Mediterranean, western Indian Ocean, and the US); whereas long-term fishery landings data are lacking for greater amberjack populations in South America (both Atlantic and Pacific), Atlantic coast of Africa, and most of the

western Pacific. According to the Food and Agriculture Organization (FAO), global landings data for greater amberjack date back to 1967 (~100 mt; FAO 2021). Global catch increased in the mid-1990s (from 400 mt to 1,500 mt) and has continually increased to ~3,500 mt in 2020 (Figure 2.3).

The landings data by region indicated the US and Mediterranean are the countries with the highest catch with different trends by region (Figure 2.3). This review will focus on highlighting the trends in catch of these two regions. In the Mediterranean, there is a gradual increase in catches since 1967 which follows the same pattern as the global trend in catch (Figure 2.3). The catches from the Mediterranean are likely driving the global trend in an increase in landings. Among countries in the Mediterranean region, Greece has the highest recorded landings followed by Italy (Figure 2.3). Landings in Greece and Italy are opposite; with catches in Greece increasing from 1967 to the present while catches in Italy have generally declined over the same time period (Figure 2.3). In contrast, the reported landings of greater amberjack in the US began in 1995 at 800 tonnes. Landings data fluctuated for the first 20 years of reporting, but since 2014, reports of greater amberjack catches in the US are progressively declining (Figure 2.4).

Estimates of stock size are limited for greater amberjack fisheries globally, and formal stock assessments are largely absent outside U.S. waters. Stock assessments for greater amberjack are routinely reported for the Gulf of Mexico and US Atlantic stocks (SEDAR 2020). The greater amberjack stock in the Gulf of Mexico has been estimated to be overfished; spawning stock biomass (SSB) was estimated to be 10% of the virgin stock biomass (SEDAR 2020), much lower than the SBB at maximum sustainable yield for many marine fishes (40%; Thorson et al. 2012). Interestingly, the SBB for greater amberjack in the Gulf of Mexico has remained below 15% of virgin stock biomass since 1987 (SEDAR 2020). For comparison, the US Atlantic greater amberjack stock is not overfished or undergoing overfishing, and SSB is estimated at 69% of virgin stock biomass (SEDAR 2020). The SSB in the US Atlantic has shown a consistent increase since the 2000s (SEDAR 2020). Current efforts are underway in the US Atlantic and Gulf of Mexico "to develop additional data sources, assessment approaches, and knowledge to improve agency and agency-independent estimates of abundance of the greater amberjack throughout the greater amberjack range" (MASGO-20-0249, The Mississippi-Alabama Sea Grant Consortium). Efforts such as these in the Gulf of Mexico and US Atlantic may be useful for management of greater amberjack fisheries around the globe.

Concern over the status of the greater amberjack stock has prompted efforts to address population status in the Mediterranean (Pereira et al. 2023; Hamwi et al. 2024). Researchers in the eastern Mediterranean in Syrian waters assessed the population dynamics of greater amberjack using artificial neural networks and fuzzy logic (Hamwi et al. 2024). Results from their analyses suggested that greater amberjack in the eastern Mediterranean are experiencing high exploitation rates from fishing pressure along the Syrian Coast, potentially leading to decreased abundance. Similarly, a study conducted by Pereira et al. (2023) analysing landings data for 72 years (1950 – 2021) found a continuous decline in greater amberjack landings attributed to fishing pressure, reduced fishing capacity, and decreased nutrient flux in the western Mediterranean. These studies highlight another example of depletion of greater amberjack stocks across their range.

Several management tools have been applied to greater amberjack fisheries including size restrictions, restrictive harvest limits, and temporal closures. Legislation has been passed in the Gulf of Mexico closing the fishery for greater amberjack during the spawning season (SEDAR 2020), and in the

Mediterranean to protect recruiting juveniles from purse seines (Pereira et al. 2023). Size restrictions are in place in both the Gulf of Mexico and US Atlantic waters with a minimum size limit of 864 mm (34") and 711 mm (28") fork length, respectively (SEDAR 2020). Little management information is available for small-scale fisheries targeting greater amberjack; however, recent studies report that the greater amberjack is an important part of such fisheries in the Mediterranean (Pereira et al. 2023; Palmer et al. 2024) and Brazil (Seminara et al. 2024). In the Balearic Islands off the coast of Spain, greater amberjack catch represents the fifth most valued fisheries resource where the purse seine fleet accounts for the highest percentage of catches (Pereira et al. 2023). Similarly, in small-scale fisheries in Cyprus, greater amberjack are one of five species comprising 70% of total landing mass (Palmer et al. 2024). The scarcity of information on small-scale fisheries may impair management decisions and protection of these types of fisheries for greater amberjack.

Addressing future research needs

Greater amberjack populations are highly fragmented across their geographic range and there is evidence of decline in subpopulations. Local declines are not considered a threat to its global population, and therefore the greater amberjack is classified as least concern by the International Union for Conservation of Nature (IUCN) red list of threatened species. Numerous studies contribute to a general understanding of the biology and ecology of the greater amberjack; however, there is a geographical bias to subpopulations in the Gulf of Mexico, US Atlantic waters, and in the Mediterranean, with fewer studies in the Pacific, South Atlantic, and Indian Oceans. Many tools and techniques have been used to investigate the biology and ecology of greater amberjack including telemetry (Murie et al. 2011; Tone et al. 2021; Tone et al. 2022), histology (Harris et al. 2011; Murie et al. 2020), underwater video (Bacheler et al. 2022), habitat mapping (Mammel et al. 2022), and genetics (Šegvić-Bubić et al. 2016). Despite these efforts, it remains unclear how well biological and ecological information from one region translates to other parts of the species' range, and key knowledge gaps persist. Throughout most of the greater amberjack range there is a lack of knowledge about age and growth, reproductive biology, and feeding ecology. Aspects of the biology and life history of greater amberjack are best understood for the adult life stage, with minimal available data on early life and juvenile stages across the range. Fundamental questions remain unanswered for the major stocks of greater amberjack, such as 1) spawning seasonality, frequency, and estimates of fecundity 2), spawning site fidelity and aggregation sites, 3) survival throughout the life cycle of the species, particularly for larval and juvenile fish post-settlement, and 4) larval dispersion and connectivity patterns with adult stocks.

The dynamics of greater amberjack spawning and their eventual contribution to the adult stock represent an emergent and critical area of research essential for effective management. Across their range, estimates of spawning seasonality and age at maturity have been reported, offering foundational inputs for population assessments. Yet reproductive success is influenced by more than egg production and timing alone; it is fundamentally tied to the survival of offspring to reproductive age, a process shaped by a suite of factors including spawning location, larval dispersal and retention, nursery habitat availability, recruitment success, and ontogenetic habitat shifts. The precise location of offshore spawning aggregations and the return movements of adults remain long-standing uncertainties in greater amberjack biology. Recent studies indicate potential spawning grounds across the shelf edge in the Gulf of Mexico (Gruss et al. 2018) and around the continental shelf in the southern ECS (Hasegawa et al. 2020), underscoring the importance of shelf edge habitats to the species reproductive ecology across their distribution. Observations of long-distance adult movements, potentially related to

spawning migrations, contrast with evidence that some individuals may reproduce within localized home ranges. These patterns raise the possibility of migratory contingents within populations which have implications for both connectivity and stock structure. While large-scale movements of adults may contribute to genetic and demographic connectivity within oceanic basins, information on stock mixing during early life stages remains absent. Studies employing otolith microchemistry or large-scale telemetry studies would provide critical insights into natal origin, migration routes, and spatial structuring of spawning contingents. and should be prioritized to fill knowledge gaps about the reproductive dynamics of greater amberjack across their range and support the development of spatially informed management strategies.

Greater amberjack undergo an ontogenetic habitat shift from an offshore pelagic larval life to a benthic reef associated adult life (Wells and Rooker 2004; Sinopoli et al. 2006). Habitat shifts influence population connectivity and structure. These habitat shifts play a key role in shaping population connectivity and structure. During the pelagic larval stage, dispersal across large spatial scales serves as the primary mechanism for recruitment and colonization of new habitats, making this life stage essential for maintaining gene flow and population resilience. Ensuring connectivity among larval, juvenile, and adult habitats is therefore critical for sustaining healthy reef fish populations. Models that do not account for the complexity of population connectivity may lead to misinterpretation of fishery status (Frisk et al. 2013). There is little knowledge on the dispersal and migration patterns, home range size, and habitat use of the greater amberjack. Given the importance of fish movement in population structuring, more studies are needed to examine greater amberjack daily movement patterns, habitat use across life stages, and turnover of individuals within habitats. Understanding movement behavior is particularly relevant for marine management and conservation, especially in the context of marine protected areas (MPAs). MPAs are generally more effective at rebuilding spawning biomass for species with limited juvenile and adult mobility and small home ranges (Botsford et al. 2003). The mobility of juvenile greater amberjack, though, remains poorly understood. Gaining insight into these dynamics is essential to evaluate the potential benefits of localized habitat protections and to predict how spatial management measures might influence overall population productivity.

Population assessments for greater amberjack are currently available from US waters and parts of the Mediterranean, but formal evaluations are lacking across much of the species' broader distribution. A recent population assessment of the Syrian stock, using neural networks and fuzzy logic, indicated high fishing mortality and signs of vulnerability (Hamwi et al. 2024). Although no formal stock assessment has been conducted in the eastern Mediterranean, Pereira et al. (2023) reported a continuous decline in landings over the past several decades, signalling potential stock depletion. In contrast, assessments in US waters present a regional dichotomy: the Gulf of Mexico stock is considered overfished and subject to overfishing, prompting multiple rebuilding plans, whereas the South Atlantic stock is currently not overfished and is managed under sustainable harvest levels. These disparate assessments underscore the influence of socio-economic and ecological factors on population dynamics and highlight the need for sustainable management, particularly in regions where greater amberjack represents a valuable marine resource, such as the Balearic Islands. Population assessments in the Gulf of Mexico, Atlantic, and Eastern Mediterranean all assume spatially invariant life history traits within a given stock. Yet numerous age and growth studies suggest considerable variability in growth rates and age-at-maturity among regions (see Table 2.1), challenging this assumption. Environmental gradients and spatially heterogeneous fishing pressures likely drive this geographic variation in life history traits and may also

contribute to demographic variability within a single stock. Because growth and maturity directly influence stock productivity, failure to account for spatial variation in these traits may result in biased estimates of stock status and yield potential. As such, future research should prioritize the investigation and incorporation of spatially explicit life history parameters to better characterize potential stock substructure and inform more spatially nuanced management strategies for greater amberjack.

Integrating Ecosystem Approaches into Greater Amberjack Management

Future studies on greater amberjack should integrate multiple data streams into ecosystem-based models to improve fisheries management (Chagaris et al. 2019; Lowerre-Barbieri et al. 2019; Thorson et al. 2021). The greater amberjack is a strong candidate for ecosystem approaches to management due to its socio-economic importance, broad and complex habitat use, trophic role as a mid-level predator, and evidence for regional variation in life history traits. Current assessments rely on single-species models that assume uniform biological characteristics across the species' range, and do not account for predator-prey interactions, habitat connectivity, or broader ecosystem effects. Despite interest in applying ecosystem models, data limitations currently prevent the implementation of such models for greater amberjack. This review provides the first comprehensive synthesis of the biology and ecology of the greater amberjack, and offers a solid foundation on which to build either enhanced single species models with ecosystem considerations or fully integrated multi-species or ecosystem-based models.

Current approaches for ecosystem models for marine fish include options such as Atlantis (Audzijonyte et al. 2019) and EcoPath and EcoSim (Keramidas et al. 2023), among others. These models are typically composed of submodels or modules that incorporate human activities, biophysical dynamics, oceanographic variability, and management decisions. To parameterize these models, data are required on abundance-at-age by spatial unit, consumption and production rates, diet composition, individual growth trajectories, maximum age and age-at-maturity, recruitment parameters and variability, habitat preferences, and dispersal or migratory behavior—many of which are covered or referenced in this review. Key spatially explicit biological data, such as habitat-specific growth and maturity, movement patterns, and regional recruitment dynamics, remain poorly understood for greater amberjack. Notably, recruitment is a critical input for both single-species and ecosystem models, yet remains poorly characterized for greater amberjack, with limited empirical data available to estimate variability across regions or years. A key input into ecosystem-based models is trophic ecology and detailed diet composition across life stages and habitats. Greater amberjack diet in the Mediterranean and the Korea Strait in the Pacific has been described in detail, but diet composition for other stocks in lacking. Addressing these gaps should be a priority for future research and must be addressed to support the development of robust Atlantis or EcoSim or EcoPath models and ensure ecologically robust management strategies

Conclusions

Greater amberjack has not historically been considered a species of conservation concern globally, but it is now emerging as a reef-associated fish that may require more specialized management attention. Recent studies suggest that greater amberjack is vulnerable to fishing pressure, despite characteristics typically associated with resilience, such as rapid growth and high fecundity. Connectivity, both between and within stocks, may play a key role in maintaining population stability and resilience. Therefore, future research should prioritize understanding patterns of adult and juvenile connectivity to improve management strategies and predict population responses to fishing pressure and environmental

change. Ongoing efforts to refine abundance estimates, such as the project currently underway in US waters (MASGO-20-0249, Mississippi-Alabama Sea Grant Consortium), offer promising avenues for enhancing data availability, particularly in regions where greater amberjack is an important fisheries resource. This review provides a comprehensive synthesis of available biological and ecological information and highlights emerging research needs, offering a valuable reference for supporting sustainable management of greater amberjack fisheries globally.

Tables

Table 2.1: Known aspects of the reproductive biology of wild greater amberjack (*Seriola dumerili*).

Reference	Region	Size at smallest maturity (FL mm)		Size at 50% maturity (FL mm)		Sex ratio (M:F)	Peak spawning season
		Male	Female	Male	Female		
Harris et al. (2007)	Southeastern US	464	514	644	733	1:1.1	April – May
Murie et al. (2008)	Gulf of Mexico		700	-	900	1:2.3	March – May
Marino et al. (1995)	Mediterranean	610 SL	800 SL	1090 SL	1120 SL	-	May – July
Sley et al. (2014)	Mediterranean	930	814	937	953	1:1.44	June – July
Murie et al. (2020)	Gulf of Mexico	-	710	-	827 – 873	-	March – May
Kozul et al. (2001)	Mediterranean	750 TL	750 TL			1:0.96	May – June

Table 2.2: Known aspects of age and growth of the wild greater amberjack (*Seriola dumerili*) including the number of samples (N), the length range (mm), the age range, the estimated von Bertalanffy growth parameter L∞, and the length-weight equation.

Reference	Region	Ageing structure	N	Length range (mm)	Age range	L∞	Length – weight equation
Burch 1979	Southeastern US	Scales	1643	411 – 1386	1 – 10	1643	W = 6.40 x 10-5 L2.842
Manooch and Potts (1997)	Gulf of Mexico	Sectioned otoliths	340	345 – 1552	0 – 15	1109	W = 2.4 X 10-8 TL2.86 W = 5.3 X 10-8 FL2.81
Manooch and Potts (1997)	Southeastern US	Sectioned otoliths	323	345 – 1598	1 – 17	1648	W= 7.2 × 10–8L2.70
Thompson et al. (1999)	Gulf of Mexico	Sectioned otoliths	597	167 – 1441	0 – 15	1389	W= 4.2 × 10–5L2.83
Harris et al. (2007)	Southeastern US	Sectioned otoliths	1985	233 – 1445	1 – 13	1241.5	
Murie and Parkyn (2008)	Gulf of Mexico	Sectioned otoliths and fin rays	1787	74 – 1829	0 – 15	1240	Log10W= -7.1207 + 2.747 log10FL
Kozul et al. (2001)	Mediterranean	Scales	298	320 – 1600	1 – 10	1746	W = 0.000123 X L2.847
Murie et al. (2020)	Gulf of Mexico	Sectioned otoliths	6801	141 – 1829	0 – 19	1307	

Table 2.3: Known feeding intensity by measure of vacuity of greater amberjack (*Seriola dumerili*). Lengths were recorded in total length (TL), fork length (FL), and standard length (SL).

Reference	Region	Vacuity	Length (mm)
Manooch and Haimovici 1983	Southeastern US	12	397 – 1386 TL
Andaloro and Pipitone 1997	Mediterranean	60	520 - 1480 SL
Pipitone and Andaloro 1995	Mediterranean	32	90 – 1480 SL
Badalamenti et al. 1995	Mediterranean	3.63	30 – 120 SL
Sley et al. 2016	Mediterranean	37.9	155 – 1660 TL
Kim et al. 2014	North Pacific Ocean	35.3	175 – 499 FL

Figures



Figure 2.1: Map showing (colored lines) the geographical distribution of the greater amberjack (*Seriola dumerili*) that includes six main areas that have been identified by genetics: western Atlantic (Gold and Richardson 1998), Gulf of Mexico (Hargrove et al. 2018), eastern Atlantic, Mediterranean Sea(Šegvić-Bubić et al. 2016), south China sea and east China sea (Araki et al. 2018). The grey lines represent areas in which occurrences have been recorded without genetic information.

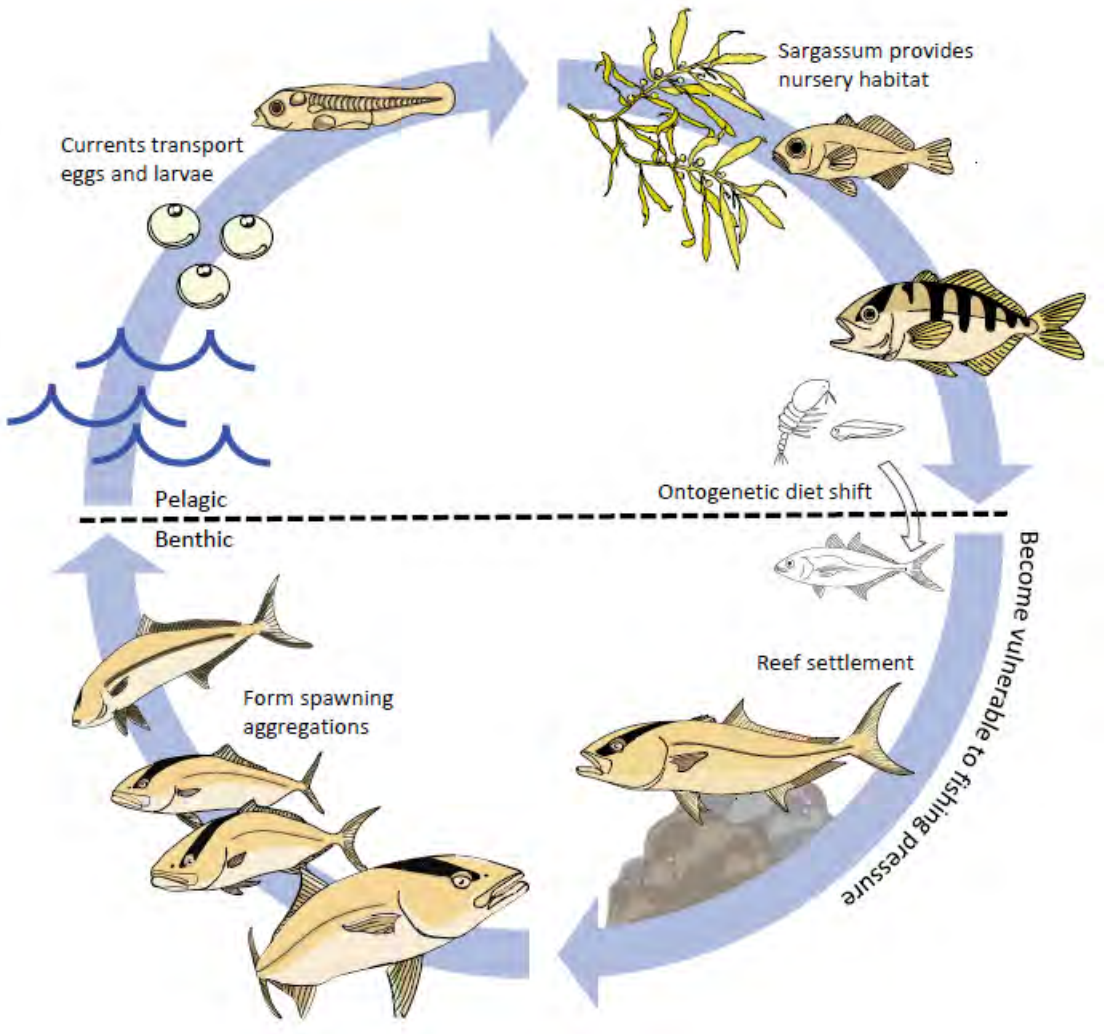


Figure 2.2: Conceptual diagram of the life cycle of the greater amberjack (*Seriola dumerili*). Dashed line separates pelagic and benthic life cycle stages. Illustration by Mischa Schultz.

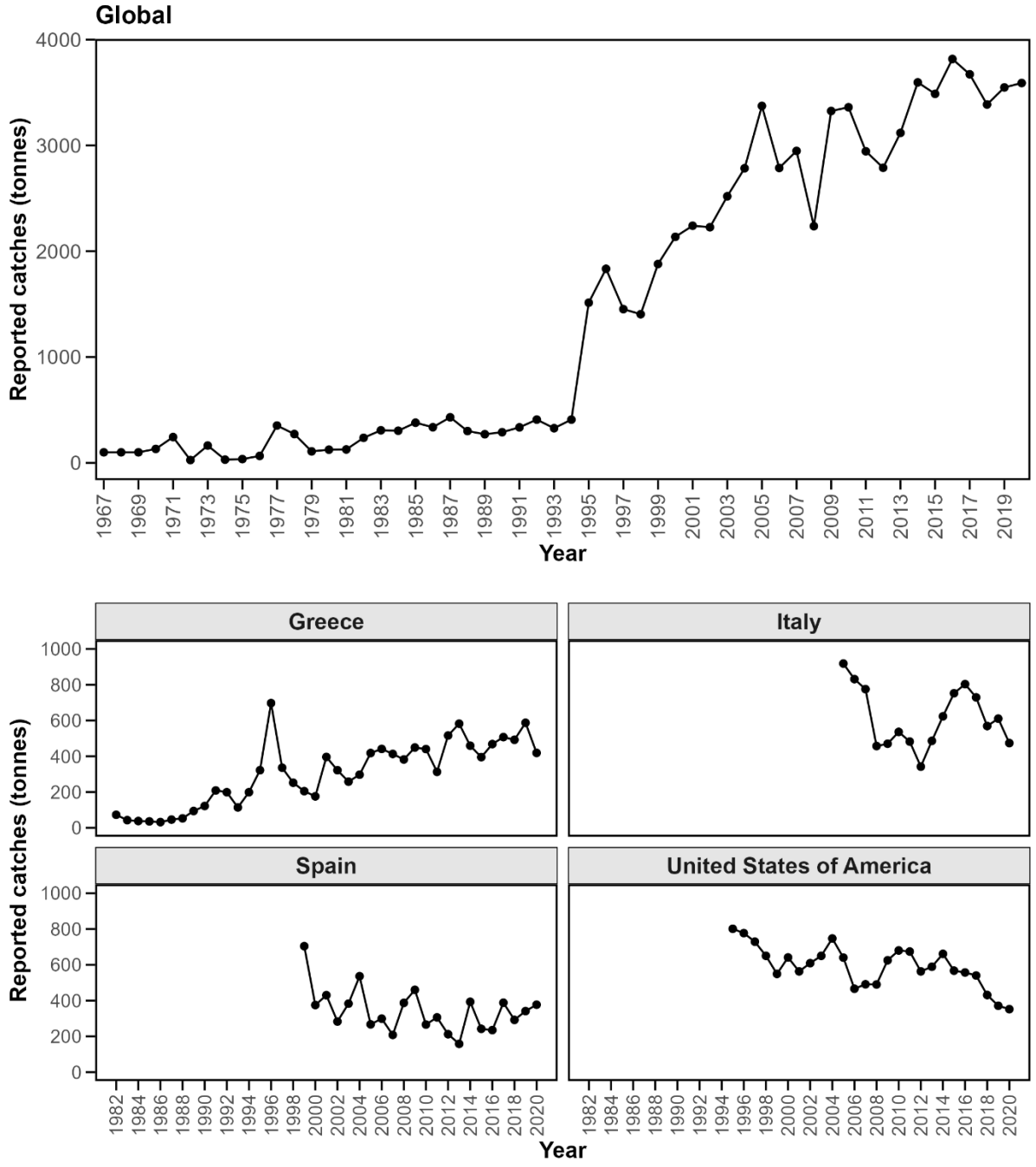


Figure 2.3: Temporal variation reported catch (tonnes) of greater amberjack (*Seriola dumerili*) compiled from 2021 FAO database. Top panel shows global reported catch (1967 – 2020) followed by regional reported catch separated into Mediterranean (1967 – 2020) and United States of America (USA; 1995 – 2020). Bottom panel shows reported catch for top recording countries in the Mediterranean region, Greece (1982 – 2020) and Italy (2005 – 2020).

REFERENCES

Andaloro F, Pipitone C. 1997. Food and feeding habits of the amberjack, *Seriola dumerili* in theCentral Mediterranean Sea during the spawning season. Cahiers De Biologie Marine. 3:91–96.

Aoki R, Chuda H, Washio Y, Masuma S, Kato K. 2019. Sex discrimination of cultured greater amberjack *Seriola dumerili* using steroid hormones. Fisheries Science. 86(1):57–64.

Aoufi S, Valls M, Bienentreu OC, Pereira IR, Huang Y, Mourre B, Cabanellas-Reboredo M. 2023. Preferential aggregation areas of *Seriola dumerili*: Estimation of environmental influences and insights towards sustainable exploitation. Fisheries Research. 266: 106793.

Araki K, Aokic JY, Kawase J, Hamada K, Ozaki A, Fujimoto H, Yamamoto I, Usuki H. 2018. Whole Genome Sequencing of Greater Amberjack (*Seriola dumerili*) for SNP Identification on Aligned Scaffolds and Genome Structural Variation Analysis Using Parallel Resequencing. International Journal of Genomics. 2018:7984292.

Arena PT, Jordan LKB, Spieler RE. 2007. Fish assemblages on sunken vessels and natural reefs in southeast Florida, USA. Hydrobiologia. 580(1):157–171.

Arrington DA, Winemiller KO, Loftus WF, Akin S. 2002. How often do fishes "run on empty"?

Audzijonyte A, Pethybridge H, Porobic J, Gorton R, Kaplan I, Fulton EA, Poisot T. 2019. Atlantis: A spatially explicit end-to-end marine ecosystem model with dynamically integrated physics, ecology and socio-economic modules. Methods in Ecology and Evolution. 10(10):1814–1819.

Auster PJ, Kracker L, Price V, Heupel E, McFall G, Grenda D. 2013. Behavior Webs of Piscivores at Subtropical Live-Bottom Reefs. Bulletin of Marine Science. 89(1):377–396.

Bacheler NM, Gregalis KC, Gillum ZD, Pickett EP, Schobernd CM, Schobernd ZH, Teer BZ. 2022. Using stationary video data to infer relative abundance and distribution of four *Seriola* species along the southeast United States Atlantic coast. Fisheries Research. 249:106238.

Badalamenti F, Anna GD, Mazzola A. 1995. Feeding habits of young-of-the-year greater amberjack *Seriola dumerili* (Risso, 1810) along the N/W Sicilian Coast. Scientia Marina. 59(3-4):317–323.

Baeyaert J, Abecasis D, Afonso P, Graça G, Erzini K, Fontes J. 2018. 'Solo datasets': unexpected behavioural patterns uncovered by acoustic monitoring of single individuals. Marine and Freshwater Behaviour and Physiology. 51(3):183–201.

Benfield MC, Kupchik MJ, Palandro DA, Dupont JM, Blake JA, Winchell P. 2019. Documenting deepwater habitat utilization by fishes and invertebrates associated with Lophelia pertusa on a petroleum platform on the outer continental shelf of the Gulf of Mexico using a remotely operated vehicle. Deep Sea Research Part I: Oceanographic Research Papers. 149:103045.

Berry FH, Burch RK. 1979. Aspects of the amberjack fisheries. Proceedings of the Gulf and Caribbean Fisheries Institute, 31. Gulf and Caribbean Fisheries Institute.

Biggs CR, Heyman WD, Farmer NA, Kobara S, Bolser DG, Robinson J, Lowerre-Barbieri SK, Erisman BE. 2021. The importance of spawning behavior in understanding the vulnerability of exploited marine fishes in the U.S. Gulf of Mexico. PeerJ. 9:e11814.

Boeuf G, Payan P. 2001. How should salinity influence fish growth? Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology. 130(4):411–423.

Bolser DG, Egerton JP, Grüss A, Loughran T, Beyea T, McCain K, Erisman BE. 2020. Environmental and Structural Drivers of Fish Distributions among Petroleum Platforms across the U.S. Gulf of Mexico. Marine and Coastal Fisheries. 12(2):142–163.

Botsford LW, Micheli F, Hastings A. 2003. Principles for the Design of Marine Reserves. Ecological Applications. 13(sp1):25–31.

Boyle KS, Hightower CL, Powers SP. 2022a. Diel patterns of depth use and swimming activity of post-release greater amberjack (*Seriola dumerili*) in the northern Gulf of Mexico. Environmental Biology of Fishes. 106(3):491–518.

Boyle KS, Hightower CL, Powers SP. 2022b. Evaluation of factors contributing to post release mortality of greater amberjack (*Seriola dumerili*) in the northern Gulf of Mexico with depth and acceleration data from acoustic tags. Fishery Bulletin. 120(3-4):218–233.

Burch RK. 1979. The greater amberjack, *Seriola dumerili*: its biology and fishery off southeastern Florida. Coral Gables, FL. p. 112.

Cadrin SX. 2020. Defining spatial structure for fishery stock assessment. Fisheries Research. 221:105397.

Campanella F, Auster PJ, Taylor JC, Munoz RC. 2019. Dynamics of predator-prey habitat use and behavioral interactions over diel periods at sub-tropical reefs. PLoS One. 14(2):e0211886.

Cardoso APLR, Matos MRSBC, Rosa RS, Alvarado F, Medeiros APM, Santos BA. 2020. Increased fish diversity over day and night in structurally complex habitats of artificial reefs. Journal of Experimental Marine Biology and Ecology. 522:151244.

Chagaris D, Sagarese S, Farmer N, Mahmoudi B, de Mutsert K, VanderKooy S, Patterson WF, Kilgour M, Schueller A, Ahrens R et al. 2019. Management challenges are opportunities for fisheries ecosystem models in the Gulf of Mexico. Marine Policy. 101:1–7.

Chen C, Ji R, Huang J, He H, Liao Z. 1997. The relationship between the salinity and the embryonic, early larval development in *Seriola dumerili*. Journal of Shanghai Ocean University. 6(1).

Chen WM, Umeda N, Mitsuboshi T, Hirazawa N. 2007. Circadian self-feeding rhythms in greater amberjack *Seriola dumerili* (Risso). Journal of Fish Biology. 70(2):451–461.

Claisse JT, Pondella DJ, 2nd, Love M, Zahn LA, Williams CM, Williams JP, Bull AS. 2014. Oil platforms off California are among the most productive marine fish habitats globally. Proceedings of the National Academy of Sciences. 111(43):15462–15467.

Colloca F, Crespi V, Cerasi S, Coppola SR. 2004. Structure and evolution of the artisanal fishery in a southern Italian coastal area. Fisheries Research. 69(3):359–369.

Corriero A, Wylie MJ, Nyuji M, Zupa R, Mylonas CC. 2021. Reproduction of greater amberjack (*Seriola dumerili*) and other members of the family Carangidae. Rev Aquacult. 13(4):1781–1815. English.

Crandall CAC, Parkyn DC, Murie DJ. 2013. Regional stock structure of greater amberjack in the southeastern United States using otolith shape analysis. North Charleston, SC.

Dance MA, Patterson III WF, Addis DT. 2011. Fish Community and Trophic Structure at Artificial Reef Sites in the Northeastern Gulf of Mexico. Bulletin of Marine Science. 87(3):301–324.

Dillon KS, Fleming CR, Slife C, Leaf RT. 2021. Stable Isotopic Niche Variability and Overlap across Four Fish Guilds in the North-Central Gulf of Mexico. Marine and Coastal Fisheries. 13(3):213–227.

dos Santos Silva Amaral L, Bastos ASA, de Carvalho-Junior L, Maciel MDR, Teixeira-Neves TP, Araújo FG, Neves LM. 2023. Diel changes in fish assemblages of Southwest Atlantic rocky reefs. Environmental Biology of Fishes. 106(4):627–639.

Duffy HJ, Letessier TB, Irving RA. 2017. Significant range extensions for two fish species at Pitcairn Island, South Pacific. Journal of Fish Biology. 91(2):669–672.

FAO. 2021. FAO FishStatJ: Landings reports from 1950-2020. FAO, Rome.

Feeley MW, Morley D, Acosta A, Barbera P, Hunt J, Switzer T, Burton M. 2018. Spawning migration movements of Mutton Snapper in Tortugas, Florida: Spatial dynamics within a marine reserve network. Fisheries Research. 204:209–223.

Fernández-Montero A, Caballero MJ, Torrecillas S, Tuset VM, Lombarte A, Ginés RR, Izquierdo M, Robaina L, Montero D. 2018. Effect of temperature on growth performance of greater amberjack (*Seriola dumerili* Risso 1810) Juveniles. Aquaculture Research. 49(2):908–918.

Flavio H, Seitz R, Eggleston D, Svendsen JC, Stottrup J. 2023. Hard-bottom habitats support commercially important fish species: a systematic review for the North Atlantic Ocean and Baltic Sea. PeerJ. 11:e14681.

Foss KL. 2016. Feeding Ecology of Red Snapper and Greater Amberjack at Feeding Ecology of Red Snapper and Greater Amberjack at Standing Platforms in the Northern Gulf of Mexico: Disentangling the Effects of Artificial Light the Effects of Artificial Light. Louisiana State University.

Frisk MG, Jordaan A, Miller TJ. 2013. Moving beyond the current paradigm in marine population connectivity: are adults the missing link? Fish and Fisheries. 15(2):242–254.

Galbraith JK, Bemis KE, Bemis WE, Cook HS, Wuenschel MJ. 2022. Identifications, distributions, and life history of four species of *Seriola* (Carangiformes: Carangidae) in the western North Atlantic based on contemporary and historical data.

Gallaway BJ, Raborn SW, McCain KA, Beyea RT, Dufault S, Heyman W, Putman NF, Egerton J. 2021. Absolute Abundance Estimates for Red Snapper, Greater Amberjack, and Other Federally Managed Fish on Offshore Petroleum Platforms in the Gulf of Mexico. North American Journal of Fisheries Management. 41(6):1665–1690.

Gallaway BJ, Szedlmayer ST, Gazey WJ. 2009. A Life History Review for Red Snapper in the Gulf of Mexico with an Evaluation of the Importance of Offshore Petroleum Platforms and Other Artificial Reefs. Reviews in Fisheries Science. 17(1):48–67.

Garner SB, Boswell KM, Lewis JP, Tarnecki JH, Patterson WF. 2019. Effect of reef morphology and depth on fish community and trophic structure in the northcentral Gulf of Mexico. Estuarine, Coastal and Shelf Science. 230:106423.

Gold JR, Richardson LR. 1998. Population structure in greater amberjack, *Seriola dumerili*, from the Gulf of Mexico and the western Atlantic Ocean. Fishery Bulletin. 96(4):767–778.

Graham RT, Castellanos DW. 2005. Courtship and spawning behaviors of carangid species in Belize. Fisheries Bulletin. 103:426–432.

Grüss A, Biggs C, Heyman WD, Erisman B. 2018. Prioritizing monitoring and conservation efforts for fish spawning aggregations in the U.S. Gulf of Mexico. Scientific Reports. 8(1):8473.

Hamwi NI, Ali-Basha N, Altajer H, Salem J. 2024. Population assessment of greater amberjack *Seriola dumerili* along the Syrian waters in the eastern Mediterranean Sea using expert systems. Journal of Marine Science and Research. 3(2).

Hargrove JS, Murie DJ, Parkyn DC, Saarinen EV, Austin JD. 2018. Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico. Genetica. 146(4-5):393–402.

Harris PJ, Wyanski DM, White DB, Mikell PP, Eyo PB. 2011. Age, Growth, and Reproduction of Greater Amberjack off the Southeastern U.S. Atlantic Coast. Transactions of the American Fisheries Society. 136(6):1534–1545.

Hasegawa T, Lu C-P, Hsiao S-T, Uchino T, Yeh H-M, Chiang W-C, Chen J-R, Sassa C, Komeyama K, Kawabe R et al. 2020. Distribution and genetic variability of young-of the-year greater amberjack (*Seriola dumerili*) in the East China Sea. Environmental Biology of Fishes. 103(7):833–846.

Hasegawa T, Takatsuki N, Kawabata Y, Kawabe R, Nishihara GN, Ishimatsu A, Soyano K, Okamura K, Furukawa S, Yamada M et al. 2017. Continuous behavioral observation reveals the function of drifting seaweeds for *Seriola* spp. juveniles. Marine Ecology Progress Series. 573:101–115.

Heyman WD, Kjerfve B. 2008. Characterization of Transient Multi-species Reef Fish Spawning Aggregations at Gladden Spit, Belize. Bulletin of Marine Science. 83(3):531–551.

Honryo T, Sakurai Y, Yamao S, Okada T, Agawa Y, Sawada Y. 2021. The problem of marine litters for cultured teleost. Marine Pollution Bulletin. 170:112679.

Humphreys RL Jr. 1980. Feeding habits of the Kahala, *Seriola dumerili*, in the Hawaiian archipelago. Proceedings of the Symposium on Status of Resource Investigations in the Northwestern Hawaiian Islands

Grigg RW, Tanoue KY, editors. Ciguatera and the feeding habits of the greater amberjack, *Seriola dumerili*, in the Hawaiian Archipelago. Proceedings of the Symposium on Resource Investigations in the Northwestern Hawaiian Islands; 1984; University of Hawaii, Honolulu, Hawaii. UNIHI-SEAGRANT-MR-84-01.

Ibrahim MA, El-Bary KA, Al-Khayat JA. 1989. By-catch of Commercial Bottom Trawl Fishery from Qatar Waters, Arabian Gulf. Qatar University Science Bulletin. 9:309–319.

Ingram GW, Patterson WF III. 2001. Movement pattems of Red Snapper (Lutjanus campechanus), Greater Amberjack (*Seriola dumerili*), and Gray Triggerfish (Balistes capriscus) in the Gulf of Mexico and the utility of Marine Reserves as management tools. Proceedings of the 52nd Gulf and Caribbean Fisheries Institute.

Jackson LS, Drymon JM, Nelson TR, Powers SP. 2018. Biotelemetry based estimates of greater amberjack (*Seriola dumerili*) post-release mortality in the northern Gulf of Mexico. Fisheries Research. 208:239–246.

Jerez S, Samper M, Santamaría FJ, Villamandos JE, Cejas JR, Felipe BC. 2006. Natural spawning of greater amberjack (*Seriola dumerili*) kept in captivity in the Canary Islands. Aquaculture. 252(2-4):199–207.

Kasai A, Sakamoto W, Mitsunaga Y, Yamamoto S. 2001. Behaviour of immature yellowtails (*Seriola* quinqueradiata) observed by electronic data-recording tags. Fisheries Oceanography. 9(3):259–270.

Kendall NW, Dieckmann U, Heino M, Punt AE, Quinn TP. 2014. Evolution of age and length at maturation of Alaskan salmon under size-selective harvest. Evolutionary Applications. 7(2):313–322.

Keramidas I, Dimarchopoulou D, Ofir E, Scotti M, Tsikliras AC, Gal G. 2023. Ecotrophic perspective in fisheries management: a review of Ecopath with Ecosim models in European marine ecosystems. Frontiers in Marine Science. 10:1182921.

Kikkawa BS, Everson AR. 1984. Gonadal Maturation, Fecundity, and Spawning of the Greater Amberjack, *Seriola dumerili* (Risso), in Hawaiian Waters with References to Ciguatoxin Incidences. Proceedings of the second Symposium on Resource Investigations in the Northwestern Hawaiian Islands; Honolulu, Hawaii.

Kim DG, Soh HY, Baeck GW. 2023. Feeding Habits and Trophic Level of Juvenile Greater Amberjack, *Seriola* dumereili (Carangidae) in the Southern Waters of Korea. Journal of Ichthyology. 64(1):117–126.

Kozul V, Skaramuca B, Kraljevicâ M, Dulcï Icâ J, Glamuzina B. 2008. Age, growth and mortality of the Mediterranean amberjack *Seriola dumerili* (Risso 1810) from the south eastern Adriatic Sea. Journal of Applied Ichthyology. 17(3):134–141.

Lowerre-Barbieri SK, Catalán IA, Frugård Opdal A, Jørgensen C. 2019. Preparing for the future: integrating spatial ecology into ecosystem-based management. ICES Journal of Marine Science. 76(2):467–476.

Lowerre-Barbieri SK, Ganias K, Saborido-Rey F, Murua H, Hunter JR. 2011. Reproductive Timing in Marine Fishes: Variability, Temporal Scales, and Methods. Marine and Coastal Fisheries. 3(1):71–91.

Lucchetti A, Carbonara P, Colloca F, Lanteri L, Spedicato MT, Sartor P. 2017. Small-scale driftnets in the Mediterranean: Technical features, legal constraints and management options for the reduction of protected species bycatch. Ocean & Coastal Management. 135:43–55.

Luzhniak VA, Chepurnaia TA, Zhivoglyadov AA. 2020. First Discovery of Greater Amberjack *Seriola dumerili* (Carangidae) in Russian Waters of the Black Sea Coast of the Caucasus. Journal of Ichthyology. 60(2):335–338.

Mammel M, Lee M-A, Wang Y-C, Ghosh A, Chen Y-K. 2023. Spatial Variations and Size Dependent Feeding Strategies of the Greater Amberjack in Association of Environmental Factors in Taiwanese Waters.

Mammel M, Lee M-A, Wang Y-C, Ghosh A, Chen Y-K, Matilda MV, Navus SG. 2024. Unraveling the feeding strategies of the greater amberjack: insights into size-dependent dietary patterns and environmental influences in Taiwanese waters. Fisheries Science. 90(6):931–952.

Mammel M, Naimullah M, Vayghan AH, Hsu J, Lee M-A, Wu J-H, Wang Y-C, Lan K-W. 2022. Variability in the Spatiotemporal Distribution Patterns of Greater Amberjack in Response to Environmental Factors in the Taiwan Strait Using Remote Sensing Data. Remote Sensing. 14(12):2932.

Manooch CS, III, Haimovici M. 1983. Foods of Greater Amberjack, *Seriola dumerili*, and Almaco Jack, *Seriola rivoliana* (Pisces:Carangidae), from the South Atlantic Bight. Journal of the Elisha Mitchell Scientific Society. 99(1):1–9.

Manooch CS, III, Potts JC. 1997. Age, growth, and mortality of greater amberjack, *Seriola dumerili*, from the U.S. Gulf of Mexico headboat fishery. Bulletin of Marine Science. 61(3):671–683.

Marino G, Mandich A, Massari A, Andaloro F, Porrello S, Finoia MG, Cevasco F. 1995. Aspects of reproductive biology of the Mediterranean amberjack (*Seriola dumerili*i Risso) during the spawning period. Journal of Applied Ichthyology. 11:9–24.

Massari A, Bottero S, Pizzicori P, Goos H, Marino G, Mandich A. 2004. Plasma sex steroid and vitellogenin profiles during gonad development in wild Mediterranean amberjack (*Seriola dumerili*i). Marine Biology. 144(1):127–138.

Masuma S, Kanematu M, Teruya K. 1990. Embryonic and Morphological Development of Larvae and Juveniles of the Amberjack, *Seriola dumerili*. Japanese Journal of Ichthyology. 37:164–169.

McClellan DB, Cummings NJ, 1997. Preliminary analysis of tag recapture data of Greater Amberjack, *Seriola dumerili*, in the southeastern United States. Proceedings of the Gulf nd Caribbean Fisheries Institute, 49.

Micale V, Maricchiolo G, Genovese L. 1999. The reproductive biology of the amberjack *Seriola dumerili*i Risso 1810. I. Oocyte development in captivity. Aquaculture Research. 30(5), 349–355.

Mohan JA, Sutton TT, Cook AB, Boswell KM, David Wells RJ. 2017. Influence of oceanographic conditions on abundance and distribution of post-larval and juvenile carangid fishes in the northern Gulf of Mexico. Fisheries Oceanography. 26(5):526–541.

Morales NA, Coghlan AR, Easton EE, Friedlander AM, Herlan J, Gaymer CF. 2022. Now you see me: “first” records of the greater amberjack *Seriola dumerili* at Rapa Nui range extension or increased scientific effort? Journal of Fish Biology. 100(3):835–842.

Munnelly RT, Reeves DB, Chesney EJ, Baltz DM. 2020. Spatial and Temporal Influences of Nearshore Hydrography on Fish Assemblages Associated with Energy Platforms in the Northern Gulf of Mexico. Estuaries and Coasts. 44(1):269–285.

Murie D, Parkyn D, Austin J. 2011. Seasonal Movement and Mixing Rates of Greater Amberjack in the Gulf of Mexico and Assessment of Exchange with the South Atlantic Spawning Stock. SEDAR, North Charleston, SC.

Murie D, Parkyn DC. 2008. Age, Growth and Sex Maturity of Greater Amberjack (*Seriola dumerili*) in the Gulf of Mexico. North Charleston, SC.

Murie D, Parkyn DC, Fischer A. 2020. Is the BOFFF (Big Old Fat Fecund Females) hypothesis applicable to Gulf of Mexico greater amberjack? SEDAR70-RD-02, North Charleston, SC.

Murua H, Saborido-Rey F. 2003. Female Reproductive Strategies of Marine Fish Species of the North Atlantic. Journal of Northwest Atlantic Fishery Science. 33:23–31.

Nakada Mk. 2010. Yellowtail Culture Development and Solutions for the Future. Reviews in Fisheries Science. 10(3-4):559–575.

Nugroho E, Taniguchi N, Kato K, Miyashita S. 2000. Genetic Difference among Seed Populations of Greater Amberjack Used in Aquaculture Farm of Japan. Aquaculture Science. 48(4):665–674.

Ovando D, Hilborn R, Monnahan C, Rudd M, Sharma R, Thorson JT, Rousseau Y, Ye Y. 2021. Improving estimates of the state of global fisheries depends on better data. Fish and Fisheries. 22(6):1377–1391.

Palmer JL, Armstrong C, Akbora HD, Beton D, Çağlar Ç, Godley BJ, Metcalfe K, Özkan M, Snape RTE, Broderick AC. 2024. Behavioural patterns, spatial utilisation and landings composition of a small-scale fishery in the eastern Mediterranean. Fisheries Research. 269:106861.

Papadaki M, Mandalakis M, Anastasiou TI, Pouli M, Asderis M, Katharios P, Papandroulakis N, Mylonas CC. 2021. Histological evaluation of sex differentiation and early sex identification in hatchery-produced greater amberjack (*Seriola dumerili*) reared in sea cages. Fish Physiology and Biochemistry. 47(6):1777–1792.

Papandroulakis N, Mylonas CC, Maingot E, Divanach P. 2005. First results of greater amberjack (*Seriola dumerili*) larval rearing in mesocosm. Aquaculture. 250(1-2):155–161.

Parker JRO, Mays RW. 1998. Southeastern US deepwater reef fish assemblages habitat characterisics and life history summaries. NOAA Technical Reports.

Pavlov D, Kasumyan AO. 2000. Patterns and mechanisms of schooling behavior in fish: A review.

Paxton AB, Newton EA, Adler AM, Van Hoeck RV, Iversen ES, Jr., Taylor JC, Peterson CH, Silliman BR. 2020. Artificial habitats host elevated densities of large reef-associated predators. PLoS One. 15(9):e0237374.

Pereira IR, Valls M, Aoufi S, Bienentreu OC, Huang Y, Cabanellas-Reboredo M. 2023. Socio Ecological Overview of the Greater Amberjack Fishery in the Balearic Islands. Fishes. 8(4):184–184.

Pipitone C, Andaloro F. 1995. Food and feeding habits of juvenile greater amberjack, *Seriola dumerili* (Osteichthyes, Carangidae) in inshore waters of the central Mediterranean Sea. Cybium: International Journal of Ichthyology. 19(3):305–310.

Rajesh KM, Thomas S, Sukumaran S, Purushottama GB, Nataraja GD, Dineshbabu AP, Rohit P, Zacharia PU. 2021. New Distributional Record of Greater Amberjack *Seriola dumerili* (Risso, 1810) (Perciformes: Carrangidae) from the Eastern Arabian Sea, India. Indian Journal of Animal Research. 56:243–248.

Raya V, Sabatés A. 2015. Diversity and distribution of early life stages of carangid fishes in the northwestern Mediterranean: responses to environmental drivers. Fisheries Oceanography. 24(2):118–134.

Reed JK, Pomponi SA, Weaver D, Paull CK, Wright AE. 2005. Deep-water sinkholes and bioherms of South Florida and the Pourtales terrace-habitat and fauna. Bulletin of Marine Science. 77(2):267–296.

Robichaud D, Rose GA. 2004. Migratory behaviour and range in Atlantic cod: Inference from a century of tagging. Fish and Fisheries. 185–214.

Rodriguez-Marin E, Busawon D, Luque PL, Castillo I, Stewart N, Krusic-Golub K, Parejo A, Hanke A. 2022. Timing of Increment Formation in Atlantic Bluefin Tuna (Thunnus thynnus) Otoliths. Fishes. 7(5).

Rooker JR, Kitchens LL, Dance MA, Wells RJ, Falterman B, Cornic M. 2013. Spatial, temporal, and habitat-related variation in abundance of pelagic fishes in the Gulf of Mexico: potential implications of the Deepwater Horizon oil spill. PLoS One. 8(10):e76080.

Rooker JR, Turner JP, Holt SA. 2006. Trophic ecology of Sargussum-associated fishes in the Gulf of Mexico determined from stable isotopes and fatty acids. Marine Ecology Progress Series. 313:249–259.

Sakakura Y, Tsukamoto K. 1999. Ontogeny of aggressive behaviour in schools of yellowtail, *Seriola* quinqueradiata. Environmental Biology of Fishes. 56:231–242.

Sanchez-Cueto P, Stavrakidis-Zachou O, Clos-Garcia M, Bosch M, Papandroulakis N, Llado S. 2023. Mediterranean Sea heatwaves jeopardize greater amberjack's (*Seriola dumerili*) aquaculture productivity through impacts on the fish microbiota. ISME Communication. 3(1):36.

Sanderson SL, Chesnutt CR, Lobe PS. 1996. Evidence for ram suspension feeding by the piscivore, *Seriola dumerili* (Carangidae). Environmental Biology of Fishes. 46:365–373.

Secor DH. 2015. Migration Ecology of Marine Fishes. Johns Hopkins University Press.

Secor DH, Kerr LA, Cadrin SX. 2009. Connectivity effects on productivity, stability, and persistence in a herring metapopulation model. ICES Journal of Marine Science. 66(8):1726–1732.

Secor DH, O'Brien MHP, Gahagan BI, Fox DA, Higgs AL, Best JE. 2020. Multiple spawning run contingents and population consequences in migratory striped bass Morone saxatilis. PLoS One. 15(11):e0242797.

SEDAR. 2020. SEDAR 70 – Gulf of Mexico Greater Amberjack Stock Assessment Report.

Sedberry GR, Pashuk, O, Wyanski DM, Stephen JA, Weinbach P. 2006. Spawning locations for Atlantic reef fishes off the southeastern U.S. Item. Proceedings of the Gulf and Caribbean Fisheries Institute, 57; 2006; Charleston, South Carolina.

Šegvić-Bubić T, Marrone F, Grubišić L, Izquierdo-Gomez D, Katavić I, Arculeo M, Lo Brutto S. 2016. Two seas, two lineages: How genetic diversity is structured in Atlantic and Mediterranean greater amberjack *Seriola dumerili* Risso, 1810 (Perciformes, Carangidae). Fisheries Research. 179:271–279.

Seminara CI, Souza GBG, Zapelini C, Barbosa-Filho MLV, Profice CC. 2024. Opportunities and challenges for small-scale fisheries co-management: A case study of the unofficial landing records system in a Brazilian fishing community. Marine Policy. 160:105991.

Sicuro B, Luzzana U. 2016. The State of *Seriola* spp. Other Than Yellowtail (S. quinqueradiata) Farming in the World. Reviews in Fisheries Science and Aquaculture. 314–325.

Sinopoli M, D'Anna G, Badalamenti F, Andaloro F. 2006. FADs influence on settlement and dispersal of the young-of-the-year greater amberjack (*Seriola dumerili*). Marine Biology. 150(5):985–991.

Sley A, Taieb A, Jarboui O, Ghorbel M, Bouain A. 2014. Reproductive biology of greater amberjack *Seriola dumerili* (Risso, 1810) from the Eastern Mediterranean Sea (Tunisia, Gulf of Gabes). Cahiers de Biologie Marine. 55:421–430.

Sley A, Taieb AH, Jarboui O, Ghorbel M, Bouain A. 2016. Feeding behaviour of greater amberjack *Seriola dumerili* (Risso, 1810) from Central Mediterranean (Gulf of Gabes, Tunisia). Journal of the Marine Biological Association of the United Kingdom. 96(6):1229–1234.

Smith GH, Murie DJ, Parkyn DC. 2014. Nonlethal Sex Determination of the Greater Amberjack, with Direct Application to Sex Ratio Analysis of the Gulf of Mexico Stock. Marine and Coastal Fisheries. 6(1):200–210.

Smith GH, Murie DJ, Parkyn DC. 2018. Effects of sex-specific fishing mortality on sex ratio and population dynamics of Gulf of Mexico greater amberjack. Fisheries Research. 208:219– 228.

Stuart-Smith J, Pecl G, Pender A, Tracey S, Villanueva C, Smith-Vaniz WF. 2016. Southernmost records of two *Seriola* species in an Australian ocean-warming hotspot. Marine Biodiversity. 48(3):1579–1582.

Swart BL, von der Heyden S, Bester-van der Merwe A, Roodt-Wilding R. 2015. Molecular systematics and biogeography of the circumglobally distributed genus *Seriola* (Pisces: Carangidae). Molecular Phylogenetics and Evolution. 93:274–280.

Tara Marshall C. 2016. Implementing Information on Stock Reproductive Potential in Fisheries Management: The Motivation, Challenges and Opportunities. Fish reproductive biology: Implications for assessment and management. Wiley, p.438–464.

Tharp RM, Hostetter NJ, Paxton AB, Taylor JC, Buckel JA. 2024. Artificial structure selection by economically important reef fishes at North Carolina artificial reefs. Frontiers in Marine Science. 11:1373494.

Thompson BA, Beasley M, Wilson CA. 1999. Age, distribution, and growth of greater amberjack, *Seriola dumerili*, from the north central Gulf of Mexico. Fishery Bulletin. 97:362–371.

Thorson James T, Cope Jason M, Branch Trevor A, Jensen Olaf P, Walters CJ. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences. 69(9):1556–1568.

Thorson JT, Hermann AJ, Siwicke K, Zimmermann M, Grand M. 2021. Grand challenge for habitat science: stage-structured responses, nonlocal drivers, and mechanistic associations among habitat variables affecting fishery productivity. ICES Journal of Marine Science. 78:1956–1968.

Tone K, Chiang WC, Yeh HM, Hsiao ST, Li CH, Komeyama K, Kudo K, Hasegawa T, Sakamoto T, Nakamura I et al. 2022. Two-way habitat use between reefs and open ocean in adult greater amberjack: evidence from biologging data. Marine Ecology Progress Series. 699:135–151.

Tone K, Nakamura Y, Chiang WC, Yeh HM, Hsiao ST, Li CH, Komeyama K, Tomisaki M, Hasegawa T, Sakamoto T et al. 2021. Migration and spawning behavior of the greater amberjack *Seriola dumerili* in eastern Taiwan. Fisheries Oceanography. 31(1):1–18.

Tracey SR, Wolfe BW, Hartmann K, Pepperell J, Williams SM. 2023. Movement behavior of swordfish provisions connectivity between the temperate and tropical southwest Pacific Ocean. Scientific Reports. 13(1):11812.

Ulman A, Çiçek A, Salihoglu I, Petrou A, Patsalidou M, Pauly D, Zeller D. 2013. The reconstruction and unification of Cyprus's marine fisheries catch data, 1950–2010. Vancouver, BC.

Valls A, Gascuel D, Guénette S, Francour P. 2012. Modeling trophic interactions to assess the effects of a marine protected area: case study in the NW Mediterranean Sea. Marine Ecology Progress Series. 456:201–214.

Véron M, Duhamel E, Bertignac M, Pawlowski L, Huret M, Baulier L. 2020. Determinism of Temporal Variability in Size at Maturation of Sardine Sardina pilchardus in the Bay of Biscay. Frontiers in Marine Science. 7:567841.

Von Bertalanffy L. 1938. A Quantitative Theory of Organic Growth (Inquiries on Growth Laws II). Human Biology. 10(2):181–213.

Wells RJD, Rooker JR. 2004. Distribution, age, and growth of young-of-the-year greater amberjack (*Seriola dumerili*) associated with pelagic Sargassum. Fisheries Bulletin. 102:545–554.

Zhang Y, Hu C, McGillicuddy DJ, Jr., Barnes BB, Liu Y, Kourafalou VH, Zhang S, Hernandez FJ. 2024. Pelagic Sargassum in the Gulf of Mexico driven by ocean currents and eddies. Harmful Algae. 132:102566.

Zupa R, Rodriguez C, Mylonas CC, Rosenfeld H, Fakriadis I, Papadaki M, Perez JA, Pousis C, Basilone G, Corriero A. 2017. Comparative Study of Reproductive Development in Wild and Captive-Reared Greater Amberjack *Seriola dumerili* (Risso, 1810). PLoS One. 12(1):e0169645.

3. Population genomics

Tami Hildahl, Brooke Doucet, Eric Saillant

3.1 Introduction/rationale

In this project component, genome scans generated by the double-digest Restriction Site Associated DNA (dd-RAD) sequencing method were used to infer genetic stock structure and connectivity of Greater Amberjack (GAJ) in US waters. Available information on population genetic structure in the species is limited to a study of mitochondrial DNA variation by Gold & Richardson (1998) and a more recent study by Hargrove et al. (2018) based on 11 microsatellites. Divergence among geographic populations was very weak in both studies. Gold & Richardson (1998) found some weak support for divergence between Atlantic and Gulf of America stocks when the Florida Keys samples were grouped with those from the East coast suggesting that an Atlantic genetic stock that included the Florida Keys area was distinct from the Gulf stock. Hargrove et al. (2018) sampled the Florida Keys spawning population (assumed to represent the Atlantic breeding stock) and three Gulf locations. The distinction between the Gulf and Keys spawning stocks based on the analysis of spawning adults was weakly supported. A two-stock model for the sampled range was proposed, but without clear geographic delineation. It was suggested that GAJ may form a complex metapopulation with extended geographic mixing. These two studies were limited by incomplete geographic sampling (Hargrove et al., 2018 did not sample the western Gulf or the East Coast) and small numbers of markers which prevented assessment of divergence due to natural selection and limited the power of inferences on demographic parameters. This study aimed to address limitations of the two previous studies by using genome scans involving thousands of Single Nucleotide Polymorphism (SNP) markers and a comprehensive sampling from the Carolinas to the Western Gulf of America in conjunction with the acoustic telemetry tagging component of this project. In a first project component, the project team developed a draft reference genome assembly for West Atlantic greater amberjack that was used to map sequencing reads and locate genetic markers during the population genomics study. This report then presents the data obtained by processing 810 greater amberjack specimens from US geographic populations collected from North Carolina to South Texas, one voucher S. *zonata*, three voucher S. *rivoliana*, and 10 specimens received from the Canary Islands.

3.2 Materials and Methods

3.2.1 Draft reference genome

3.2.1.1 Sample preparation and sequencing

Genome sequencing employed the DNA from one specimen of GAJ obtained offshore the Alabama coast by the Fisheries Ecology Laboratory of the University of South Alabama Dauphin Island Sea Lab. DNA was extracted using the MagBind® Blood and Tissue DNA HDQ 96 Kit (Omega BioTek, cat. M6399-01) and sequenced on two PACBIO Sequel II 8M SMRT cells using the HiFi protocol. The same DNA was sequenced on the Illumina Novaseq 6000 platform to generate short (150 bp) paired end reads. This hybrid sequencing strategy leverages the strengths of SMRT and Illumina technologies for a robust hybrid assembly (Salzberg et al., 2012; Haghshenas et al., 2020).

3.2.1.2 Assembly

Raw long reads were assembled in the Flye v. 2.9 pipeline (Kolmogorov et al., 2019). Errors in the resulting assembly were corrected using PACBIO long reads during two iterations of the Inspector v. 1.0.2 pipeline (Chen et al., 2021). Illumina reads were trimmed to remove adaptor sequences, base calls with a quality score below 20, and resulting sequences shorter than 50 base pairs in fastp (Chen et al., 2018). The long-read corrected contigs were polished with trimmed short reads in POLCA (Zimin & Salzberg 2020) followed by NextPolish (Hu et al., 2020). Redundans (Pryszcz & Gabaldón 2016) identified and removed redundant contigs, scaffolded contigs with long reads and closed the gaps after scaffolding. The scaffolded assembly was polished twice with long reads in Racon (Vaser et al., 2017). Overlap files between the assembly and long reads were generated in (Md et al., 2019). A final short-read polishing step was performed in Pilon v. 1.24 (Walker et al., 2014). The Racon-polished assembly was divided into two segments that were run separately in Pilon. The individual segments were concatenated upon completion. The assembly was checked for bacterial contamination in FCS-GX and for residual adapter sequences and FCS-adaptor (Astashyn et al., 2024). No bacterial contamination was identified, and residual adapter sequences were removed with the "clean genome" function in fcs.py.

Quality and contiguity of the assembly were evaluated using QUAST (Gurevich et al., 2013). The assembly metrics generated include N50, the largest length in base pairs such that 50% of the total assembly length is contained in contigs of size at least N50, and L50, the minimum number of contigs necessary to contain 50% of the assembly length. Genome completeness was evaluated based on a core set of 3,640 conserved genes in the actinopterygii database using the Benchmarking Universal Single Copy Orthologs (BUSCO, Simão et al., 2015) software.

3.2.1.3 Linkage mapping

A linkage map of the *Seriola dumerili* genome was developed to anchor genome contigs obtained during assembly above in linkage groups and generate a chromosome-level reference. The map was obtained by analyzing the segregation of Single Nucleotide Polymorphism (SNP) markers in a full sib family of GAJ.

3.2.1.3.1 Production of a full sib mapping family

This project component was conducted with the collaboration of the Hellenic Center for Marine Research (HCMR) who are routinely spawning and culturing GAJ. A single pair cross was performed in aquaculture facilities of the HCMR. Larvae from the cross were incubated for 3 days and over 300 specimens were fixed in 95% ethanol. Tissue samples from the two parents were also preserved. Samples were transferred to USM and dd-RAD sequencing libraries for 276 larvae and the two parents were prepared and sequenced.

3.2.1.3.2 Reduced representation library preparation and sequencing

DNA from larvae and parents were extracted with MagBind® Blood and Tissue DNA HDQ 96 Kit (Omega BioTek, cat. M6399-01) and amplified with Qiagen's REPLI-g whole genome amplification kit following the manufacturer's instructions. The obtained preparations were purified using ethanol precipitation. DNAs were assessed using a Qubit fluorometer and agarose electrophoresis to ensure quantity and quality met the requirements to prepare ddRAD-seq libraries.

ddRAD Library preparation followed protocols developed by Peterson et al. (2012) as implemented in our laboratory to genotype other marine species (see Norrell et al., 2020 with modifications described in Dimens et al., 2023). Briefly, DNA from individual GAJ was digested with MspI and EcoRI restriction

enzymes. Adapters were ligated to the digested DNA. Adapters feature a unique 6bp barcode to each GAJ sample which allows pooling (barcoded) libraries from hundreds of samples for multiplex sequencing and retrieving (demultiplexing) sequencing reads from each individual during data processing; adapters also include an 8bp Unique Molecular Identifier that is used to remove duplicates that arise during the Polymerase Chain Reaction (PCR) amplification of library fragments.

MspI-EcoRI ligated fragments were amplified by PCR and the obtained libraries were pooled and size-selected on a Pippin Prep system to retain fragments ranging in length between 300 and 500bp. The ddRAD library pool (all individuals) was sequenced on an Illumina NovaSEQ X genome sequencer to yield a target 1 to 5M paired-end (2x150bp) reads per individual.

3.2.1.3.3 Sequencing reads processing and SNP Discovery

Barcodes were used to demultiplex reads and assign raw paired reads to individual samples. The reads were quality trimmed to retain sequences with a quality-score above 20 with Trimmomatic, and PCR duplicates identified as identical sequences sharing UMIs, were removed with fastp. BWA-MEM (Li & Durbin, 2010) aligned the resulting filtered reads to the draft genome assembly produced as described above, and variants were called with the local haplotype variant caller Freebayes (Garrison & Marth, 2012) in the dDocent pipeline (Puritz et al., 2014).

The resulting VCF file was screened with VCFtools (Danecek et al., 2011) to remove markers with more than 50% missing data, minor allele counts less than 3 and genotype quality scores below 30. The data were then filtered to retain markers with a minimum depth of 4. This step was followed by the removal of markers with minor allele frequencies below 0.001 and a progressive filtering strategy that consisted of the following: removal of SNPs with less than 30% data, individuals missing >97.5% data, SNPs with less than 40% data and individuals missing >89.1% data. Sites exhibiting extreme allelic balance and improperly paired reads were filtered via dDocent_filters (https://github.com/jpuritz/dDocent/blob/master/scripts/dDocent_filters). Additionally, sites potentially representing repetitive DNA, indicated by coverage of a mean depth greater than twice the standard deviation of the mean site depth for the entire dataset were removed. Multi-nucleotide polymorphisms (MNPs) were then decomposed into multiple SNPs, indels and multiallelic SNPs were then removed to retain only biallelic sites. The dataset was further progressively filtered to retain only individuals missing less than 20% of data and sites with less than 10% missing data. Kinship analyses (see below) revealed two additional individuals showing extreme similarity (indicating duplication) that were removed from the dataset.

3.2.1.3.4 Linkage Map construction

Linkage analysis was performed in the LepMap3 software (Rastas et al., 2013; Rastas, 2017) as implemented in the LepWrap Snakemake workflow (Dimens, 2022). The ParentCall2 module generated an input pedigree file and imputed missing or erroneous parental genotypes. Filtering2 then removed non-informative markers and loci showing segregation distortion. The SeparateChromosomes2 module was used to assign filtered loci to linkage groups. Linkage group assignment was based on a critical LOD score (Morton, 1995) determined by running SeparateChromosomes2 assuming critical LOD scores ranging from 1 to 30; The LOD score retained was the uppermost value that segregated markers into 24 major linkage groups (LOD=20). Unassigned loci were then assigned to linkage groups in the JoinSingles module using parameter settings of lodLimit = 10 and lodDifference = 5. Ordering was completed in the

OrderMarkers2 module and the Kosambi's distance between markers (Kosambi, 1943) was calculated. The ordering was repeated 100 times for each linkage group and the linkage maps with the highest likelihood were retained for each LG. Markers inflating the ends of each linkage group by more than 5 cM were removed. Fifty additional iterations of OrderMarkers2 were then performed following pruning and the obtained LG were evaluated using evaluateOrder to calculate map distances and improveOrder to improve map ordering. The package Red v. 2.0 (Girgis, 2015) was used to mask repeats present in the draft assembly, then an input chain file for LepAnchor software was generated in LASTZ v. 1.02. Linkage information was then used to order and orient contigs of the genome assembly in LepAnchor (Rastas, 2019). Input distances between markers and their physical positions were provided by LepMap3. PACBIO long reads were mapped to the draft genome assembly using Minimap2 v. 2.24 and the obtained PAF file was input in LepAnchor along with the linkage map and the assembly to support anchoring. Three iterations of the PlaceAndOrientContigs procedure of Lep-Anchor were performed to produce the final anchored genome.

3.2.1.4 Hi-C anchoring

3.2.1.4.1 Hi-C sequencing

Liver tissue was collected from a single greater amberjack specimen in the northern US Gulf and stored in 70% ethanol prior to sample preparation. High quality DNA was extracted and the sample was processed using Phase Genomics' Proximo Hi-C Kit (Animal) following instructions from the manufacturer. The paired-end Hi-C library was sequenced at the Johns Hopkins Genetic Resource Core Facility on an Illumina Novaseq X sequencer to produce 444,087,944 reads.

3.2.1.4.2 Hi-C Scaffold Anchoring

The hybrid assembly was scaffolded using the Hi-C reads obtained from the liver sample. Briefly, reads were first aligned to the assembly using the Arima mapping pipeline that uses bwa mem (0.7.17-r1188) and samtools (v. 0.1.20) for alignment and 5´-end filtering. Scaffolding was performed using YaHS (v.1.2.2), the assembly was visualized and hand-curated in Juicebox (v.2.17.00), and the final edited assembly was exported and produced with juicer (v.3.0). The collinearity of the Hi-C scaffolding with the linkage map anchoring was assessed in Allmaps (Tang et al., 2015).

3.2.2 Population genomics of greater amberjack in the South Atlantic and Gulf of America

3.2.2.1 Sampling

Collections from US populations occurred in 2021, 2022, and 2023 in conjunction with the acoustic telemetry component of the project. Table 3.1 summarizes the collection information for the samples received at USM (total of 914). Sampling aimed to obtain 50 individuals per sub-region per year in two successive years. Because most US samples were collected in 2022 and 2023, sample processing focused on the 2022 and 2023 sampling years.

We also pursued samples from other parts of the Atlantic basin to assess connectivity of US stocks with other GAJ geographic stocks. Contacts were made with collaborators to seek assistance collecting GAJ tissue samples from Belize, Brazil, and Canary Islands but only samples from Canary Islands were received.

We initially selected 810 of the US samples and 10 Canary Island samples for sequencing based on tissue quality, sampling year, and availability of geographic location data. In addition, because amberjack species are very similar morphologically, especially at the juvenile stage, and hybridization of *Seriola* species has been reported in previous genetic studies (Takahashi et al., 2021), voucher specimens of other *Seriola* species were sequenced to facilitate detection of mis-identified specimens or hybrids in the dataset. Three *Seriola rivoliana* samples were collected during the project, one *Seriola zonata* sample was procured from our laboratory collection, and two *Seriola fasciata* samples were donated by Dr. N. Phillips.

Ten of the 820 selected greater amberjack samples and the two voucher specimens of *Seriola fasciata* (lesser amberjack) did not yield DNA of sufficient quality for library preparation. The remaining 810 greater amberjack samples were processed alongside the voucher *Seriola rivoliana*, and *zonata* specimens and 10 of the specimens from Canary Islands (which were lesser amberjacks instead of greater amberjacks due to mis identification in the field, see results). Sampling locality-areas in the US included Texas (Galveston and Corpus Christi), Louisiana, Alabama, the Florida panhandle, the Gulf and Atlantic coasts of the Florida peninsula, the Florida Keys and North and South Carolina. Collection locations for the GAJ samples are reported on Figure 3.1. For all collections, fin clipped tissues were preserved in 95% molecular grade ethanol or 20% DMSO-EDTA salt saturated fixative.

3.2.2.2 DNA isolation and Double-Digest RAD sequencing libraries preparation

Genomic DNA from each sample was isolated using Mag-Bind® Blood & Tissue DNA HDQ kit and EZ-96 tissue kits from Omega Bio-Tek (cat. M6399-01 and D1196-01). Samples were processed according to a modified double-digest restriction associated DNA (ddRAD) protocol (Peterson et al., 2012) as described in section 4.2.1.3.2. Library pools of up to 192 libraries consisted of samples representative of each sampling region and survey year to reduce bias that could result from sequencing across different Illumina flow cells.

Paired-end sequencing (2x150bp) was performed on the Illumina NovaSeq X platform at the Johns Hopkins Genetic Resource Core Facility with a target of 5M paired reads per sample.

3.2.2.3 Sequencing reads processing, SNP discovery and filtering

Raw sequence reads were demultiplexed and duplicates were removed as described in section 3.2.1.3.3. Demultiplexed reads were trimmed, and initial quality filtering performed using the dDocent pipeline (Puritz et al., 2014). Trimmed reads were mapped using BWA-MEM (Li & Durbin, 2010) onto the Hi-C anchored greater amberjack assembly developed in section 3.2.1.4.1.

SNPs were identified using FreeBayes (Garrison & Marth, 2012) and filtered with VCFtools where markers with 50% missing data, minor allele counts below 3, and quality scores below 30 were initially removed. Sites with minimum mean read depths below 15 and minimum depths below 10 were also removed followed by a progressive filtering process for sites and individuals: sites missing >60% data, individuals missing more than 60% data were removed, followed by sites missing >50% data, individuals missing more than 50% data, and sites missing >40% data. Sites exhibiting extreme allelic balance and improperly paired reads were filtered via dDocent_filters (https://github.com/jpuritz/dDocent/blob/master/scripts/dDocent_filters). Additionally, sites

potentially representing repetitive DNA, indicated by coverage of a mean depth greater than twice the standard deviation of the mean site depth for the entire dataset, were removed.

Multi-nucleotide polymorphisms (MNPs) were then deconstructed into multiple SNPs, before filtering indels, and the dataset was then limited to biallelic sites. Sites were then restricted to those with less than 10% missing data, and individuals missing more than 20% data were removed. Sites were then filtered to retain those with minor allele frequencies above 0.01. To further control for sequencing bias affecting samples loaded on different Illumina Novaseq sequencing lanes, an analysis of molecular variance was computed clustering samples in their sequencing lane within a geographic group-sampling year sample. Sites for which the among-sequencing lane component of variance was significantly greater than zero ($P < 0.05$) were removed from the dataset. Individuals identified as duplicates or close kin relationships were removed during Kinship analyses as described in section 3.2.2.5.2. The data for each geographic population (aggregated as described in the results section) was tested independently for Hardy Weinberg Equilibrium and sites exhibiting significant departure in at least 25% of the groups were removed. Sites were then restricted to a maximum of 5% missing data and individuals missing less than 10% data.

3.2.2.4 Mitochondrial DNA sequencing

Partial sequences of the Cytochrome Oxidase 1 sub-unit 1 (COI-1) gene were obtained to confirm some of the species identifications inferred from the SNP dataset. A 706 bp fragment was amplified using the primers FF2d and FR1d described by Ivanova et al. (2007). PCR reactions was conducted in a 50 ul total volume including (4.4 ng template DNA, 1X Promega 5X Colorless Go Taq Flexi Buffer, 1.75 mM MgCl2, 0.2 mM dNTP, 0.5 uM of forward and reverse primers, 2.5 units of GoTaq Flexi DNA Polymerase, ddH2O to 50 ul. Amplification employed a thermos-cycling protocol that consisted in an initial denaturation at 94°C for 2 min, followed by 35 cycles at 94°C for 30 s, 54°C for 30 s, and 72 °C for 1 min, final extension at 72°C for 10 min.

PCR products were submitted to Eurofin Genomics for clean-up and sequencing using the forward and reverse primers. Sequencing reads were obtained on an ABI 3730 XL sequencer (Applied Biosystems). The forward and reverse sequences were edited to remove low quality base calls and assembled in a contig in BioEdit v. 7.0 (Hall, 1999). The consensus sequence was submitted to Blastn online (v. 2.16.1+) of the National Center for Biotechnology Information for a BLAST search of the Blastn core database using the MEGABLAST algorithm to identify closest matches. Matches were ranked based on E values (Kerfeld & Scott, 2011).

3.2.2.5 Population genomic analysis

3.2.2.5.1 Species identification

The dataset was examined using Principal Component Analysis (PCA, Patterson et al., 2006) in Adegenet 2.1.6 (Jombart, 2008) and Bayesian clustering in ezsturcture (Raj et al., 2014) to visualize the 4 *Seriola* species clusters and detect potential mis-identified specimens within the greater amberjack dataset or hybrids of greater amberjack with other species. Species identifications based on Bayesian clustering ancestry assignments and PCA clustering were further confirmed for a subset of samples assayed at mitochondrial DNA as described in section 3.2.2.4 by comparing the matching species at mitochondrial

DNA with the species inferred from clustering based on the SNP profiles. Samples that were matched to another *Seriola* species were removed from the dataset.

3.2.2.5.2 Kinship analyses

Pairwise relatedness coefficients were estimated in KING v. 2.3.2 (Manichaikul et al., 2010) for all pairwise relationships. Unrelated pairs were separated from close relatives (clones/duplicates, 1st degree relatives such as parent offspring or full sibs, and 2nd degree relatives such as half sibs). The KING output was used as input to NAToRA (Leal et al., 2022) to identify the smallest number of individuals to be removed to achieve acceptable levels of relatedness while retaining the largest sample size. The obtained dataset did not feature any close kinships (1st and 2nd degree relationships).

3.2.2.5.3 Population comparisons and multivariate clustering

Samples were aggregated in 10 collection areas (Galveston, TX, Corpus Christi, TX, Louisiana, Alabama, the Florida panhandle, Florida's Gulf coast, Florida's Southern Atlantic coast, Florida's Northern Atlantic Coast, and North Carolina and South Carolina) for analysis.

Occurrence of cryptic structure within the dataset was examined using Principal Component Analysis (PCA, Patterson et al., 2006) in Adegenet 2.1.6 (Jombart, 2008) and Bayesian clustering accounting for spatial autocorrelation in TESS 3.0 (Caye et al., 2016). Discriminant analyses of Principal Components (DAPC, Jombart et al., 2010) was applied to identify population clusters. The number of principal components to retain was determined by cross-validation. K-means clustering was used to infer the number of clusters, DAPC was also performed using a priori groups (the 10 geographic populations; Gulf and Atlantic; and three major groups described below, Figure 3.2).

F_{ST} was computed between all pairs of populations and significance was assessed using an exact test comparing allele distributions between groups in Arlequin v.3.5.1 (Excoffier & Lischer, 2010). The exact probability value was estimated using a Markov chain approach (Raymond & Rousset, 1995) using 10,000 steps. Probability values were adjusted to account for multiple testing using a False Discovery Rate correction (Benjamini & Hochberg, 1995).

Analyses of molecular variance (AMOVA, Excoffier et al., 1992) was used to partition genetic variance among geographic groups. AMOVAs were computed comparing the 10 geographic groups described above and hierarchical models accounting for two (Gulf and Atlantic) or three (Western Gulf, Eastern Gulf, Atlantic) major groups. The Atlantic group included the Carolinas, North Florida Coast, and South Florida Coast localities (nested in the Atlantic group), the Western Gulf included samples from Texas and Louisiana, and the Eastern Gulf included samples from Alabama to the Florida Gulf side, Figure 3.2). Hierarchical AMOVAs were conducted for the whole dataset and separately for the 2022 and 2023 sampling years. Additional AMOVAs were also conducted accounting for sampling year (within locality) to test the significance of temporal variations within a geographic region.

Finally, the dataset was restricted to individuals collected during the spawning period (March to June (Schultz et al., 2025) and aggregating samples in regional groups and sampling year as above.

3.2.2.5.4 Spatial analyses

For spatial analyses, least-cost paths between capture locations of each individual were calculated in the R package marmap (Pante & Simon-Bouhet, 2013) and used as the measure of geographic distance.

Occurrence of isolation by distance was first examined using a Mantel test (Smouse et al., 1986) testing the correlation between genetic distance and geographic distances. The multilocus genetic distance described in Smouse and Peakall (Smouse & Peakall, 1999) was used as a measure of genetic distance between individuals. Significance was assessed based on 10,000 permutations of the distance matrices in GenAlEx v. 6.5.0.3 (Peakall & Smouse, 2006, 2012).

Spatial genetic variation was further assessed using spatial autocorrelation analysis (Smouse & Peakall, 1999). The spatial autocorrelation coefficient (r) was computed using genetic and geographic distances described above. Under isolation by distance, estimated values of r differ significantly from zero for geographically proximal samples, and decrease with increasing geographic distance between samples. The distance between samples at which r no longer differs significantly from zero provides an approximation of the distance at which population structure can be detected (Peakall et al., 2003). Because the estimation of spatial autocorrelation is influenced by the size of the distance class (Peakall et al., 2003), r was computed for distance classes of increasing value from 100 km classes to 2,500 km. The significance of r was determined via 10,000 random permutations of the genotypes of individuals among distance class groups. The distribution of r values under the null hypothesis of random spatial distribution of genotypes was used to determine the probability of significance of the observed values of r according to a one-tailed test. The significance of r was also tested by generating bootstrap 95% confidence intervals (CIs) for r within a distance class group. Bootstrap values were obtained by sampling with replacement pairs of individuals within a given distance class. Bootstrap resampling was performed 10,000 times, and the significance of r was inferred when the 95% CI did not overlap zero.

The autocorrelation framework was applied to study temporal autocorrelation of samples obtained within regions aggregated in 5 groups: Western Gulf (Galveston, Corpus Christi and Western Louisiana), Central Gulf (Eastern Louisiana through the Florida Panhandle), Southwest Florida (Florida Gulf), Southeast Florida (North and Southeast Florida), and the Carolinas. Autocorrelation analysis was computed as described above for spatial autocorrelation except that the distance between individual samples, instead of a least cost geographic distance used in spatial autocorrelation, was the number of days between collection dates for the two samples. Because the estimation of temporal autocorrelation can also be influenced by the size of the distance class, r was computed for temporal classes of increasing value from 10 to 100 in ten-day increments, and 30 to 300-450 days in 30-day increments.

3.2.2.5.5 Outlier analyses

The dataset was screen for loci potentially affected by divergent (natural) selection using an outlier analysis in outFLANK (Whitlock & Lotterhos, 2015). This method uses a trimmed likelihood approach to detect outlier loci showing significantly higher divergence among populations than expected under the neutral distribution. Outlier searches were conducted aggregating samples in 10 sampling regions (Figure 3.2) but also aggregating them in two main groups (Gulf of America and Atlantic), and three groups (Western Gulf, Eastern Gulf and Atlantic, Figure 3.2).

Findings

3.3.1 Reference genome

3.3.1.1 Genome Assembly

Long read sequencing yielded over 41.3 Gigabase pairs (bp) across 2,280,601 reads averaging 58,823 bp in length. Illumina sequencing generated over 76.5 Gb of sequence in 507,026,486 2 x 150 bp paired reads. The corresponding genome coverage with each platform was over 66X and 123X for PACBIO and Illumina, respectively. The length of the unanchored assembly was 625,192,197 bp across 224 scaffolds (N50 = 18,175,249 bp, L50= 11 scaffolds) and yielded a complete BUSCO score of 93.4%.

3.3.1.2 Linkage Mapping

The F1 family dataset included dd-RAD sequences for 153 individuals (male and female parent, 151 offspring). Variant calling yielded 15,356,745 initial sites. Filtration of the dataset to remove low quality candidate SNPs and individuals with low sequencing coverage resulted in 12,098 sites across the 153 individuals. The optimal LOD score for assignment of loci to linkage groups in SeparateChromosomes2 was 20. After removal of uninformative markers and those showing segregation distortion. 6,026 informative markers were clustered into 24 linkage groups. Ordering and anchoring in LepAnchor was completed leaving some ambiguities in ordering although most of the assembly (98.9%) was anchored to the linkage map (Table 3.2). The marker positions on the anchored assembly are represented on Figure 3.3.

3.3.1.3 Hi-C sequencing

The paired-end Hi-C library generated from liver tissue produced 444,087,944 reads. Reads were processed in the Arima Genomics mapping pipeline, followed by scaffolding with the YAHS pipeline (Zhou et al., 2023). The contact map was visualized and hand-curated in Juicebox (v. 2.17.00, Figure 3.4) and the final edited assembly was produced and exported with juicer (v.3.0).

Hi-C anchoring recovered slightly better metrics than the linkage map (Table 3.2) and was used for the population genomics analysis.

3.3.2 Population genomic analysis

3.3.2.1 Sample procurement

A total of 810 GAJ samples of suitable quality for sequencing were selected to represent the different sampling areas and the two years. Sample locations and sample sizes per region x year are provided in Figure 3.2 and Table 3.3, respectively. After filtration of the dataset, 704 samples remained for population structure analysis. The ten sampling regions are represented with sample sizes ranging from 1 (FL-panhandle 2023) to 102 (North Carolina 2022).

3.3.2.2 Species identification

The initial filtering of the dataset including the US populations along with the Canary Island samples and the voucher *Seriola rivoliana* and *zonata* yielded 36,413 SNP loci. Multivariate analyses were conducted on this dataset to identify potential species mis-identifications.

Principal Component Analysis (PCA) identified three groups of highly divergent individuals (3. 5) that corresponded to the *Seriola rivoliana* voucher specimens and one amberjack collected from the Florida Gulf coast during the study in a first outlier cluster, the *Seriola zonata* voucher and 12 of the putative greater amberjack specimens from this study in a second cluster, and the 7 Canary Island specimens remaining after filtration along with one specimen from Alabama in a third outlier cluster. Bayesian

clustering confirmed the pattern from PCA (Figure 3.6a) and also identified a potential hybrid intermediate between the *Seriola zonata* and *Seriola dumerili* clusters (Figure 3.6b).

Mitochondrial DNA sequencing was conducted to verify species identification for the 7 Canary Island samples (and the other samples received for this locality) and the 12 putative greater amberjack that clustered with the voucher *Seriola zonata*. Mitochondrial DNA sequences revealed that the Canary Island samples were all lesser amberjack (*Seriola fasciata*) and confirmed that the 12 greater amberjacks clustering with the *Seriola zonata* voucher were *Seriola zonata*. The sample found in intermediate position between the banded rudderfish and greater amberjack clusters during PCA showed an estimated 51% ancestry in the *Seriola dumerili* cluster and 49% ancestry in the *Seriola zonata* cluster in ezstructure (Figure 3.6b), consistent with the hypothesis of a first-generation hybrid. The 14 mis-identified *Seriola* (12 *Seriola zonata*, one *fasciata* and one *rivoliana*) and the *zonata-dumerili* potential hybrid were removed from further analysis.

3.3.2.3 Population genomic structure of Southeast US greater amberjack

Variant calling was performed on the dataset restricted to greater amberjack specimens only and identified 16,848,233 variable sites. Subsequent filtering refined the data to 5,188 biallelic SNPs in 781 samples. Fifty-seven individuals were identified where kinship coefficients exceeded 0.177 (full sibs), likely indicating contamination or duplicate samples and were therefore removed the dataset. Natora removed an additional 5 samples displaying lower relatedness levels for a final number of 704 individuals and 5,188 markers. There were no 1st or 2nd degree relationships in the final dataset.

3.3.2.3.1 Divergence among regions and sampling years

Pairwise F_{ST} estimates between the ten populations ranged between -0.00003 and 0.00227. No estimate differed significantly from zero (see Table 3.4).

Analysis of Molecular Variance (AMOVA) results were consistent with exact tests and F_{ST} estimates in that the among-population component of molecular variance was very low; the variance among geographic groups was less than 0.05% of the total molecular variance in the different AMOVAs examined. Only the among-regions component of molecular variance was significantly different from zero during hierarchical AMOVAs accounting for regions using the full dataset and the 2022 sub-dataset but it was not significant during analysis of the 2023 sub-dataset (Table 3.5). The temporal component of molecular variance did not differ significantly from zero but it was larger than the spatial component (Table 3.5).

Restricting the dataset to samples collected during the spawning season (March through June, Figure 3.7) and aggregating samples in 5 regions did not reveal significant divergence among regions (Table 3.5, results were unchanged when samples were aggregated in 4 groups (data not shown).

3.3.2.3.2 Multivariate and spatial analyses

Bayesian clustering cross validation scores after the removal of extraneous *Seriola* samples and the putative hybrid in TESS showed a linear increase between 1 and 12 clusters supporting inference of 1 single cluster for the 704 confirmed GAJ (Figure 3.8a). Visualization of the ancestry matrix in ezstructure for 2 clusters shows all 704 individuals with uniform ancestry assigned to one cluster (Figure 3.8b).

All samples were aggregated around the centroid during Principal Component Analysis indicating high homogeneity across the sampling surface (Figure 3.9a). Discriminant Analysis of Principal Components with a priori groups also showed large overlap between locality samples indicative of genetic homogeneity (Figure 3.9b).

No significant spatial autocorrelation was detected during analysis of the two sampling years. The correlation was slightly greater than zero for small distance classes (100 km) and for classes between 500-800 km and 1600-1800 km (Figure 3.10a) in the 2022 dataset but none of the correlation differed significantly from zero. Spatial autocorrelation analysis for sampling year 2023 were near zero or negative (Figure 3.10b).

Temporal autocorrelation was assessed within five regions: Western Gulf (Galveston and Corpus Christi; n=145), Central Gulf (Louisiana through the Florida Panhandle; n=276), Southwest Florida (Florida Gulf; n=49), Eastern Florida (North and South Florida; n=81), and the Carolinas (n=144). Positive temporal autocorrelation was detected for Eastern Florida (n=81) with the correlation highest for samples collected within a 60-days interval, and signatures stabilizing within 120 collection days (Figure 3.11). A significant positive correlation was also detected for Southwest Florida for samples collected within a 270 day window and for the central Gulf region when samples were collected within 450 days. No significant temporal autocorrelation was detected in the other regions.

3.3.2.3.3 Outlier analyses

Outlier analyses accounting for 10 sampling localities, or aggregating them in 3 regions did not detect any outliers. Analyses aggregating samples in two major groups ('Gulf' and 'Atlantic') indicated 8 candidate outliers which showed low F_{ST}, and were not supported by other loci in the same genomic region (Figure 3.12).

3.4 Interpretations and Discussion

This chapter aimed to develop a draft reference genome for genetic studies of greater amberjack and apply this resource to a population genomic survey of GAJ in US waters of the South Atlantic and Gulf of America. The reference genome was obtained by combining sequencing and assembly of long and short reads and anchoring the obtained assembly with Hi-C sequencing and linkage mapping to achieve a chromosome level reference genome with high contiguity as needed to support analysis of genome-wide linkage disequilibrium and inferences on selection. The population genetic component of the project applied this resource to document genetic stock structure of the species in US waters of the South Atlantic and Gulf of America regions.

3.4.1 Genomic resources for GAJ

Assembly yielded a highly contiguous reference with only 224 scaffolds encompassing most of the genome and N50 and L50 or 18,175,249 and 11, respectively. The obtained reference showed a high BUSCO score (above 93%), suggesting that most of the greater amberjack genome was captured during sequencing and assembly. The contiguity is substantially higher than that of other assemblies generated previously in our laboratory (Dimens et al., 2024; Hildahl and & Saillant 2025) using the same hybrid assembly approach. Contiguity was promoted in part by the high coverage obtained with long PacBIO HiFi reads as the previous assemblies did not apply the HiFi sequencing technologies and thus were impacted by lower sequencing accuracies at the end of reads; the previous assemblies also were based

on lower coverage with PacBio reads (20 to 30 X versus close to 66X of raw reads for this assembly). When this work began, the only available assembly for GAJ was highly fragmented with an N50 of only 0.25Mb (Sarropoulou et al., 2017) and it featured incomplete annotations (Chen et al., 2025). In parallel to this work, another complete assembly of the greater amberjack genome was produced by sequencing GAJ specimens from China (Chen et al., 2025). The metrics of the Chen et al. (2025) assembly are similar to those of our assembly except that it is slightly more complete with a BUSCO score of 98.5%. It will be interesting to study the synteny between the West Atlantic GAJ assembly produced in this work and that of Chen et al. (2025). The Chen et al. (2025) was only published in December 2025 and the comparison of the two assemblies has not yet been done but will be completed in the coming months. The West Atlantic assembly developed in this work was the most suitable for use in the population genomic study of greater amberjack from the Gulf and US East coast as we expect significant divergence between GAJ from the West Atlantic and those from the West Pacific sampled by Chen et al. (2025) due to the long period of isolation between these two groups.

Anchoring of an assembly in a chromosome level reference is a critical step to support population genomics inferences (Thorburn et al., 2023). Specifically, anchoring enable analyses of genome wide linkage disequilibrium, assessment of variation under selection, and selection of markers with known degree of physical linkage. In this project, two approaches to anchoring were conducted. The initial anchoring in chromosomes employed a linkage map developed using SNP genotypes derived from ddRAD sequencing of parents and offspring of a single-pair cross conducted at the Institute of Marine Biology of Crete in 2023. The linkage map anchored 98.9% of the assembly in 24 major linkage groups to produce pseudo-chromosomes. yielding a first draft chromosome level assembly. The assembly was then anchored using Hi-C chromatin contact maps which also yielded the 24 major groups with a slight improvement in contiguity and completeness. The latter assembly was used to map double-digest RAD sequencing reads for the population genomic analysis and test for the presence of outlier regions.

3.4.2 Population genomic structure and connectivity

The second objective of this chapter was to describe genetic stock structure of GAJ in US waters of the South Atlantic and Gulf of America to complement studies of movement by acoustic telemetry and provide essential data for the definition of stocks for management. Sampling focused on US waters and was conducted in conjunction with the tagging effort of this project. The project team also recognized the importance of accounting for connectivity with other stocks likely to exchange migrants with US stocks in the analysis. Therefore, efforts were made to obtain samples from other parts of the Atlantic basin. Collaborators in Brazil, Belize and Canary Island committed to collect tissues from specimens in these areas but samples were received only for the Canary Islands and, unfortunately, the 21 samples received from this area were matched to lesser amberjack using mitochondrial DNA sequencing and had to be removed from the dataset. The analysis of stock structure proceeded with samples from localities along the US coastline from North Carolina to South Texas.

3.4.2.1 Species identification and hybridization

Identification of *Seriola* specimens in the field is known to be challenging and a recent study revealed that mis-assignments at sea were most likely to affect greater amberjack; field assignments were found to be accurate in only 24% of the ‘difficult cases’ that were checked by genetic analysis (Galbraith et al., 2022). Species assignment errors did occur in this study where all the specimens received from Canary Island (matched to *Seriola fasciata*) and 14 specimens collected along the US coastline (12 *Seriola*

zonata, 1 *Seriola rivoliana*, 1 almaco and a possible hybrid *dumerili* x *zonata*) had been incorrectly identified as *Seriola dumerili* and were matched to other species using genetic data. While the proportion of mis-assigned specimens was low (2% of the individuals in the final dataset), the few misassigned specimens would have been sufficient to alter substantially conclusions of the population genomics study had they not been discovered. The 12 *Seriola zonata* were collected only in Florida and North Carolina and would have driven inference of occurrence of a divergent group that was only occurring in the eastern part of the range. This outcome could have led to the incorrect definition of management units. While genetic identification can always be used to confirm species identifications, the field characters developed by Galbraith et al. (2022) will be useful to reduce further the frequency of mis-identified *Seriola* in the future.

The detection of one possible hybrid in this work also suggests that natural hybridization between *Seriola* species could be a source of bias in genetic and other studies of this group. Hybrids are expected to be intermediate between the two parental species, making them more difficult to identify than pure species. In the case of the *Seriola* genus where parental species are already difficult to separate as discussed above, detection of hybrids in the field based on phenotypic characters would be particularly challenging. While hybrids are expected to occur in very low frequency, even very low numbers could impact inferences on population structure. Natural hybrids have been discovered in other *Seriola* species (quinqueradiata x lalandi, Takahashi et al., 2021) but have not been reported in *dumerili* x *zonata* to our knowledge. The hybrid status of the potential *dumerili* x *zonata* specimen detected during this study was inferred from ancestry proportions during Bayesian clustering and PCA. The ancestry proportions in the two genomes were close to 50%, suggesting the possible hybrid was an F1, which would be the most likely scenario considering that hybrids are often sterile or have reduced reproductive success. The alternative explanation for the mixed ancestry inferred during Bayesian and multivariate clustering would be a contamination of field samples (a S. *zonata* specimen sample contaminated by a *S. dumerili* sample or vice versa), The two scenarios (contaminated sample versus a F1 hybrid) are challenging to differentiate as in both cases high heterozygosity is expected but examination of the pattern of heterozygosity across loci may provide further information on the status of affected samples. Examination of mitochondrial DNA could also be helpful to confirm the hybrid status or reveal contamination (if multiple mitochondrial DNA haplotypes are observed). In either case, the affected samples need to be removed from a population genomic study of *Seriola dumerili* and the potential *zonata* x *dumerili* hybrid identified in this study was removed from the dataset.

3.4.2.2 Population genomic structure

Population structure was assessed using a combination of approaches including comparisons of geographic and temporal groups, multivariate clustering methods, and spatially explicit models.

The level of divergence among geographic samples was very weak and non-significant during most pairwise comparisons of localities, suggesting high connectivity across the region. When samples were grouped in regions, AMOVAs only revealed a significant among-region component when major regions were compared in the overall dataset and in the 2022 sub-dataset but divergence among the three regions was not significant during analysis of the samples collected in 2023. Multivariate analyses (PCA, DAPC and Bayesian clustering) were also consistent with the inference of high connectivity across the area and provided no evidence for discrete cryptic groups. Finally, spatial autocorrelation analysis did not reveal a significant isolation-by-distance pattern or other spatial autocorrelation of genotypes.

The lack of consistent structure across the area departs from findings reported in previous studies of GAJ population structure in US waters (Gold and Richardson, 1998; Hargrove et al., 2018). Gold and Richardson (1998) reported divergence at mtDNA haplotypes between an Atlantic group that included samples collected along the US East coast and those from the Florida Keys and a Gulf group that included samples from the Gulf of America (except the Florida Keys). The structure inferred between the two groups was weak (ϕ_{ST} = 0.009). Hargrove et al. (2018) used 11 microsatellite loci and also found limited evidence for population structure in the area with low F_{ST} estimates when comparing the Florida Keys locality sample to samples from the Gulf. Divergence was slightly higher when spawners from two breeding populations (Louisiana and Keys) were compared and used as a reference to train a Bayesian clustering model but remained low. In this study, restricting the dataset to the samples obtained during the spawning season (March to June) did not reveal significant structuring either but the maturity status of these summer samples was not assessed (individuals were tagged and released) and some may have been juveniles. While it is possible that restricting the dataset to mature breeding individuals would have revealed some structuring as it did in Hargrove et al. (2018), we note that multivariate analyses did not provide any evidence for cryptic structure within the dataset. Accordingly, if distinct breeding units occur, they likely feature very weak levels of genetic divergence.

Efforts to group samples by major regions including the Atlantic and Gulf groups defined by Gold and Richardson (1998) and Hargrove et al. (2018) or additional geographic groupings did not reveal a consistent geographic signal either. The only significant population structure signal was observed when the East coast, Central Gulf and Western Gulf were compared in the 2022 sub-dataset but this result was not repeated in the 2023 sampling year. While the non-significant outcome of the tests conducted on the 2023 sub-dataset could be due to the smaller sample size in that year, it could also signal temporal variability. The temporal (between-year) component of molecular variance was not significantly greater than zero but it was positive and larger than the among-localities component during analysis of molecular variance, suggesting temporal instability of population structure. Temporal stability within geographic regions was examined by testing the temporal autocorrelation of genotypes. This analysis suggested that the Western Gulf and Carolinas populations were more stable (no significant temporal autocorrelation at any of the time scales tested). In contrast, East Florida showed a significant autocorrelation of genotypes sampled within a 30 to 60 days window, suggesting different migrating groups were sampled in this area at different times in the project. The West Florida and Central Gulf group also showed some temporal autocorrelation but it was very weak and was only significant for long time intervals (270 days for West Florida and 450 days for the Central Gulf), suggesting slower and more gradual shifts; although, for the West Florida analysis, the time interval could be consistent with an annual spawning migration.

A potential explanation for the different findings in this study and those of Gold and Richardson (1998) and Hargrove et al. (2018) could be an on-going change in population structure in the region. Recent stock assessments indicated that the Atlantic stock was not overfished and not undergoing overfishing (SEDAR 59, 2020) while the Gulf stock has been overfished since the late 80s (SEDAR 70, 2020). The Gulf breeding stocks may have been experiencing higher levels of gene flow from Atlantic GAJ since the sampling of the two previous studies (1990s for the Gold and Richardson (1998) study, 2008 and 2009 for the study of Hargrove et al. (2018)) if the local Gulf spawning stocks experienced depletion by overfishing. This could have led to homogenization of the two gene pools and ultimately to the low or lack of divergence observed in this study. Considering the short time frame since this extended mixing

between the two groups would have been initiated, this hypothesis could be examined using denser genome scans by analyzing identity by descent tracks (Browning and Browning, 2011).

A second hypothesis explaining the different patterns of structure between this study and two previous studies could be the occurrence of residual species mis-identifications in some of the datasets. In this study, species identification in the field was conducted carefully by expert scientists and some specimens were still mis-identified. The haplotypes studied by Gold and Richardson (1998) may not have allowed identification of specimens from closely related *Seriola* species if the related species shared a high proportion of the sites recognized by the 16 restriction enzymes used to digest mitochondrial DNA in the study. Undetected species mis-identifications seem less likely in the study of Hargrove et al. (2018) who used microsatellite markers. Microsatellites feature a high mutation rate and high divergence levels would therefore be expected for at least some of the loci surveyed in that study, leading to small groups of strong outliers that are easier to detect during standard multivariate analyses.

Weak structure and high connectivity across a large range encompassing the US East coast and the northern Gulf of America is consistent with what is known of the life history of GAJ. Greater amberjacks frequent both coastal pelagic and reef associated habitats and have high movement capabilities. Tagging studies to date suggest limited movement (McCleelan and Cummings, 1997; Young, 2024), in particular across regions (Atlantic-Gulf and Eastern Gulf-Western Gulf). However, GAJ spawn on shelf edge drop-offs in offshore waters. The latter is expected to promote dispersal of pelagic eggs and larvae over broad geographic areas depending on the velocity and direction of surface currents. Larvae and juveniles have been reported near floating sargassum mats (Wells and Rooker, 2004) and can be further transported by surface currents as they follow those mats, which would further contribute to extended geographic mixing. Within the Atlantic basin, significant structuring was detected between Mediterranean and East Atlantic populations (Šegvić-Bubić et al., 2016) but no structure was detected within the Mediterranean, suggesting that the species showed abundant mixing across the entire basin. Studies of other *Seriola* species also indicate that they form large units showing genetic homogeneity across large areas (for example Eastern, Western, and Southern Pacific for *Seriola* lalandi, Purcell et al., 2015). The occurrence of such large units for GAJ would be compatible with the homogeneity across the Gulf and Atlantic regions of the US.

Finally, the possible relatively sedentary behavior of adults or spawning-site fidelity coupled with pelagic larval dispersal would lead to the prediction of an isolation by distance model of population structure. This study provided no evidence for significant spatial autocorrelation, a finding consistent with the study of Gold and Richardson (1998). This result suggests that the pelagic dispersal, coupled with sargassum-associated movements is an effective mechanism of long-distance dispersal and promotes genetic mixing across the area.

3.4.2.3 Outlier analyses

Outlier analyses revealed 8 outlier loci in the Gulf versus Atlantic analysis, but none were detected in the other comparisons. The outlier loci were not clustered in specific genomic regions in the Manhattan plot and showed only low levels of divergence between geographic groups, overall providing no support for divergent selection affecting populations within the sampling area. However, the density of the genome scan conducted in this study is low (average interval between markers 121.8kb while linkage disequilibrium may only span a few

tens of kb as reported in other species (e.g., Wang et al., 2025, Han et al., 2025) and genomic regions affected by selection may have been missed or represented only by one locus yet still exist. The deployment of higher density genome scans is therefore warranted to assess the occurrence of genomic regions undergoing selection in future studies of GAJ genetic structure.

3.4.2.4 Conclusions and management implications

The results obtained in this study did not reveal divergence of the Gulf and Atlantic stocks or other subdivision within US waters. Other sources of information, in particular tagging of adults may support the maintenance of different stock units for Gulf and Atlantic GAJ and maybe even separate management of eastern and western Gulf if demographic connectivity is limited. The lack of genetic divergence found in this study was unexpected considering the suggestion of Atlantic Gulf structure reported in previous studies but could indicate recent/ongoing mixing of previously divergent units, a hypothesis that could be examined further using higher density genome scans and analysis of identity by descent. Also, samples from the southern Gulf and from Caribbean populations that are likely connected to US stocks could not be obtained in this study. Future work should include these stocks to determine if other units occur in the West Atlantic (for example in the southwest Atlantic) and the East Atlantic and assess their degree of connectivity with US stocks.

Tables

Table 3.1. Summary of samples obtained for the study. (1a) Greater amberjack samples were collected from the Gulf of America and the western Atlantic Ocean during tagging studies throughout the calendar year from 2021 through 2023. (1b) Additional *Seriola* species were collected in the Gulf of America and eastern Atlantic Ocean during the tagging study and by fisheries dependent sampling in the Canary Islands in 2024 and 2025.

3.1a. *Seriola dumerili* samples obtained from project tagging survey

Sampling region	2021	2022	2023	2025
North Carolina	7	128	19	-
South Carolina	-	3	32	-
Southeast Florida	-	45	57	-
Southwest Florida	-	88	48	-
Alabama	15	100	106	-
Louisiana	-	74	15	-
Texas-Galv	-	75	-	-
Texas-CC	-	49	53	-

3.1b. Other *Seriola* species samples used to detect species mis-identifications (number successfully assayed)

Other *Seriola* species	
S. *fasciata* (Canary Islands 2024 & 2025)	21 (10)
S. *zonata* (Gulf of America)	1 (1)
S. *rivoliana* (Gulf of America)	3 (3)
S. *fasciata* (Gulf of America)	2 (0)

Table 3.2. Summary statistics of the greater amberjack genome assembly. Assembly results are presented prior to anchoring a hybrid assembly produced by sequencing one GAJ individual from the Gulf of America with PacBio Hi-Fi and Illumina, after anchoring the hybrid assembly with a second GAJ from the Gulf of America sequenced using Hi-C, and after anchoring the hybrid assembly with a linkage map obtained from one outbred cross produced by the Hellenic Center for Marine Research.

Assembly metric	Unanchored assembly	Hi-C anchored scaffolds	Linkage map anchored scaffolds
# contigs	224	24	24
# contigs (>=10,000bp)	224	24	24
# contigs (>=25,000bp)	209	24	24
# contigs (>=50,000bp)	189	24	24
Total length	625,192,197	618,265,221	618,066,210
Total length (>=10,000 bp)	625,192,197	618,265,221	618,066,210
Total length (>=25,000 bp)	624,934,530	618,265,221	618,066,210
Total length (>=50,000 bp)	624,113,260	618,265,221	618,066,210
Largest contig	83,166,725	34,230,163	33,086,027
GC(%)	40.81	40.8	41
N50	18,175,249	26,698,020	26,929,976
N90	1,560,132	22,061,501	22,023,817
L50	11	11	11
L90	70	21	21
Complete BUSCOs (C)	3,398 (93.4%)	3,382 (92.9%)	3,383(92.9%)
Complete Single-copy BUSCOs (S)	3,363 (92.4%)	3,356 (92.2%)	3,353 (92.1%)
Complete and duplicate BUSCOs (D)	35 (1%)	26 (0.7%)	30 (0.8%)
Fragmented BUSCOs	14 (0.4%)	12 (0.3%)	15 (0.4%)
Missing BUSCOs (M)	228 (6.2%)	24 (6.8%)	242 (6.6%)

* BUSCO refers to Benchmark Universal Single Copy Orthologs.

Table 3.3. Sample counts per location and per year. For each location x year combination, values are the final sample sizes used in population structure analysis after filtration of the SNP dataset. Values in parentheses are the initial numbers of samples assayed.

Region	Location	Year	
		2022	2023
AL	Alabama	74(86)	86 (96)
CC	Corpus Christi	46(49)	44 (51)
FLA	South Florida Atlantic	26 (30)	31 (37)
	North Florida Atlantic	14 (14)	11 (15)
FLG	Florida Gulf	30 (32)	21 (25)
FLPAN	Florida Panhandle	41 (53)	1 (1)
GAL	Galveston, TX	59 (71)	
LS	Louisiana	67 (73)	10 (12)
CAR	North Carolina	102 (120)	12 (13)
	South Carolina	3 (3)	25 (28)

Table 3.4. Pairwise F_{ST} between ten GAJ sampling regions in the Gulf of America and western Atlantic (lower diagonal) and the probability that F_{ST} estimates differ significantly from zero (upper diagonal) before correction for multiple tests. See Table 3.3 for acronym definitions.

	SFL	NFL	NC	SC	FLPan	FLG	AL	GAL	LS	CC
SFL	*	0.05019	0.75062	0.73686	0.50698	0.63637	0.26393	0.38313	0.24849	0.36016
NFL	0.00227	*	0.29651	0.67251	0.4656	0.08999	0.09009	0.46292	0.2975	0.33571
NC	0.00038	0.00129	*	0.90694	0.27047	0.09573	0.55133	0.71844	0.49005	0.80517
SC	0.00074	0.00153	0.00036	*	0.85734	0.95951	0.88853	0.89783	0.97396	0.98782
FLPan	0.00093	0.00184	0.00084	0.00074	*	0.30096	0.51945	0.51054	0.60944	0.82952
FLG	0.00065	0.00211	0.00085	0.00026	0.00113	*	0.4748	0.8712	0.1979	0.54549
AL	0.00064	0.00163	0.0003	0.00045	0.00065	0.00056	*	0.70429	0.50955	0.337
GAL	0.00081	0.00145	0.00039	0.00046	0.0009	0.00042	0.0004	*	0.77468	0.93179
LS	0.00085	0.00156	0.00045	0.00022	0.00076	0.00092	0.00042	0.00046	*	0.97466
CC	0.00063	0.00127	0.00023	-0.00003	0.00044	0.00054	0.00041	0.00019	0.00011	*

Table 3.5. Hierarchical analyses of Molecular Variance accounting for variation among regions and sampling year within region. Data are aggregated in 10 geographic groups described in Table 3.3, 3 geographic regions (Figure 3.2), and 5 geographic regions for the spawning season (Figure 3.7).

Source of variation	Degrees of freedom	Variance component (% variation)	P (rand. values > obs. value)
3 Regions (Western Gulf, Eastern Gulf, Atlantic), years combined			
Among Regions	2	0.00972 (0%)	0.10723
3 Regions (Western Gulf, Eastern Gulf, Atlantic), 10 populations within groups, years combined			
Among regions	2	0.05085 (0.01%)	0.02881
Among populations within regions	7	-0.06162(-0.1%)	0.90297
10 Populations years within Population			
Among populations	9	-0.14581 (-0.03%)	0.87277
Among years / populations	9	0.18612 (0.04%)	0.08485
3 Regions (Western Gulf, Eastern Gulf, Atlantic), 10 populations within groups, Year 2022			
Atl.region/W.Gulf region/E. Gulf region	2	0.06943 (0.01%)	0.02059
Among populations / regions	7	-0.02377 (-0.01%)	0.60762
3 Regions (Western Gulf, Eastern Gulf, Atlantic), 10 populations within groups, Year 2023			
Atl.region/W.Gulf region/E. Gulf region	2	-0.15713(-0.03%)	0.6403
Among populations / regions	6	0.23766 (0.05%)	0.15366
Spawning season sub-dataset, 5 Regions (Western Gulf, Central Gulf, Eastern Gulf, Florida Atlantic, Carolinas), sampling years combined			
Among Regions	4	0.22541 (0.05%)	0.99
Spawning season sub-dataset 5 Regions (Western Gulf, Central Gulf, Eastern Gulf, Florida Atlantic, Carolinas), sampling year within region			
Among regions	4	-0.25013 (-0.05%)	0.8199
Among years / regions	5	0.09686 (0.02%)	0.30386

Figures

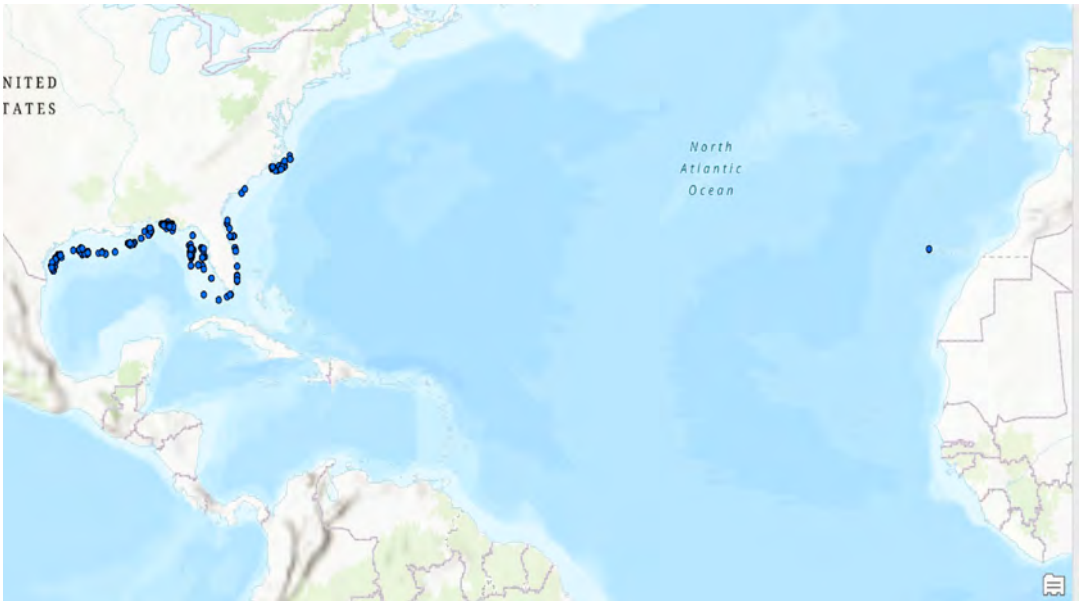


Figure 3.1. Collection sites for 810 greater amberjack samples included in the population genomic structure dataset.

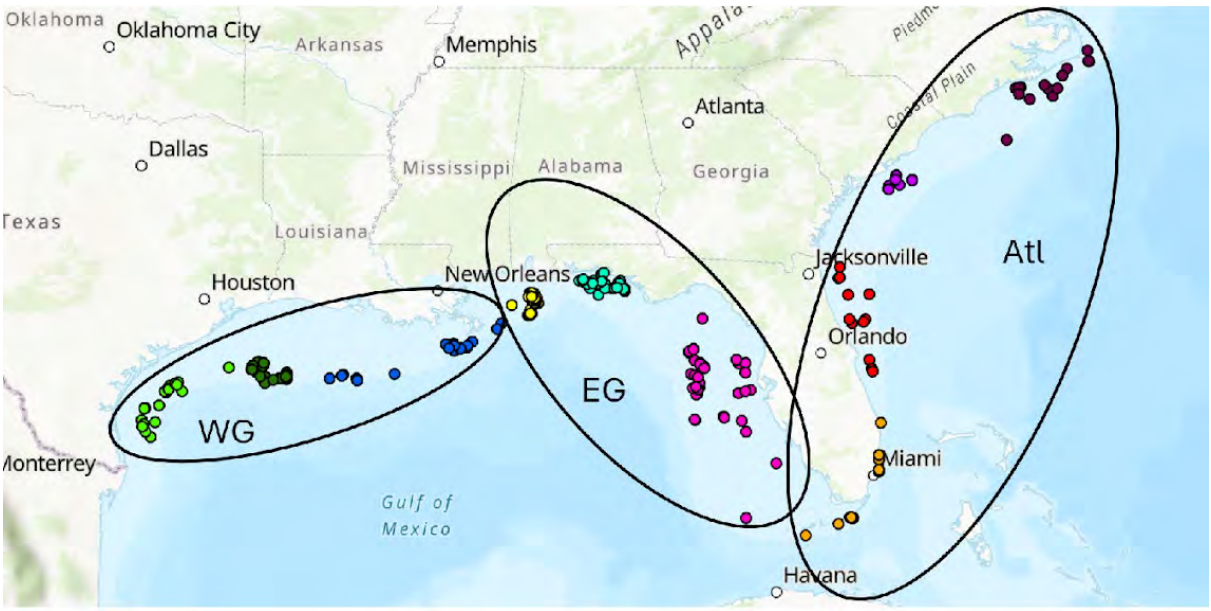


Figure 3.2. Collection sites for the 704 individuals sequenced for the population genetic structure study of greater amberjack. Sites are colored by region: Corpus Christi, TX (bright green), Galveston, TX (army green), Louisiana (blue), Alabama (yellow), Florida panhandle (turquoise), Florida Gulf (pink), Southern Florida Atlantic (orange), Northern Florida Atlantic (red), South Carolina (purple) and North Carolina (burgundy). Large ellipses denote the groups within which populations are analyzed (WG=Western Gulf; EG=Eastern Gulf; Atl=Atlantic).

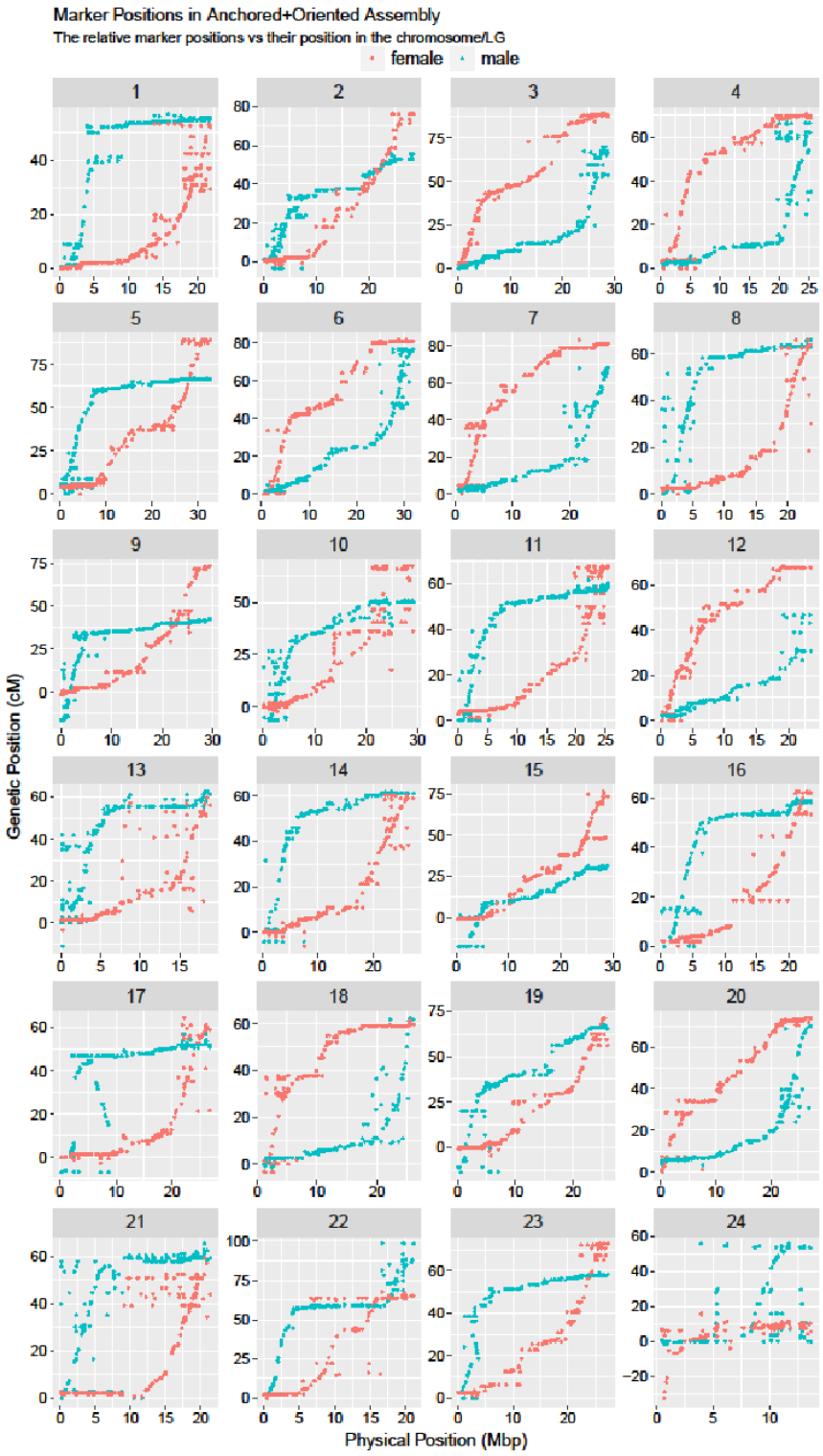


Figure 3.3. Relationships between positions on the physical and linkage maps. Correlation between marker positions on the physical map (Mbp) and position on the male (blue) and female (red) linkage maps (cM). Each plot represents a linkage group, indicated by the number in the grey rectangle.

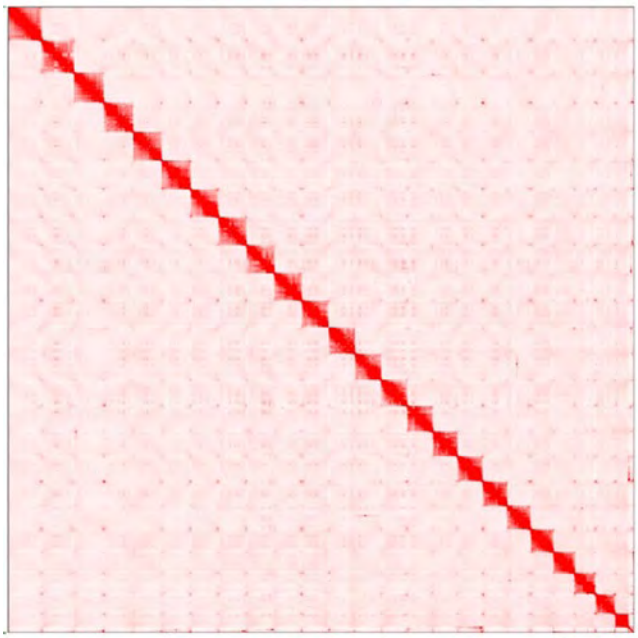


Figure 3.4. Hi-C interaction heat map for the greater amberjack genome. The level of interaction intensity is visualized using a color gradient ranging from white, indicating low density, to dark red, indicating high density. The red dark diamonds represent the 24 chromosomes showing high intensity of self-interacting regions.

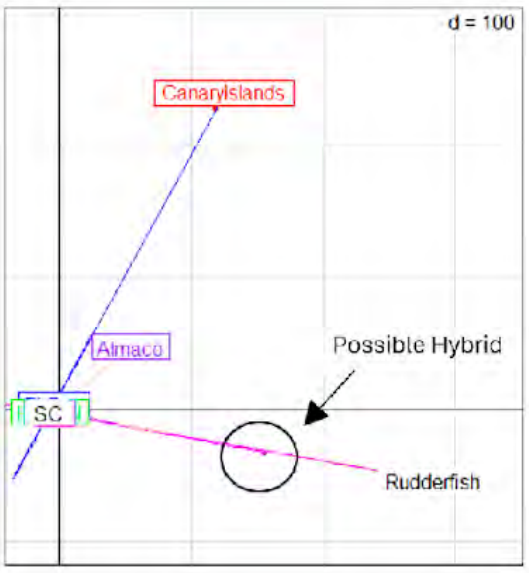
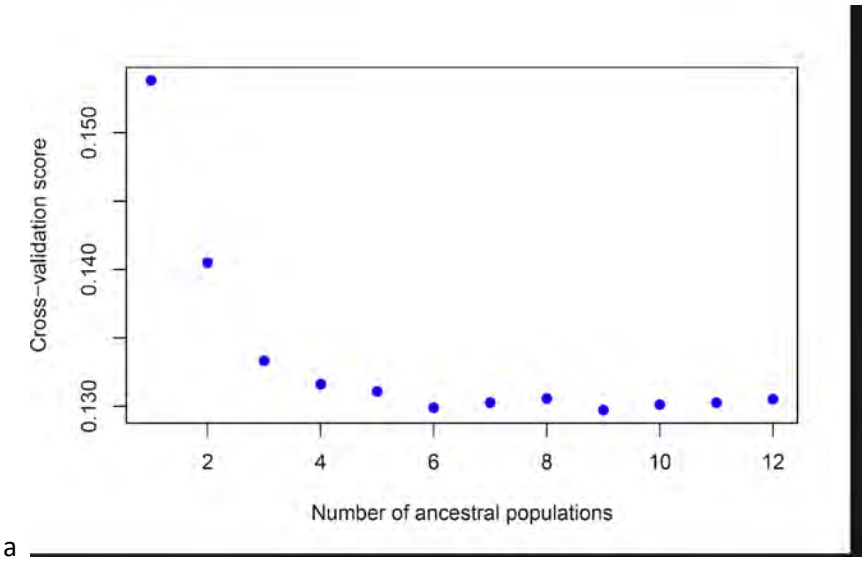


Figure 3.5. Principal Component Analysis. PCA conducted with 36,413 SNP loci on 817 specimens including US and Canary Islands (Spain) samples and vouchers of banded rudderfish and almaco jack.



a



b

Figure 3.6. Bayesian clustering and ancestry proportions for the identification of mis-identified samples. (a) Cross validation scores obtained during Bayesian clustering analysis in TESS for K values of 1 to 12 and (b) ancestry proportions obtained for all individuals for K = 4 (green = Greater Amberjack, blue = Banded Rudderfish, purple = Lesser Amberjack, red = Almaco jack) for the mixed species analysis. The inset highlights the four identified species and a sample showing mixed ancestry (potential hybrid) between greater amberjack and banded rudderfish.

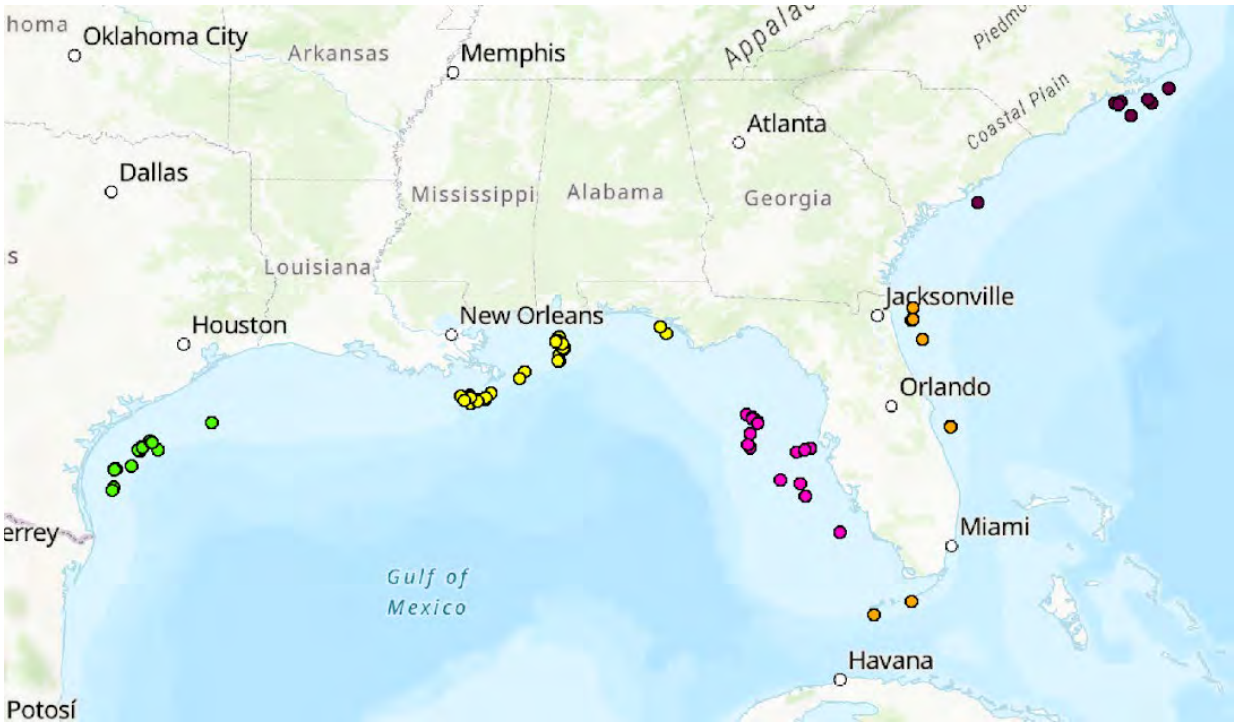
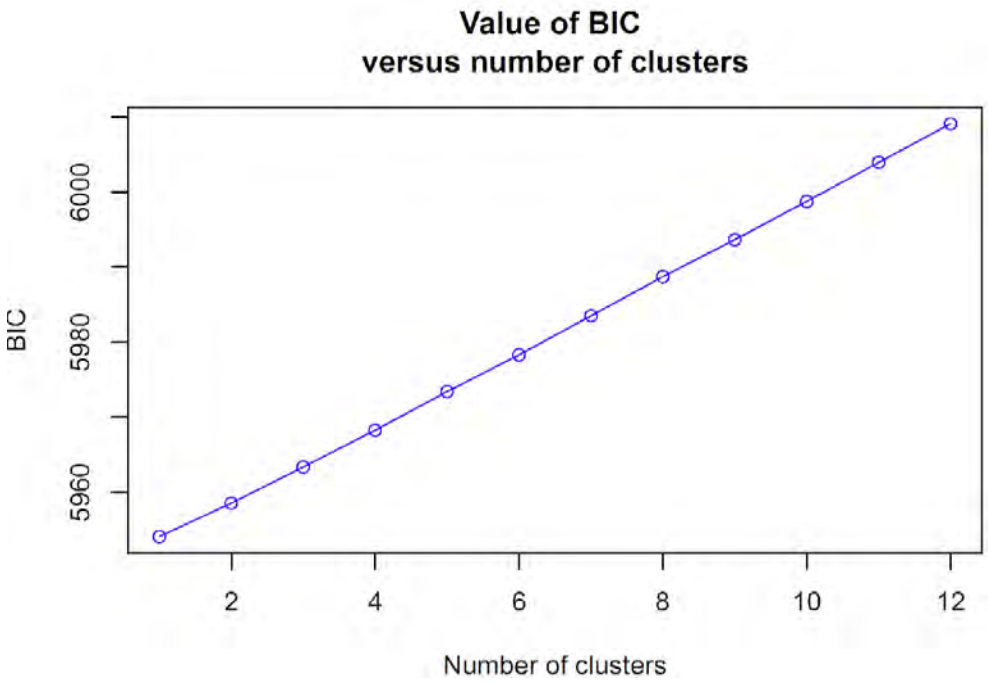


Figure 3.7. Collection sites for the 260 greater amberjack individuals collected during the spawning season (March to June) and analyzed during the study. Sites are colored by statistical region: Western Gulf (WG; bright green), Central Gulf (CG; yellow), Florida Gulf (FLG, pink), Florida Atlantic (FLA; orange), and the Carolinas (burgundy).



a



b

Figure 3.8. Bayesian clustering and ancestry proportions for the greater amberjack dataset (other *Seriola* species and potential hybrids removed). **(a)** Cross validation scores obtained during Bayesian clustering analysis in TESS for K values of 1 to 12 and **(b)** ancestry proportions obtained for all individuals for K = 2 (all individuals show ancestry in one single red cluster).

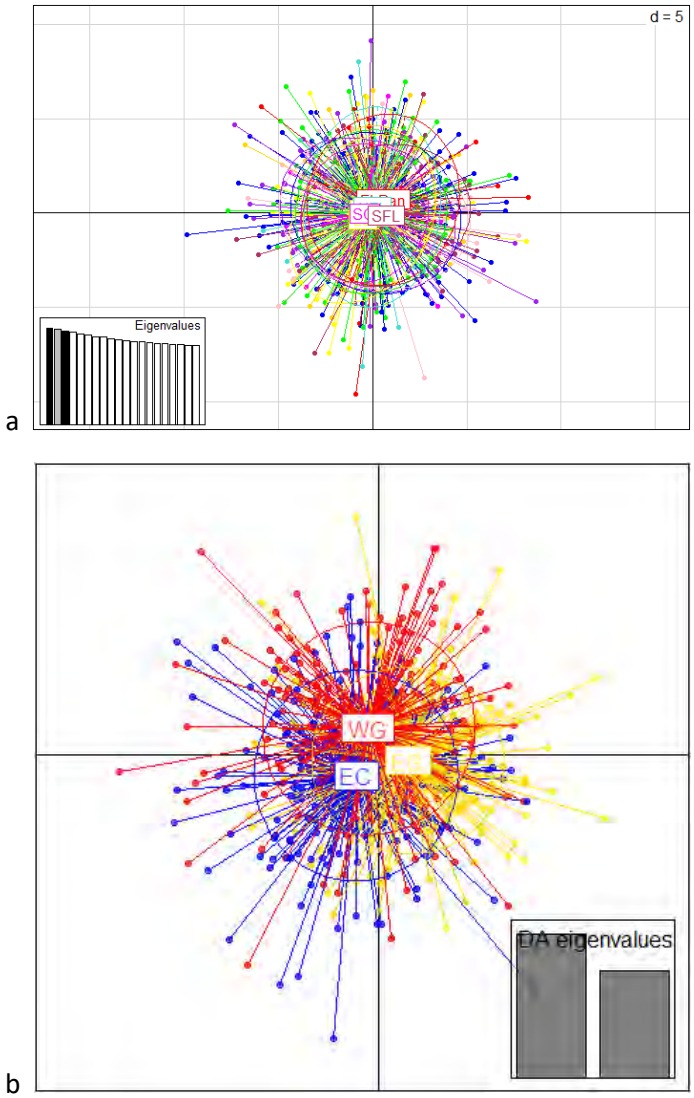
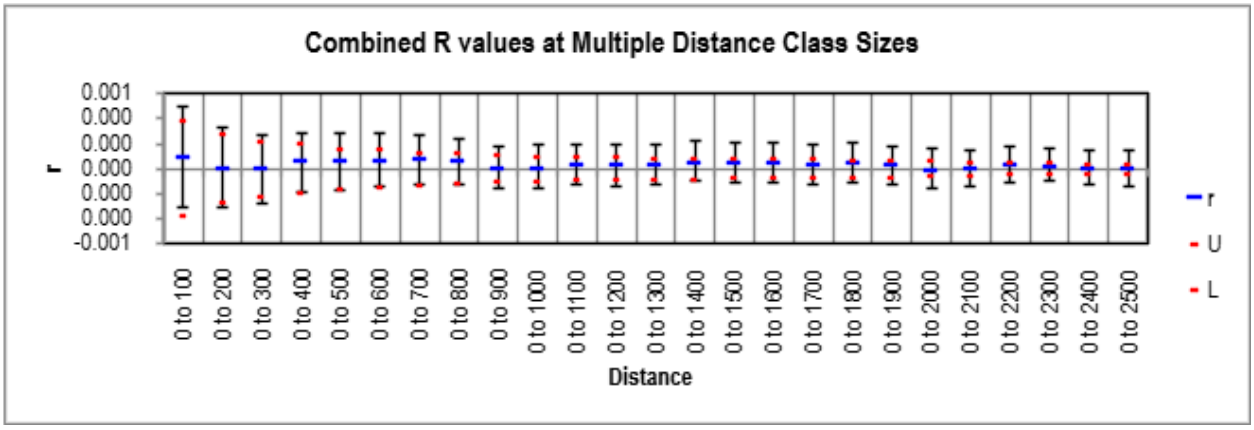


Figure 3.9. Principal Component Analysis and Discriminant Analysis of Principal Components of 704 greater amberjack samples collected in the Gulf and western Atlantic. (a) Principal Component Analysis of Greater Amberjack based on 5,188 loci and 704 individuals. Geographic localities are represented as different colors on the graph (AL=blue; CC=purple; FLG=gold; FLPan=red; GAL=pink; LS=yellow; NC=green; NFL=turquoise; SC=magenta; SFL=maroon). (b) Discriminant Analysis of Principal Components with a priori groups based on 5,188 loci and 704 individuals. Regions are represented as different colors on the graph (i.e. WG=red; EG=yellow; ATL=blue).

a



b

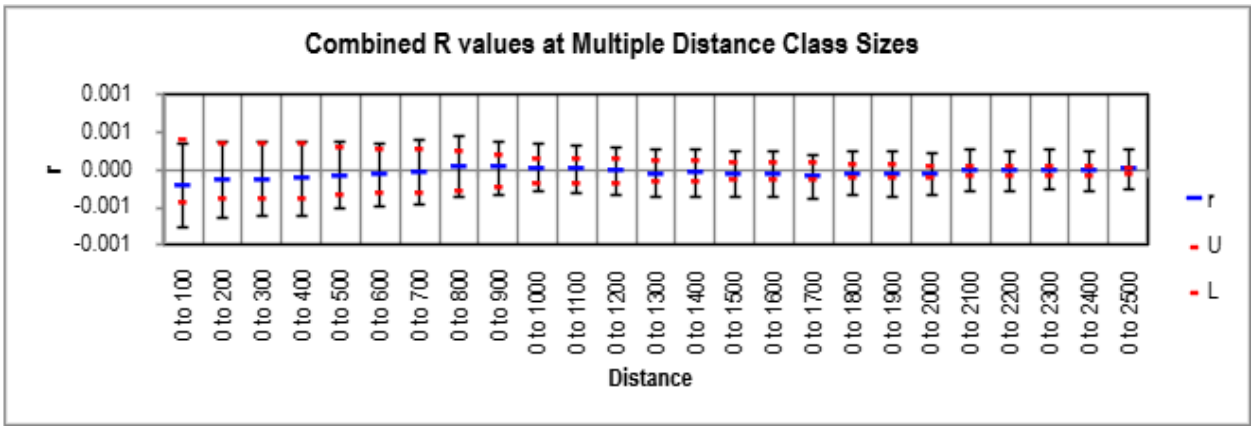


Figure 3.10. Spatial autocorrelation analysis of greater amberjack in the southeastern United States. Correlograms illustrating the influence of the size of the distance class on spatial autocorrelation. X-axis, distance class; y-axis, spatial autocorrelation (r). Spatial autocorrelation (r) is computed when the first distance class is increased in increments of 100 km. 95% confidence error bars for r were estimated by bootstrapping over pairs of samples; dashed lines represent upper and lower bounds of a 95% CI for r generated under the null hypothesis of a random geographic distribution of greater amberjack genotypes. **(a)** 2022 subdataset, **(b)** 2023 subdataset.

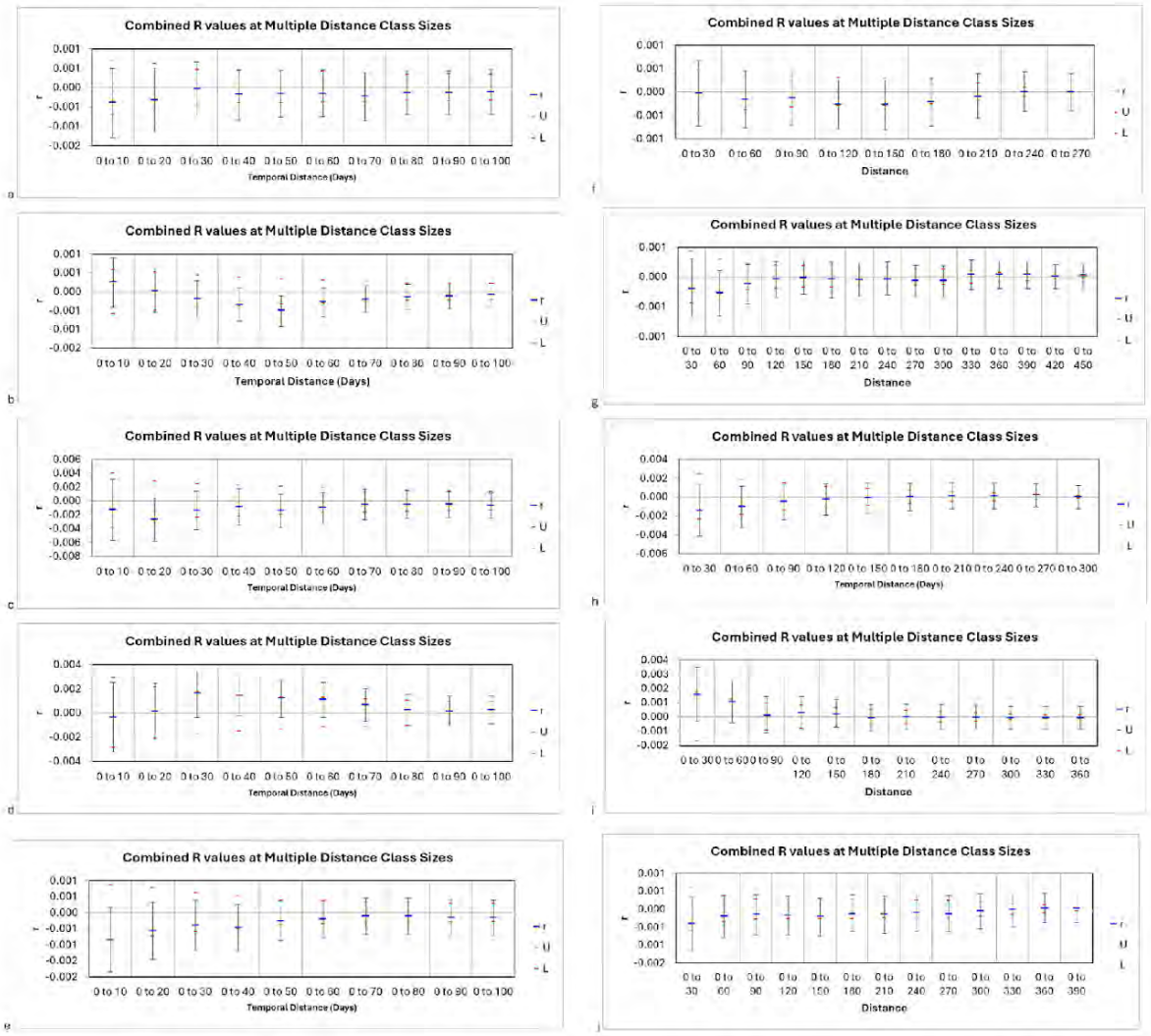


Figure 3.11. Temporal autocorrelation analysis in five subregions of the southeastern United States. Correlograms illustrating the influence of the collection date interval on autocorrelation. X-axis, temporal distance class; y-axis, autocorrelation (r). Autocorrelation (r) is computed when the first distance class is increased in increments of 10 days (a-e), and in increments of 30 days (f-j). 95% confidence error bars for r were estimated by bootstrapping over pairs of samples; dashed lines represent upper and lower bounds of a 95% CI for r generated under the null hypothesis of a random temporal distribution of greater amberjack genotypes. **(a)** Western Gulf (10-day), **(b)** Central Gulf (10-day), **(c)** Southwest Florida (10-day), **(d)** East Florida (10-day), **(e)** Carolinas (10-day), **(f)** Western Gulf (30-day), **(g)** Central Gulf (30-day), **(h)** Southwest Florida (30-day), **(i)** East Florida (30-day), **(j)** Carolinas (30-day).

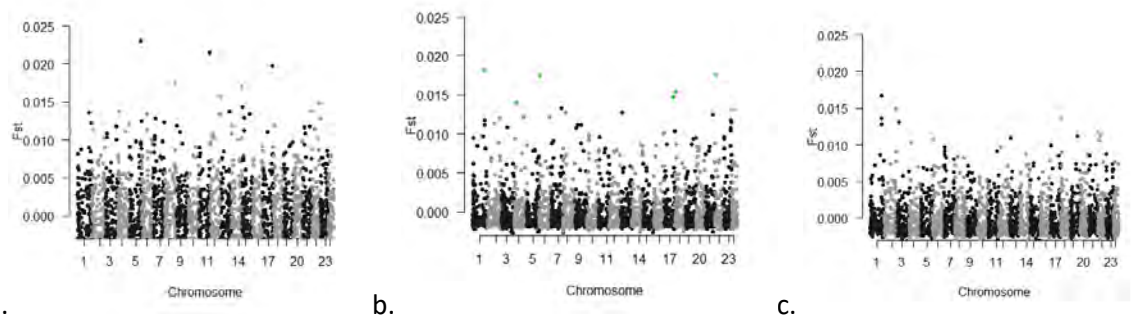


Figure 3.12. Manhattan plot visualization of outlier analysis results. Manhattan plots depicting F_{ST} of SNP loci positioned on the 24 greater amberjack chromosomes. Analysis comparing the **(a)** 10 sampling regions (outliers=0) described in Figure 3.2, **(b)** Gulf vs Atlantic (outliers=8), and **(c)** a subdivided Gulf vs Atlantic (outliers=0). Significant outlier loci inferred in outFLANK are highlighted in green.

References

Astashyn, A., E. S. Tvedte, D. Sweeney, V. Sapojnikov, N. Bouk, V. Joukov, E. Mozes, P. K. Strope, P. M. Sylla, L. Wagner, S. L. Bidwell, L. C. Brown, K. Clark, E. W. Davis, B. Smith-White, W. Hlavina, K. D. Pruitt, V. A. Schneider, and T. D. Murphy. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biology 25(1).

Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Page Source: Journal of the Royal Statistical Society. Series B (Methodological).

Browning, B. L., and S. R. Browning. 2011. A Fast, Powerful Method for Detecting Identity by Descent. The American Journal of Human Genetics 88(2):173–182.

Caye, K., T. M. Deist, H. Martins, O. Michel, and O. François. 2016. TESS3: Fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources 16(2):540–548.

Chen, L., Y. Jiang, B. Tao, Y. Song, Y. Huang, G. Li, W. Hu, C. Zhu, and J. Chen. 2025. High-quality Chromosome-level genome assembly of male and female greater amberjack (*Seriola dumerili*). Scientific Data 12(1).

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34(17):i884–i890.

Chen, Y., Y. Zhang, A. Y. Wang, M. Gao, and Z. Chong. 2021. Accurate long-read de novo assembly evaluation with Inspector. Genome Biology 22.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin. 2011. The variant call format and VCFtools. Bioinformatics 27(15):2156–2158.

Dimens, P., L. Reynal, C. Pau, F. Arocha, F. Hazin, P. Roque, N. J. Cummings, J. S. Franks, K. L. Jones, and E. A. Saillant. 2023. Weak range-wide population structure in the blackfin tuna (Thunnus atlanticus) revealed by analysis of genome-wide SNPs. ICES Journal of Marine Science 80(4):941–952.

Dimens, P. V. 2022. pdimens/LepWrap: link with zenodo.

Dimens, P. V., K. L. Jones, D. Margulies, V. Scholey, S. Cusatti, B. McPeak, T. Hildahl, and E. A. E. Saillant. 2024. Genomic resources for the Yellowfin tuna Thunnus albacares. Molecular Biology Reports 51.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10(3):564–567.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131(2):479–491.

Galbraith, J. K., K. E. Bemis, W. E. Bemis, H. S. Cook, and M. J. Wuenschel. 2022. NOAA Professional Paper NMFS 22 Identifications, distributions, and life history of four species of *Seriola* (Carangiformes: Carangidae) in the western North Atlantic based on contemporary and historical data.

Garrison, E., and G. Marth. 2012. Haplotype-based variant detection from short-read sequencing.

Girgis, H. Z. 2015. Red: An intelligent, rapid, accurate tool for detecting repeats de-novo on the genomic scale. BMC Bioinformatics 16.

Gold, J. R., and L. R. Richardson. 1998. Mitochondrial DNA diversification and population structure in fishes from the Gulf of America and Western Atlantic. Journal of Heredity 89(5):404–414.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29(8):1072–1075.

Haghshenas, E., H. Asghari, J. Stoye, C. Chauve, and F. Hach. 2020. HASLR: Fast Hybrid Assembly of Long Reads. iScience 23(8).

Hall, T. A. 1999. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series (41):95–98.

Han, S., J. Wang, J. Zhang, Y. Wang, Y. Luo, Q. Liu, and L. Chen. 2025. GWAS and selective sweep analysis reveal the genetic basis of cold tolerance in the domesticated pufferfish (Takifugu obscurus). Aquaculture 598:742018.

Hargrove, J. S., D. J. Murie, D. C. Parkyn, E. V. Saarinen, and J. D. Austin. 2018. Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico. Genetica 146(4–5):393–402.

Hildahl, T., and E. Saillant. 2025. A chromosome-anchored reference assembly for the gray snapper, Lutjanus griseus. Molecular Biology Reports 52(1):99.

Hu, J., J. Fan, Z. Sun, and S. Liu. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics 36(7):2253–2255.

Ivanova, N. V., T. S. Zemlak, R. H. Hanner, and P. D. N. Hebert. 2007. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7(4):544–548.

Jombart, T. 2008. adegenet : a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11):1403–1405.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11(1):94.

Kerfeld, C. A., and K. M. Scott. 2011. Using BLAST to teach “E-value-tionary” concepts. PLoS Biology 9(2).

Kolmogorov, M., J. Yuan, Y. Lin, and P. A. Pevzner. 2019. Assembly of Long Error-Prone Reads Using Repeat Graphs. Nature Biotechnology 37:540–546.

Kosambi, D. 1943. The estimation of map distances from recombination values. Ann Eugen (12):172–175.

Leal, T. P., V. C. Furlan, M. H. Gouveia, J. M. Saraiva Duarte, P. A. Fonseca, R. Tou, M. de O. Scliar, G. S. de Araujo, L. F. Costa, C. Zolini, M. G. C. D. Peixoto, M. R. S. Carvalho, M. F. Lima-Costa, R. H. Gilman, E. Tarazona-Santos, and M. R. Rodrigues. 2022. NAToRA, a relatedness-pruning method to minimize the loss of dataset size in genetic and omics analyses. Computational and Structural Biotechnology Journal 20:1821–1828.

Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26(5):589–595.

Manichaikul, A., J. C. Mychaleckyj, S. S. Rich, K. Daly, M. Sale, and W.-M. Chen. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics 26(22):2867–2873.

McCleelan, D. B., and N. J. Cummings. 1997. Preliminary Analysis of Tag and Recapture Data of the Greater Amberjack, *Seriola dumerili*, in the Southeastern United State. Pages 25–45 Proceedings of the Gulf and Caribbean Fisheries Institute.

Md, V., S. Misra, H. Li, and S. Aluru. 2019. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems.

Morton, N. E. 1995. Sequential Tests for the Detection of Linkage. American Journal of Human Genetics (7):277–318.

Norrell, A. E., K. L. Jones, and E. A. Saillant. 2020. Development and characterization of genomic resources for a non-model marine teleost, the red snapper (Lutjanus campechanus, Lutjanidae): Construction of a high-density linkage map, anchoring of genome contigs and comparative genomic analysis. PLoS ONE 15(4).

Pante, E., and B. Simon-Bouhet. 2013. marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. PLoS ONE 8(9):e73051.

Patterson, N., A. L. Price, and D. Reich. 2006. Population Structure and Eigenanalysis. PLoS Genetics 2(12):e190.

Peakall, R., M. Ruibal, and D. B. Lindenmayer. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. Evolution 57(5):1182–1195.

Peakall, R., and P. E. Smouse. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6(1):288–295.

Peakall, R., and P. E. Smouse. 2012. Genalex 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28:2537–2539.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLoS ONE 7(5).

Pryszcz, L. P., and T. Gabaldón. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Research 44(12):e113–e113.

Purcell, C. M., C. L. Chabot, M. T. Craig, N. Martinez-Takeshita, L. G. Allen, and J. R. Hyde. 2015. Developing a genetic baseline for the yellowtail amberjack species complex, *Seriola* lalandi sensu lato, to assess and preserve variation in wild populations of these globally important aquaculture species. Conservation Genetics 16(6):1475–1488.

Puritz, J. B., C. M. Hollenbeck, and J. R. Gold. 2014. dDocent: a RADseq , variant-calling pipeline designed for population genomics of non-model organisms. PeerJ.

Raj, A., M. Stephens, and J. K. Pritchard. 2014. FastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197(2):573–589.

Rastas, P. 2017. Genome analysis Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. Bioinformatics 33(23):3726–3732.

Rastas, P. 2019. Lep-Anchor: Automated construction of linkage map anchored haploid genomes. Bioinformatics 36(8):2359–2364.

Rastas, P., L. Paulin, I. Hanski, R. Lehtonen, and P. Auvinen. 2013. Lep-MAP: fast and accurate linkage map construction for large SNP datasets. Bioinformatics 29(24):3128–3134.

Raymond, M., and F. Rousset. 1995. An Exact Test for Population Differentiation 49(6):1280–1283.

Salzberg, S. L., A. M. Phillippy, A. Zimin, D. Puiu, T. Magoc, S. Koren, T. J. Treangen, M. C. Schatz, A. L. Delcher, M. Roberts, G. Marçais, M. Pop, and J. A. Yorke. 2012. GAGE: A critical evaluation of genome assemblies and assembly algorithms. Genome Research 22(3):557–567.

Sarropoulou, E., A. Y. M. Sundaram, E. Kaitetzidou, G. Kotoulas, G. D. Gilfillan, N. Papandroulakis, C. C. Mylonas, and A. Magoulas. 2017. Full genome survey and dynamics of gene expression in the greater amberjack *Seriola dumerili*. GigaScience 6(12):1–13.

Schultz, M. K., M. A. Dance, S. Powers, and S. R. Midway. 2025. Global Review of Greater Amberjack (*Seriola dumerili*) Life History. Reviews in Fisheries Science & Aquaculture:1–22.

SEDAR 59. 2020, March. SEDAR – South Atlantic Greater Amberjack Stock Assessment Report. North Charleston.

SEDAR 70. 2020. SEDAR - Gulf of Mexico Greater Amberjack. North Charleston.

Šegvić-Bubić, T., F. Marrone, L. Grubišić, D. Izquierdo-Gomez, I. Katavić, M. Arculeo, and S. Lo Brutto. 2016. Two seas, two lineages: How genetic diversity is structured in Atlantic and Mediterranean greater amberjack *Seriola dumerili* Risso, 1810 (Perciformes, Carangidae). Fisheries Research 179:271–279.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31(19):3210–3212.

Smouse, P. E., and R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82(5):561–573.

Takahashi, H., T. Kurogoushi, R. Shimoyama, and H. Yoshikawa. 2021. First report of natural hybridization between two yellowtails, *Seriola* quinqueradiata and S. lalandi. Ichthyological Research 68(1):139–144.

Tang, H., X. Zhang, C. Miao, J. Zhang, R. Ming, J. C. Schnable, P. S. Schnable, E. Lyons, and J. Lu. 2015. ALLMAPS : robust scaffold ordering based on multiple maps. Genome biology:1–15.

Thorburn, D. J., K. Sagonas, M. Binzer-Panchal, F. J. J. Chain, P. G. D. Feulner, E. Bornberg-Bauer, T. B. H. Reusch, I. E. Samonte-Padilla, M. Milinski, T. L. Lenz, and C. Eizaguirre. 2023. Origin matters: Using a local reference genome improves measures in population genomics. Molecular Ecology Resources 23(7):1706–1723.

Vaser, R., I. Sović, and N. Nagarajan. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Research 27(5):737–746.

Walker, B. J., T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. Zeng, J. Wortman, S. K. Young, and A. M. Earl. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE 9(11).

Wang, S.-L., L.-F. Luo, Y. Yu, Y.-Y. Fu, and Z.-X. Gao. 2025. Whole-genome re-sequencing reveals genetic structure and selection signals of different populations in Megalobrama amblycephala. Aquaculture 595:741548.

Wells, R. J. D., and J. R. Rooker. 2004. Distribution, age, and growth of young-of-the-year greater amberjack (*Seriola dumerili*) associated with pelagic Sargassum. Fishery Bulletin 102(3):545.

Whitlock, M. C., and K. E. Lotterhos. 2015. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of F(ST). The American Naturalist 186.

Young, S. S. 2024, December 14. Estimating Fishing Mortality and Movement of Greater Amberjack in the Gulf of Mexico and the Atlantic Ocean off of the Southeastern United States. Auburn University, Auburn.

Zhou, C., S. A. McCarthy, and R. Durbin. 2023. YaHS: yet another Hi-C scaffolding tool. Bioinformatics 39(1).

Zimin, A. V., and S. L. Salzberg. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLOS Computational Biology 16(6):e1007981.

4. Environmental DNA

A novel ddPCR™ assay for eDNA detection and quantification of Greater Amberjack *Seriola dumerili* and three congeners in US waters: challenges and application to fisheries independent surveys

P Joana Dias, Nicole M. Phillips

Abstract

Environmental DNA (eDNA) tools aimed at the detection and quantification of target species are emerging in support of marine fisheries surveys. Key to progressing the field is Droplet Digital™ PCR (ddPCR™), a highly sensitive technique with advanced multiplexing and direct quantification capabilities, able to provide fisheries scientists with improved interpretation of eDNA data. Here, we developed and validated a novel tetraplex ddPCR™ assay able to detect and distinguish between *Seriola dumerili*, S. *fasciata*, S. *rivoliana*, and S. *zonata* from seawater eDNA samples. To groundtruth eDNA ddPCR™ data, we compared eDNA ddPCR™ detections and quantifications to abundance data inferred from multiple camera (ROV, S-BRUV, chevron trap) and acoustic (VPS array) gears deployed during a fisheries research gear-calibration cruise in Alabama waters in 2022. Both eDNA and camera gears detected the presence of *S. dumerili* and S. *rivoliana* at both sites and all sampling days, but not S. *fasciata* and S. *zonata*. EDNA concentrations were higher for *S. dumerili* than S. *rivoliana* at both sites for all sampling days, in line with abundance patterns obtained from camera gears. The highest correlation (r = 0.97) was obtained between the measures of eDNA concentrations sampled 'between' other gear deployments and ROV estimates of abundance. This high correlation adds to a growing number of studies demonstrating the potential of eDNA as an indicator of abundance for fisheries stock assessments. To further explore the potential of integrating eDNA approaches into routine fisheries surveys, follow-up eDNA surveys were conducted in 2023 during stock assessment cruises in Alabama. During these surveys, a preliminary investigation into the performance of passive eDNA samplers (i.e., Speci-sponges) found the passive sampling approach yielded more positive detections for *S. dumerili* and S. *rivoliana* compared to niskin water sampling with filtration. This suggests that passive sampling approaches may improve on species eDNA detections, owing to the ability for the sampler to 'go to the fish' via deployment on the ROV, combined with longer deployment/soak times. This finding is preliminary due to small sample sizes, but is a promising, cost-efficient outcome to support the future development of streamlined eDNA sampling methods. Incorporating eDNA in fisheries surveys could improve precision in species detection and abundances, being an especially valuable complement to camera gears deployed in geographic areas or seasons with poor visibility conditions. Investment into follow-up studies with increased sample sizes, and spatial and temporal replication are necessary to validate the robustness of the eDNA quantification estimates and abundance correlations observed in this study and to further examine the potential of passive sampling to estimate abundance of marine fishes.

Introduction

After years of attracting attention as a promising tool (Hansen et al. 2017), environmental DNA (eDNA) is now being explored with the experimental dimension necessary to improve the cost-efficiency of fisheries research survey work (Kelly et al. 2023, Saborido-Rey et al. 2023). EDNA provides a non-invasive alternative for surveying aquatic communities with a demonstrated capacity to inform on

species biodiversity, distribution, seasonal migration patterns, and abundance/biomass (Rourke et al. 2022, Kelly et al. 2023, Saborido-Rey et al. 2023). However, to date, very few species-specific eDNA assays have been developed and applied to the detection and quantification of eDNA from target species from studies that parallel marine fisheries surveys (Kirtane et al. 2021, Shelton et al. 2022, Ledger et al. 2024, Maes et al. 2024).

The reliable capture and detection of eDNA in highly dynamic coastal and open-ocean environments and the use of eDNA to estimate fish abundance and biomass remain two of the most challenging goals of marine eDNA research (Yates et al. 2019, Ramirez-Amaro et al. 2022, Rourke et al. 2022, Kelly et al. 2023). EDNA shed rate, transport and degradation are complex and are influenced by biotic factors such as spawning season, size, and life-stage, and abiotic factors such as currents, water temperature, salinity, oxygen, and pH (Jerde 2019, Takeuchi et al. 2019, Lamb et al. 2022). Because eDNA shed rate is taxa and life-stage specific, and eDNA transport and degradation vary depending on season and geographic area, understanding these processes in each fisheries context is essential for planning efficient sampling, circumventing challenges, and meeting research aims. The need for more empirical studies (with increased temporal sampling and replication) on how eDNA in marine environments correlates to the presence, abundance, and biomass of fish stocks has, therefore, been pointed as the biggest caveat to progressing the field of eDNA research in applied fisheries science (Kirtane et al. 2021, Maes et al. 2024).

Key to progressing the integration of eDNA as a fisheries management tool is also Droplet Digital™ PCR (ddPCR™), which has emerged as an improved, user-friendly, and cost-efficient alternative to real-time PCR (qPCR), with increased multiplexing capabilities, and demonstrated advantages in sensitivity and quantification (Nathan et al. 2014, Doi et al. 2015, Wood et al. 2019, Huggett et al. 2020). DdPCR™ partitions samples into thousands of nanoliter droplets (~20,000), with PCR-amplification of target DNA occurring within each individual droplet. DdPCR™ software quantifies the number of positive and negative droplets for each fluorophore in each sample and fits the fraction of positive droplets to a Poisson algorithm to determine the starting concentration of the target DNA molecule in units of copies/µL (Huggett et al. 2020). By providing direct quantification, ddPCR™ has the potential to provide fisheries scientists with improved and straightforward interpretation of eDNA data. Despite the advantages, only one study has been published to date applying ddPCR™ to marine fisheries research (Maes et al 2024) and none exploring its advanced multiplexing capabilities.

Objectives

In the present study, we aimed to: 1) develop and validate a tetraplex ddPCR™ assay able to detect and distinguish between *S. dumerili*, S. *fasciata*, S. *rivoliana*, and S. *zonata*. By comparing ddPCR™ detections and quantifications to relative abundance data inferred from multiple camera and acoustic gears deployed during a fisheries research gear-calibration cruise in 2022, we further aimed to: 2) groundtruth ddPCR™ data, and 3) explore the capacity of eDNA to provide abundance estimates and support fisheries independent surveys. For eDNA to become scalable and see widespread use in fisheries surveys, sample collection and processing need to be streamlined for use by non-experts with the ability to be integrated into routine stock assessment surveys. Therefore, in 2023, we conducted a follow-up study during fisheries monitoring cruises routinely conducted by the fisheries team at Dauphin Island Sea Lab to: 4) determine the feasibility of conducting 'contamination-free' eDNA surveys during a routine stock assessment cruise, 5) explore the use of simple passive eDNA samplers as an alternative to

the 'traditional' water sample collection and filtration techniques used in 2022 surveys, and 6) test the robustness of *Seriola* spp. species-specific ddPCR™ detections. By detailing assay development and validation, best practices in the prevention of contamination, and results of basic in situ experiments in the field, we also provide useful guidance for future studies using ddPCR™ applied to marine fisheries eDNA research.

Methods

EDNA contamination controls and best practice protocols

All work was conducted under strict procedures to reduce the risk of contamination by exogenous DNA and cross-contamination between samples in both the field and laboratory (Goldberg et al. 2016, Bockrath et al. 2022). All genomic DNA (gDNA) and eDNA laboratory work was temporally and spatially separated and followed a unidirectional workflow. GDNA and eDNA samples were extracted in separate laboratories and all gDNA, synthetic DNA (sDNA), and eDNA samples, extracts, and aliquots were stored in separate rooms. All PCR reagents and aliquots were stored in a dedicated freezer, separate from all DNAs. Reaction setup for gDNA and sDNA PCR amplifications were performed under a PCR hood in a separate room from the PCR hood and ddPCR™ equipment used to perform eDNA PCR amplifications. Separate pipette sets were maintained in each laboratory space, with project-specific pipette sets dedicated to eDNA extractions and PCR setups. Barrier filter tips were used when conducting all reagent aliquoting, PCR, and eDNA work. Personal hygiene (showering and changing clothes between laboratory days and between working with gDNA, sDNA or eDNA), changing of disposable gloves, and deep cleaning laboratory spaces between temporally separate project phases (gDNA, sDNA, eDNA) were also implemented as a best practice. Benches were regularly cleaned with 10% bleach (before, between, and after work), all reusable equipment was soaked in 10% bleach for at least 20 minutes and rinsed thoroughly with deionized water, and all sensitive equipment was cleaned with DNA AWAY™ Surface Decontaminant (Thermo Scientific), or a CloroxPro™ Healthcare® Bleach Germicidal Wipes and UV sterilized for at least 20 minutes.

All field equipment was handled, cleaned, and stored in a dedicated room in a separate building from all genetic work. Water collection and filtering equipment were cleaned ahead of fieldwork and between uses in the field using a combination of two methods of sterilization: cleaning with 10% bleach followed by either autoclaving at 120°C for 20 min or exposure to UV light for 20 min, depending on the materials. Personal hygiene between fieldwork days (showering, clean clothes), changing of disposable gloves and separation of contamination-prone and clean areas both at sea and on shore were maintained throughout field activities (Drymon et al. 2021, Lehman et al. 2022). No fieldwork equipment was exposed to any target species for the duration of this study. Negative controls were incorporated into each stage of sample processing to test for the possibility of contamination, as detailed below.

Aim 1: Development of a tetraplex ddPCR™ assay

One set of degenerate primers (Eurofins Scientific, Luxemburg) and four species-specific internal PrimeTime® double-quenched ZEN™/IOWA Black™ FQ probes (Integrated DNA Technologies, Coralville, US) were manually designed in two mitochondrial DNA (mtDNA) loci, cytochrome oxidase c subunit 1 (COI) and NADH dehydrogenase 5 and 6 (ND5-ND6), of the four target species (Klymus et al. 2020, BioRad Droplet Digital PCR Applications Guide Bulletin 6407). Briefly, sequences of the four target

species and genetically similar co-occurring fish species were downloaded from GenBank and aligned with BioEdit© (Hall 1999, Alzohairy 2011). Assays designed for the mtDNA ND5-ND6 region were primarily based on the 39 sequences generated by Renshaw and Gold (2009) for the four target *Seriola* species in the U.S. Atlantic Ocean and Gulf of Mexico (GenBank Accession numbers GU014709–GU014747). Primers were designed in regions conserved for the four target species but variable across other genetically similar and co-occurring fish species, while probes were designed to include base pair (bp) differences between all species. All primer and probe sets were checked for oligo dimers using the Thermo Fisher Multiple Primer Analyzer web tool and checked for specificity in silico using NCBI BLAST (Altschul et al. 1991, Syngai et al. 2013).

All primer sets were first tested for specificity in vitro with qPCR on a Bio-Rad® C1000 Touch™ Thermal Cycler (CFX96 Optics Module, instrument no. 785BR21404) using 10-fold dilutions starting from 5–25 ng of DNA (undiluted, 1:10, 1:100, 1:1000) of gDNA extracted from one individual of each of the four target species and 25 exclusion species (Table 4.1). Reactions contained 1X Power SYBR® Green PCR Master Mix (Life Technologies LTD, UK) and 200 nanomolar (nM) of each primer, adjusted to 20 µL using PCR-grade water. Cycling conditions consisted of 95 °C for 10 min and 44 cycles of: 94 °C for 15 s, 55 to 60 °C for 30 s and 60 °C for 1 min, followed by melting curve analysis between 65 °C and 95 °C with a plate read every 0.5°C after holding the temperature for 5 s. The primer set that best amplified DNA from the target species, but not from exclusion species, was subsequently tested on all target and exclusion individuals with the respective species-specific probes using the optimized ddPCR™ conditions. All qPCR and ddPCR™ reactions, including the non-template negative controls (NTC, PCR-grade water added instead of template) were conducted in duplicate. All qPCR and ddPCR™ runs were performed separately on target and non-target species, with no positive controls added to prevent cross-contamination.

DdPCR™ optimization

DdPCR™ reactions and cycling conditions were optimized using gDNA for each of the four target species and amplitude multiplexing on the Bio-Rad® QX200™ AutoDG™ Droplet Digital™ PCR System (Automated Droplet Generator instrument no. 773BR3222, T100™ Thermal Cycler instrument no. 621BR18911, QX200™ Droplet Reader instrument no. 771BR1496, Bio-Rad, Hercules, CA, USA). The QX200 droplet reader enables detection of fluorescence in two different channels (FAM and VIC/HEX). Primer and probe concentrations, annealing temperature, ramp rate, and cycle number were adjusted to produce positive clusters of droplets with high relative fluorescence units (RFUs), optimal separation of clusters corresponding to each pair of probes within each channel, and little to no droplet "rain" (Huggett et al. 2020, Lehman et al. 2020, Kokkoris et al. 2021). While other multiplexing strategies are available, tetraplexing using fluorescence amplitude provides a clearer output (von Ammon et al. 2022). The optimized ddPCR™ reactions contained 1X Bio-Rad® ddPCR™ Multiplex Supermix for probes, 910 nM of each primer, 227 nM of each probe, and 1 µL of ~0.2 ng DNA from each of the four species, adjusted with PCR-grade water to a final volume of 22 µL. Using an automated droplet generator, 20 µL of each ddPCR™ reaction was combined with ~70 µL of automated droplet generation oil for probes to create up to 20,000 nanoliter-sized droplets prior to PCR, as per the manufacturer protocol (BioRad Droplet Digital™ PCR Applications Guide Bulletin 6407). Optimal ddPCR™ cycling conditions consisted of an initial step at 95 °C for 10 min, followed by 40 cycles of: 94 °C for 30 s and 61 °C for 2 min, and a final step at 98 °C for 10 min, using a ramp rate of 1 °C/s.

To avoid incorrectly calling artifacts (i.e., droplets in the absence of target DNA resulting from assay components interacting in a way that causes premature probe cleavage), a plate with 48 NTC ddPCR™ reactions was run. This ddPCR™ run was prepared using freshly made aliquots of all reagents, after all gDNA samples were archived at -80°C and all laboratory spaces and equipment were deep cleaned. A NTC threshold and an optimal range of the positive droplet population were established for the optimized assay (Lehman et al. 2020, Kokkoris et al. 2021, Figure 4.1). All ddPCR™ data were analyzed using Absolute Quantification (ABS), the QX Manager 1.2 Standard Edition software and the guidelines for amplitude multiplexing on the BioRad Droplet Digital™ PCR Applications Guide Bulletin 6407.

DdPCR™ limit of quantification (LoQ) and limit of detection (LoD)

Assay sensitivity is determined by its limit of quantification (LoQ), indicated by the lowest target copy number in a sample that can be reliably quantified with an acceptable level of precision and accuracy, and limit of detection (LoD), which defines the assay's ability to detect the target sequence at low levels (Dobnik et al. 2016, Klymus et al. 2019). LoQ and LoD are generally determined using dilution series of synthetic DNA, as it can be less variable than gDNA, allowing for interlaboratory comparisons of assay performance (Langlois et al. 2020, Thomson-Laing 2021, Xia et al. 2021). In this study, we designed gBlocks™ (Integrated DNA Technologies, Coralville, USA) for a 700-bp gene region covered by the assay, with specific primer and probe binding sites for the four target species. To differentiate synthetic DNA from tissue-derived DNA in the event of cross-contamination, each gBlock™ was modified by reversing a 12-bp fragment between the forward and reverse primers of the assay, without overlapping into the probe region. The optimized tetraplex ddPCR™ assay was evaluated for sensitivity on a ten-fold series of 10X dilutions from a starting gBlock™ concentration of 2 ng/µL, in triplicate reactions. Target DNA was detected in all replicates for all target species down to 2x10-8 ng/µL. At the 2x10-9ng/µL dilution, target DNA was detected in only one or two of the three replicates across all four species. No target DNA was detected in the 2x1010 ng/µL dilutions.

To further refine the LoQ and LoD, ddPCR™ reactions were performed on a series of 2X dilutions consisting of 1x10-8, 0.5x10-8 and 0.25x10-8, and on 2x10-9 and 1x10-9. Because replicates with no detections add significant variability, most ddPCR™ assays establish their LoQ at the lowest copy number in the dynamic range where ≥ 90 - 100% of replicates are detected, or the coefficient of variation (CV) of the measured copy number ≤ 25 - 35% (Dobnik et al. 2016, Mauvisseau et al. 2019, Thomson-Laing et al. 2021). In this study, target DNA of all four species was efficiently (Figure 4.1) detected in all three replicates down to 0.5x10-8 ng/µL, however at this lower end the CV was ≥55% across the four target species. For this reason, we set the LoQ at 1x10-8 ng/µL corresponding to 0.35 copies/µL with a CV of 25% for *S. dumerili*, 0.64 copies/µL with a CV of 22% for S. *fasciata*, 0.47 copies/µL with a CV of 42% for S. *zonata*, and 0.49 copies/µL with a CV of 30% for S. *rivoliana*. Estimated eDNA concentrations below the LoQ should only be evaluated qualitatively, as positive detections or non-detections (Klymus et al. 2019).

QPCR guidelines (Bustin et al. 2009, Klymus et al. 2019) have been useful in guiding ddPCR™ LoQ, and the Minimum Information for the Publication of Digital PCR Experiments (dMIQE, Huggett et al. 2013) has been recently updated (Huggett et al. 2020). However, clear guidelines are still needed for ddPCR™ LoD determination to maximize the benefit of using a highly sensitive platform. The power to detect as little as a single copy of target DNA in one amplified droplet is a major strength of ddPCR™ and the driver behind the increased application of (including transfer of qPCR assays to) ddPCR™ in eDNA

research (Hunter et al. 2018, Dimond et al. 2022, King et al. 2022). In this study, to establish much needed reasoning and confidence in the 'one-droplet-power' of ddPCR™, we proposed a new approach to LoD. We also include efficiency curves which, although not needed for quantification purposes in ddPCR™, are still a useful performance indicator. In ddPCR™, below the LoQ and when there are far fewer target DNA copies than droplets, a single droplet partition will contain no more than one target DNA copy at the start of the reaction (BioRad Droplet Digital PCR Applications Guide Bulletin 6407). Because each droplet represents its own PCR reaction, ddPCR™ amplification efficiency, accuracy, repeatability, and overall reliability can be (re)determined at the droplet level (Figure 4.1). Below the LoQ, in the 0.25x10-8, 2x10-9, and 1x10-9 dilutions, one, two, and three ddPCR™ droplets were detected in at least one replicate across all four species. For these one, two, and three droplets detections, ddPCR™ efficiency was high R2= 0.999 and the 'one droplet/target DNA copy' LoD corresponded to 0.065 ± SE 0.003 target DNA copies/µL for *S. dumerili*, 0.063 ± SE 0.002 copies/µL for S. *fasciata*, 0.067 ± SE 0.003 copies/µL for S. *zonata* and 0.064 ± SE 0.001 copies/µL for S. *rivoliana* (Figure 4.1).

Aims 2, 3: eDNA validation & grouthtruthing surveys

To validate the developed ddPCR™ assay in situ and groundtruth eDNA field data, seawater sampling was conducted during a fisheries survey gear calibration cruise from August 29th – September 2nd, 2022. Near-concurrent eDNA, acoustic, and camera imagery data were collected from baited chevron trap cameras, stereo baited remote underwater video (S-BRUV), a remotely operated vehicle (ROV), and virtual positioning system (VPS) arrays. Camera and eDNA gears were allocated approximately two hours each for sampling and, for logistic reasons, distributed on two vessels (Figure 4.2). Twenty-four hours prior to fieldwork, a deep clean was conducted of the research vessel to prevent contamination of eDNA samples from previous cruises. Exterior working areas were sprayed with 10% bleach for ≥20 minutes, followed by a thorough freshwater rinse. Inside areas were cleaned with 10% bleach wipes.

All gears were deployed at two artificial reef 'super pyramids' (~8 m tall with a 4.5 m base) in ~30 m depths off the coast of Dauphin Island, Alabama (AL), U.S. (site 1 at 29° 52.785 N -87° 58.328 W and site 2 at 29° 47.725 N -87° 57.245 W). These two sites were chosen due to the relatively large number of *S. dumerili* present, and the success of acoustic tagging efforts, where four fish in good condition were affixed with acoustic transmitters (n = 41; Innovasea V16P, V9P, or V9, random delay from 60-120 seconds) and conventional tags (Floy Tag BFIM-96, double-head dart). The VPS arrays consisted of a ring of inner receivers (Innovasea VR2AR, n = 4) placed approximately 125 m from the reef, while an outer array of receivers (n = 4) was placed ~250 m from the inner receivers, resulting in ~1 km2 of VPS coverage between both sites. The two sites were visited on alternate days by both vessels, and gear deployment order and sampling vessel were randomly selected on all five days of the calibration cruise. Each camera sampling gear (ROV, S-BRUV, chevron trap mounted cameras) was deployed following standard survey-specific sampling protocols. ROV sampling followed Powers et al. (2018).

To account for the potential impact of the baited survey gears on eDNA data, seawater sampling was conducted each day at three timepoints: before, between, and after S-BRUV and chevron trap deployments. Seawater sampling for eDNA analysis occurred as close as possible to the other gears deployment/retrieval locations, and all time and GPS coordinates were recorded. During each sampling time point, two duplicate 10 L Niskin samplers attached to a metal frame were deployed using the vessel winch line and A-frame, followed by a Sea-Bird SBE25 CTD to collect environmental parameters (salinity,

temperature, oxygen) along the depth profile. A total of 15 seawater samples were collected in duplicate (i.e., a total of 30 x 10 L samples) over the five days of the cruise.

The Niskin samplers were cleaned with 10% bleach and rinsed with deionized water before, between, and after deployments. To test for contamination during sample collection, 10 L of autoclaved deionized water were processed through each Niskin sampler and filtered at the end of each sampling day. During the first four days of sampling, seawater samples were emptied into a sterile 10 L jerry can and stored at -20 °C until water filtration, which occurred within 18 h of collection in a designated water filtration laboratory, where fish tissue had never been present. Water samples were filtered using a bench top GAST® DOA-P704-AA vacuum pump (Gast, Benton Harbor, MI, USA) and a self-preserving filter housing (Smith-Root, WA, USA; Thomas et al. 2019) fitted with Whatman® (Maidstone, United Kingdom) 47 mm 1.2 micron (µm) nitrocellulose filters, which were replaced when clogged (1–2 filters were used for each 10 L water sample). On the last sampling day, seawater samples were filtered at-sea directly from each Niskin using a Grover-Go portable vacuum pump (Grover Scientific, QLD, Australia) and the Smith-Root self-preserving filter housings. All filter housings were stored short-term (≤6 months, Thomas et al. 2019) at room temperature per manufacturers' protocols until eDNA extraction.

EDNA extraction, ddPCR™ screening and data analysis

Total eDNA was extracted from a half portion of each filter using the QIAGEN® DNeasy® Blood & Tissue Kit (Hilden, Germany) following the Goldberg et al. (2011) protocol incorporating QIAshredder™ spin columns and a final elution step of 50 µL. During eDNA extraction, designated sterile forceps and blades were used to handle filters for each sample to prevent cross-contamination. Extraction negative controls containing only reagents were processed alongside the field samples. Field negative controls were extracted, screened, and deemed negative for contamination before eDNA extraction of field samples occurred. The optimized tetraplex ddPCR™ assay developed in this study was used to screen all eDNA samples and negative controls (field, extraction, PCR) for the presence of the four target species using 4.5 µL of template per reaction, conducted in triplicate, and therefore screening 27% of the total eDNA extract. All remaining half filters and eDNA extracts were stored at -80°C.

All ddPCR™ reactions were run on the BioRad QX200 AutoDG™ Droplet Digital™ PCR System and data were analyzed using Absolute Quantification (ABS) and the QX Manager 1.2 Standard Edition software, following guidelines on the BioRad Droplet Digital PCR Applications Guide Bulletin 6407. Field samples and negative controls were analyzed using a three-criterion approach to reduce the likelihood of false positives (Schweiss et al. 2020; Lehman et al. 2020). Samples were defined as positive detections if, in at least one ddPCR™ replicate per sample, droplet(s) were above the threshold established for NTCs, within the normal range of positive droplets for the assay, and the concentration of target DNA was at or above the LoD. For controls to be considered free from contamination, none of the criteria could be met.

Camera gear and VPS array data analysis

MaxN for each *Seriola* species and genus level MaxN for *Seriola* were determined for each ROV, S-BRUV and chevron trap deployment. For all camera types, MaxN for each species was defined as the maximum number of individuals of each *Seriola* species in any frame within the video, and genus level MaxN was defined as the maximum number of *Seriola* in any frame in the video. Data from the VPS array were downloaded and processed using Innovasea Fathom software to calculate virtual positions (latitude and

longitude) of acoustically tagged fish within the arrays across the five-day sampling period. The abundance of *S. dumerili* at the two sites was then estimated using a mixed logit-normal mark-resight model (Program MARK). For model inputs, the number of acoustically tagged fish in the core region (within approximately 125 m from reef) at each site was determined from the VPS position data, and we used MaxN counts of confirmed tagged and confirmed untagged individuals from the ROV videos for the resighting events. The numbers of tagged fish seen on the other camera gears was too low and inconsistent to use for the analysis.

Statistical analysis

Fisher's exact tests were conducted to evaluate whether detection probabilities were statistically distinguishable between all the gears used. To quantify the relationship between quantitative eDNA estimates of *S. dumerili* with other acoustic and camera gears abundance measures, we calculated Spearman correlation, using a Bonferroni correction to adjust for the multiple comparisons. Spearman correlation is a nonparametric measure of the relationship between two measures and is calculated as the Pearson correlation of ranks of the measures rather than their values, which can be useful when data are skewed or have outliers.

Aims 4, 5, 6: Simplifying eDNA sampling for integration into routine stock assessment research

Follow-up eDNA surveys were conducted in 2023 during stock assessment cruises in Alabama to further explore the potential of integrating eDNA approaches into routine fisheries surveys, including preliminary tests comparing the performance of passive eDNA samplers to the water sample collection and filtration techniques established during 2022 gear calibration surveys. Although the traditional eDNA sampling was completed quickly (i.e., in only 15 minutes) during 2022 surveys, the multi-step process and number of supplies required makes eDNA sampling cumbersome and/or prone to contamination via exogenous DNA or cross-contamination between samples, rendering the data unreliable.

Stock assessment cruises with eDNA integration were conducted in spring (March, April) and fall (September) 2023, outside of the *S. dumerili* fishing season. To assess the viability of integrating eDNA sampling onboard a fisheries cruise (rather than a dedicated eDNA cruise) no 'deep cleaning' of the vessel was conducted prior to eDNA sampling, notwithstanding routine cleaning completed by the crew after each outing. All eDNA work otherwise followed the protocols developed during 2022 surveys to reduce the risk of contamination, but without interfering with the established operations and schedule. Near-concurrent eDNA, acoustic, and camera imagery data were collected from an EK80 echosounder, ROV, and, when time allowed, S-BRUV.

Seawater collection for eDNA analysis was conducted using two duplicate 10 L Niskin samples deployed sequentially using the vessel deep sea fishing rod and line. Sampling occurred down current from the reef to avoid entanglement, and as close as possible to the other gear deployment/retrieval locations. Water samples were filtered directly from each Niskin using a Grover-Go portable vacuum pump and Smith-Root self-preserving filter housings, as outlined for 2022 surveys, with filters stored at room temperature until DNA extraction. For comparisons between eDNA sampling approaches, passive eDNA sampling was conducted using Whirlpak® Cellulose Speci-sponges (Nasco Sampling, Pleasant Prairie, Wisconsin, U.S.). These passive eDNA samplers were deployed in duplicate by attaching with clean cable ties to each side of the ROV, and 'soaked' for at least 15 minutes. New sterile gloves were used to

deploy and retrieve the passive samplers, which were stored in a sterile plastic bag and frozen until DNA extraction. To assess the possibility of contamination while the Speci-sponges were handled on-board the vessel, a negative control was processed each sampling day, consisting of one sponge being removed from the sterile bag using sterile gloves and saturated with sterile water. The sponge was then returned to the bag and frozen until lab processing.

Lab processing and data analysis for filters followed previously described methods developed during 2022 surveys. The Speci-sponge samples were extracted using similar methods, with the exception that the Qiagen® DNeasy® Blood & Tissue Maxi Kit (Hilden, Germany) was used to accommodate larger sample volumes. DNA extracts were screened and analyzed using the optimized tetraplex ddPCR™ assay and quantitative criteria designed as part of this study.

Results

Aim 1: Assay development

Assay design options were limited for COI and primers were less specific during both in silico and in vitro testing, therefore, COI was ruled out as a viable target gene. From the two ND5-ND6 primer sets designed, one cross-amplified DNA from all exclusion species at <40 PCR cycles, even at higher annealing temperatures (Ta 60°C), and was therefore eliminated. The selected primer set (Table 4.2) successfully amplified the 124-bp target locus in *Seriola* spp. but did not cross-amplify DNA from exclusion species. The addition of the four species-specific probes provided further specificity of the assay, with each amplifying DNA from a single target *Seriola* species. From all exclusion species tested (Table 4.1), only S. lalandi produced non-specific amplification droplets via the *S. dumerili* probe.

To design the best possible assay, two of the probe oligos were tested using both FAM and HEX fluorophores. Compared to HEX, the FAM fluorophore increased RFUs by up to 12000, which is in line with this being primarily advised for ddPCR™ experiments (BioRad Droplet Digital™ PCR Applications Guide Bulletin 6407). On the HEX/VIC channel, IDT SUN (VIC equivalent) fluorophores provided slightly better amplitude differences between droplet clusters than HEX. However, detections on the HEX/VIC channel were never above 6000 RFUs, limiting the fluorescence amplitude between the two target species on this channel. This issue has been reported in similar work, where droplets on the edge of the range for a target species could be incorrectly assigned to another species, resulting in false positive detections (Dobnik et al. 2016, von Ammon et al. 2022). To minimize this issue, distinct fluorophores were designated for the detection of the two most common species *S. dumerili* and S. *rivoliana*. On the FAM channel, the optimal range of positive droplets for *S. dumerili* and S. *fasciata* was 2500 to 4500 RFUs and 12000 to 14000 RFUs above the average RFUs of the negative cluster, respectively. On the VIC channel, the optimal range of positive droplets for S. *zonata* and S. *rivoliana* was 1500 to 3500 RFUs and 4000 to 6000 RFUs above the average RFUs of the negative cluster, respectively (Figure 4.3). To guard against false positive detections of S. *rivoliana*, when positive droplets generated on the VIC channel fell towards the lower end of its positive droplet RFU range, the ddPCR™ was repeated using only the S. *zonata* probe to rule out the presence of this species.

The highest amplitude of an artifact droplet of the optimized assay across the 48 NTC ddPCR™ reactions was 500 RFUs above the negative droplets cluster across the FAM and VIC channels. Therefore, to be conservative, 1000 RFUs above the average RFUs of the negative droplet clusters was chosen as the NTC threshold for this assay (Figure 4.3). The absence of dimers between oligos (primers and probes) in this

assay facilitated adequate fluorescence amplitude between negative and positive droplet clusters for all target species. In addition, despite the ability to multiplex on the QX200 using the Bio-Rad® ddPCR™ supermix for probes, the use of the ddPCR™ Multiplex Supermix for probes (a concentrated version developed for the QX600) allowed for a 1000 – 2000 RFUs increase in probe amplitude.

Aims 2, 3: EDNA field sampling data

Niskin sampler deployments took an average of three minutes, and eDNA 'gear' operations, including the CTD cast, never exceeded 15 minutes of the two hours allocated for eDNA sampling. Gear deployment distance from the reef varied from 4 to 92 m (average 31.5 ± SD 23.8 m), with the vessel drifting 17 to 84 m between deployment and retrieval (average 38.3 ± SD 23.16 m). Environmental parameters at the depth of sampling throughout all five sampling days presented averages of 35.8 ± SD 0.09 salinity, 26.9 ± SD 0.25 temperature (°C), 4.2 ± SD 0.41 dissolved oxygen (mg/L) and 7.8 ± SD 0.03 pH.

Analysis of negative controls (field, extraction, PCR) found no evidence of target DNA across ddPCR™ replicates. We failed to include a filtration control in the laboratory, which means that if target DNA had been detected in the field negative controls, we would not have been able to determine if contamination had occurred during seawater sampling or filtering. Both processes were time consuming and prone to contamination during the first four days of sampling. On the fifth day of sampling, using a portable Grover pump for filtering the seawater samples on board directly from the Niskin samplers was faster, reduced the number of areas where samples were handled (vessel, storage freezer, laboratory) and equipment needed, greatly simplifying the process and reducing the potential for contamination.

From the 15 eDNA samples collected, 14 were positive for *S. dumerili* and 12 were positive for S. *rivoliana*. No positive detections were obtained for S. *fasciata* and S. *zonata*. In general, *S. dumerili* detections were above the LoQ at more sampling timepoints and DNA concentrations were greater than those for S. *rivoliana*, indicating a greater relative abundance of *S. dumerili*. EDNA detections varied depending on whether samples were collected before, between, or after other gears, particularly for *S. dumerili*. EDNA detections 'before' any baited gears entered the water were above the LoQ for only one or two out of the five sampling timepoints for S. *rivoliana* and *S. dumerili*, respectively (Figure 4.4). EDNA detections were then above the LoQ for *S. dumerili* for three of the five 'between' sampling timepoints, and all five 'after' deployment of baited gears sampling timepoints, although the concentrations of target DNA for the 'after' timepoints were only higher than other timepoints on two sampling days (Figure 4.4). In contrast, S. *rivoliana* eDNA was not detected in samples collected 'between' baited gears at site 1, and while it was detected each day at site 2, this detection was only above LoQ on one day of sampling. Samples collected 'after' all baited gears were deployed were above LoD for S. *rivoliana* at both sites and all sampling days, but were only above LoQ on one sampling day.

Camera gear and VPS array data

All camera gears made full use of the two hours allocated for sampling. All camera gears detected S. dumeril and S. *rivoliana* (when unsure these were assigned to *Seriola* genus level) at both sites and all five days of sampling, but did not detect S. *fasciata* or S. *zonata*, in agreement with the eDNA survey data. The highest MaxN counts observed during the study were 24 *S. dumerili* and four S. *rivoliana*, also in agreement with the eDNA data that *S. dumerili* were relatively more abundant. The VPS receivers confirmed that there were at least six *S. dumerili* on the sites throughout the five days. For site 1, six

individuals were observed in the core region on the first day of sampling, and seven were observed in this region from day two to five of sampling. For site 2, eight individuals were present in the core region (within ~125 m of the target reef) during all five days.

Statistical analysis

All the gears and eDNA measures were effective at detection of *S. dumerili*. All gears except SBRUVs, and all eDNA timepoints except one ('before') showed an abundance measure greater than zero on all five sampling occasions. The two exceptions detected the species on four of the five occasions. This difference (between 80% and 100% detectability on five sampling occasions) is not statistically distinguishable, as Fisher's exact test failed to reject a hypothesis of equal detection probabilities (p-value = 1). However, the similarity of the abundance measures themselves does vary.

The Spearman correlation matrix shows that the eDNA 'between' timepoint, ROV, and Chevron Trap abundance measures are closely aligned with each other, while the other two eDNA timepoints show little, or even negative, relationships with the three abundance measures (Figure 4.5). Hypothesis tests to detect presence of positive correlations in any of the 15 pairs of measures were conducted, using a Bonferroni correction to adjust for the multiple comparisons. The highest correlation among any pair of measures ($r = 0.97$) was between the eDNA 'between' timepoint and ROV measures (Figure 4.5). The hypothesis tests showed that this was the only statistic large enough to confirm a positive correlation at significance level $a = .05$ (since its p-value = $0.0024 < 0.05/15 = 0.0033$), likely as a result of the lack of power from the small sample size.

Aims 4, 5, 6: Simplifying eDNA sampling for integration into routine stock assessment research

Negative controls for the Niskin water sampling with filtration were deemed free from contamination (i.e., did not meet any of the defined analysis criteria) during routine stock assessment surveys conducted in spring or fall 2023. Negative controls for the passive samplers were not developed/integrated into spring surveys, but were included and found to be free from contamination during fall surveys. These findings are promising, indicating that eDNA can be integrated relatively easily into existing established surveys without interference.

During comparisons of eDNA sampling methods, passive samplers (i.e., Speci-sponges) yielded more positive detections for *S. dumerili* and S. *rivoliana*. Niskin water sampling with filtration detected *S. dumerili* and S. *rivoliana* at only one of the seven sites ('War Tank', Site 3), a site with a higher abundance of *S. dumerili* (Figure 4.6). The passive samplers, however, detected both *S. dumerili* and S. *rivoliana* DNA at four of the seven sites, and one additional site where only *S. dumerili* DNA was detected. Using the passive samplers, *S. dumerili* DNA was detected at all of the sites with higher 'known' *S. dumerili* abundances and at two of the four sites with lower 'known' abundances of *S. dumerili*. This suggests that passive sampling approaches may improve on species eDNA detections, owing to the ability for the sampler to 'go to the fish' via deployment on the ROV, combined with longer deployment/soak times when compared to the collection (via Niskin sampler) and filtration of large volumes of water. This finding is preliminary due to small sample sizes, but is a promising outcome to support the future development of streamlined eDNA sampling methods to improve the integration of eDNA tools in fisheries surveys.

Discussion

Here, we developed and validated a novel tetraplex ddPCR™ eDNA assay that can detect and distinguish between *S. dumerili*, S. *fasciata*, S. *rivoliana* and S. *zonata*. To our knowledge, this is the first ddPCR™ assay able to detect four aquatic species simultaneously in a single sample, and only the second published application of ddPCR™ in marine fisheries research (the first being Maes et al. 2024). This novel assay was highly specific in silico against sequences for other organisms in GenBank, in vitro against an extensive list of co-occurring and related exclusion species, and in situ, supported by datasets from multiple fisheries survey gears. There was no evidence of contamination during laboratory and field processing suggesting that detections are valid. Throughout this study, we have demonstrated that with strict cleaning and handling protocols and thorough testing of negative controls, it is viable to conduct robust eDNA research as part of a fisheries survey cruise. The validation of the species-specific assay for *S. dumerili*, S. *rivoliana*, S. *zonata* and S. *fasciata* in the field allowed for the development of a robust and efficient methodology for field use, notably the ability to quickly filter and preserve samples, limiting potential contamination that can result from multiple handling during these processes.

This tetraplex eDNA assay was developed using mtDNA loci alone, therefore, it will not be able to identify potential hybrids. However, at the time of this study, no nuclear sequences of *Seriola* species were available on Genbank and, although hybrids are known to occur in this genus (Takahashi et al. 2021), there is not yet evidence of hybridization among the target species of this study in U.S. waters. The assay specificity was tested on regionally co-occurring species, with the aim of making it suitable for eDNA research in the U.S. Atlantic Ocean and Gulf of Mexico. Further testing of co-occurring and closely related exclusion species must be conducted prior to using the assay on eDNA samples from other regions. The fact that the assay showed non-specific detection for S. lalandi was not considered relevant for the current work as this species does not occur in the U.S. Atlantic Ocean and Gulf of Mexico (Swart et al. 2015) but could be an issue in areas where it does co-occur. Because the developed primers amplify DNA across all *Seriola* species, the assay could be adjusted via probe redevelopment and optimization to detect species within this genus in other areas of the world. Sensitivity (LoQ and LoD) must be reported for any new markers and/or areas for the sake of assay transparency and reliability, and to allow future users to make informed decisions on whether to adopt existing markers or design new ones (Xia et al. 2021).

Comparison of eDNA detections to other gears

Both eDNA and camera gears detected the presence of *S. dumerili* and S. *rivoliana*, but not S. *fasciata* and S. *zonata*. This result was expected, as *S. dumerili* and S. *rivoliana* are the most common species in the surveyed region of Alabama, where S. *fasciata* (mostly present in the western Gulf of Mexico) and S. *zonata* (mostly present along the Atlantic) are rarely seen.

Seriola dumerili eDNA detection data aligned with detections from other gears, with detections on all sampling days in eDNA surveys, the acoustic array, ROV and chevron trap cameras. *Seriola dumerili* were also detected on S-BRUV videos on four of the five sampling days, but this could be explained by camera malfunctions that resulted in a lower sample size on one day. Further, *S. dumerili* were detected at almost all eDNA sampling timepoints except one, which could be explained by the timing and distance of deployment. Specifically, this sample was collected before any baited gear was deployed and at the furthest distance (72 m) recorded from the reef. After an overnight storm, increased currents were experienced on this sampling day, resulting in a higher gear deployment distance from the reef. This

could account for the overall lower detections for *S. dumerili*, as some of the lowest detections were observed on this sampling day across gears.

EDNA detection data for S. *rivoliana* also aligned well with overall detections from other gears. *Seriola rivoliana* were detected in eDNA surveys, ROV, chevron trap and S-BRUV cameras on all sampling days. Further, S. *rivoliana* eDNA was detected at 12 out of the 15 sampling timepoints. Six of these positive detections were 'one droplet' detections, demonstrating the importance of LoD and validating assay sensitivity in the field. *Seriola rivoliana* were detected after all baited gear deployments, but not between deployments on two of the five sampling days, nor before gear deployment on one day. Baited gear deployment is likely to have influence on eDNA detection, given its ability to attract *Seriola* species.

Comparison of eDNA concentrations to other gears

In this study, eDNA relative concentration data was higher for *S. dumerili* than S. *rivoliana* at both sites for all sampling days, in line with relative abundance patterns obtained from camera gears. Both *S. dumerili* and S. *rivoliana* eDNA relative concentrations were higher at site 2 compared to site 1. For *S. dumerili*, higher abundance was also observed on acoustic and camera gears at site 2 compared to site 1. For S. *rivoliana*, higher abundance at site 2 compared to site 1 was not observed across all camera gears, possibly due to the high variance inherent to fisheries gears.

The LoQ establishes the level above which eDNA concentrations can be reliably used for quantification comparisons to other gears and potential abundance estimates (Dobnik et al. 2016, Klymus et al. 2019). *Seriola dumerili* eDNA concentrations were above the LoQ at 10/15 sampling timepoints, across all sampling days. These included all timepoints after multiple baited camera gears were deployed, three out of five timepoints between gear deployments and two out of five timepoints before any gear was deployed. Higher *S. dumerili* eDNA concentrations were observed on day one and day five, when baited camera gears were deployed over two times each. These results suggest the deployment of bait and/or gears, and its ability to attract *Seriola* species, has an influence on eDNA concentration.

Seriola rivoliana eDNA relative concentrations were above the LoQ at only one sampling timepoint, before gear deployment on day two. Higher eDNA concentration data before gear deployment on day two for both S. *rivoliana* and *S. dumerili* can result from the species proximity to the Niskin sampler. This is noted also on day four after all deployments and where, despite sampling at the furthest distance to the reef (92 m), one of the highest *S. dumerili* eDNA relative concentration was obtained.

For *S. dumerili*, eDNA concentrations above the LoQ allowed for a better understanding of the relation to baited gear deployment and sampling distance from the reef. There were, however, contradictions in the data that highlight a challenge of eDNA, where *S. dumerili* proximity to the sampling device needs to be disentangled from higher fish abundance. Only by increasing sampling coverage and replication (e.g., by using a rosette sampler, not available in this study) could target species eDNA be modeled over relevant gradients, disentangling proximity from abundance and reducing variance and uncertainty in eDNA data.

This study took place outside of the known *Seriola* sp. spawning season (March to June), thus being unlikely that this biological factor could contribute to any contradictions observed. The size range of *S. dumerili* captured during the tagging cruise and the size range of *S. dumerili* and S. *rivoliana* observed on the cameras indicated young adult fishes to be mostly present. The fact that eDNA sampling took place

immediately after the camera gears, limited the influence of eDNA decay with time and dispersal with currents, as much as possible. All other environmental parameters were fairly constant throughout the five days of sampling, and therefore unlikely to be responsible for any variation in eDNA data between sampling days.

Incorporating eDNA into fisheries surveys

The non-invasive, relatively simple, fast, and overall cost-effective methodology represents some of the most compelling advantages of incorporating eDNA sampling in fisheries surveys. Assay specificity has also often proven eDNA to be a more reliable method for species detection than traditional survey methods (Jerde 2019, Ramírez-Amaro et al. 2022, Yao et al. 2022). Historic misidentification of these four *Seriola* species has been an impediment to effective fisheries management (Berry and Burch 1979, Renshaw and Gold 2009, Galbraith et al. 2022). Although taxonomic status of *S. dumerili*, S. *fasciata*, S. *rivoliana*, and S. *zonata* has been reasonably established based on the morphological identification of adults, similarities in their shape, coloration, and ontogenetic changes make them difficult to distinguish at earlier life stages and/or via camera gears. Among all gears, only VPS and eDNA provided 100% species-level detections. Acoustic arrays are mostly temporary, require continued investment to maintain, and are only able to detect tagged individuals, with the number of species and sample size of tagged fish being limited by the cost of purchasing and affixing of acoustic transmitters. In contrast, ddPCR™ assay development represents a one-off initial investment for a given biogeographic area of sampling. In a ddPCR™ assay, individuals of at least four species can be detected with little additional cost, and other species can be included with subsequent one-off partial investments. EDNA can therefore be a valuable complement to camera gears (such as chevron trap and S-BRUV cameras) deployed in geographic areas and/or seasons with poor visibility conditions, where fish may be present but cannot be confidently identified to the species level.

Mobile camera gears such as ROV are less constrained by deployment distance and visibility conditions than stationary gears, as they are controlled by an operator at all times and able to track habitat and fish. They are the gear of choice in fisheries surveys in Alabama, where they are known to provide reliable data for the type of reefs sampled in this study. The high correlation obtained between ROV and eDNA data (collected between gear deployments) adds to a growing number of studies demonstrating the potential of eDNA data as a cost-effective indicator of abundance for fisheries stock assessments (see review by Rourke et al. 2022). EDNA data collected between gears are likely to represent a more realistic comparison to abundance data by any gear, while eDNA data collected after multiple baited gears deployment are likely to be exacerbated by recurrent attraction and activity of fish. The fact that baited gears appear to attract fish that is detectable via eDNA sampling, does however, suggest that it might be interesting to explore eDNA as a complementary method in surveys where baited gears are commonly used.

Investment into follow-up studies with increased sample size and spatial and temporal replication would be necessary to validate the robustness of the correlation observed. A robust correlation fitted into a modeling approach, such as the one developed by Chambert et al. (2018) aimed at inferring animal density from eDNA and animal count data from a subset of sites, could be more confidently incorporated into existing fisheries management frameworks for Alabama. Time-series data from a carefully designed eDNA survey that estimates relative abundance could be used as an index of relative abundance for the *S. dumerili* stock assessment. EDNA holds potential for improving fisheries surveys, as

a reliable and cost-effective support ‘gear’. Incorporating this ‘gear’ with other common fisheries survey techniques would not require additional days at sea and could provide improved precision for estimated fish detection and abundance. Additionally, all archived eDNA samples can be used to investigate other species in future analyses, either by targeted assays or metabarcoding of fish assemblages (Jerde 2019, Ramírez-Amaro et al. 2022, Yao et al. 2022).

Tables and Figures

Table 4.1. Target (shaded) and exclusion species used to develop species-specific environmental DNA assays for Greater Amberjack, *Seriola dumerili*, Almaco Jack, *Seriola rivoliana*, Lesser Amberjack *Seriola fasciata*, and Banded Rudderfish, *Seriola zonata*

Type	Family	Species	N	Location	Ocean Region
Target Species	Carangidae	Greater Amberjack *Seriola dumerili*	5	Gulf of Mexico, Alabama, USA	WNAO
			1	Cape Lambert, Western Australia, AUS*	EIO
			1	Madeira, Portugal *	ENAO
			1	Key Largo, Florida, USA*	WNAO
		Almaco Jack *Seriola rivoliana*	3	Gulf of Mexico, Mississippi, USA	WNAO
			1	John's Island, South Carolina, USA*	WNAO
			2	South Carolina, USA*	WNAO
		Lesser Amberjack *Seriola fasciata*	2	Gulf of Mexico, Louisiana, USA	WNAO
			1	John's Island, South Carolina, USA*	WNAO
			1	Cape Hatteras, North Carolina, USA*	WNAO
		Banded Rudderfish *Seriola zonata*	2	North Carolina, USA	WNAO
			1	Panama City, Florida, USA*	WNAO
			1	South Carolina, USA*	WNAO
Exclusion Species		Guinean Amberjack *Seriola* carpenter	1	Angola*	ESAO
		Samson Fish *Seriola* hippos	1	Eyre Peninsula, South Australia, AUS*	SO
			1	New South Wales, AUS *	WSPO
			1	Rottnest Island, Western Australia, AUS*	EIO
		Yellowtail Amberjack *Seriola* lalandi	1	Southwest Rift, South Africa*	WIO
			2	New South Wales, Australia*	WSPO
		Japanese Amberjack *Seriola* quinqueradiata	3	Kochi Prefecture, Japan*	WNPO
		Crevalle Jack Caranx hippos	1	Gulf of Mexico, USA	WNAO
		Horse-eyed Jack Caranx latus	1		WNAO
		Florida Pompano Trachinatus carolinus	1		WNAO
		Permit Trachinatus falcatus	1		WNAO
		Rough Scad Trachurus lathami	1		WNAO
		Round Scad Decapterus punctatus	1		WNAO
		Rainbow Runner Elagatis bipinnulata	1		WNAO
		Atlantic Moonfish Selene setapinnis	1		WNAO

	Other	Great barracuda Sphyraena barracuda	1		WNAO
		Cobia Rachycentron canadum	1		WNAO
		King Mackerel Scomberomorus cavalla	1		WNAO
		Swordfish Xiphias gladius	1		WNAO
		Gag Grouper Mycteroperca microlepis	1		WNAO
		Red Snapper Lutjanus campechanus	1		WNAO
		Tilefish Lopholatilus chamaeleonticeps	1		WNAO
		Tripletail Lobotes surinamensis	1		WNAO
		Red Drum Sciaenops ocellatus	1		WNAO
		Gulf Menhaden Brevoortia patronus	1	Bait on vessel	NA
		Atlantic Mackerel Scomber scombrus	1		NA
		Squid	1		NA
		Canned tuna	1	Sandwich on board vessel	NA

*From Swart et al. 2015; AUS=Australia, WNAO=Western North Atlantic Ocean, EIO=Eastern Indian Ocean, ENAO= Eastern North Atlantic Ocean, ESAO= Eastern South Atlantic Ocean, SO=Southern Ocean, WSPO=Western South Pacific Ocean, WIO=Western Indian Ocean, WNPO=Western North Pacific Ocean

Table 4.2. Primers and species-specific probes used to amplify a 124-base pair fragment of the mitochondrial NADH dehydrogenase 5 and 6 gene in target *Seriola* spp

Species	Oligo Type	Oligo sequence	Fluorophore
Seriola spp.	Forward primer	5’-CAACRTYCAACGRGGTATR-3’	NA
	Reverse primer	5’-CGTGGGTTCTTCTYTTGAC-3	NA
Seriola dumerili	Probe	5’-CAAAACCTACCTCTCTCTCTTCCTACTTACC-3’	6-FAM
Seriola rivoliana	Probe	5’-TCTCTCTTCCTGCTCACCCTTG-3’	SUN
Seriola fasciata	Probe	5’-CCTCGTCCTCATGGTCTTATTAGTTACC-3’	6-FAM
Seriola zonata	Probe	5’-ATTCTATTAGTCACCCACTAGACAGCTC-3’	SUN

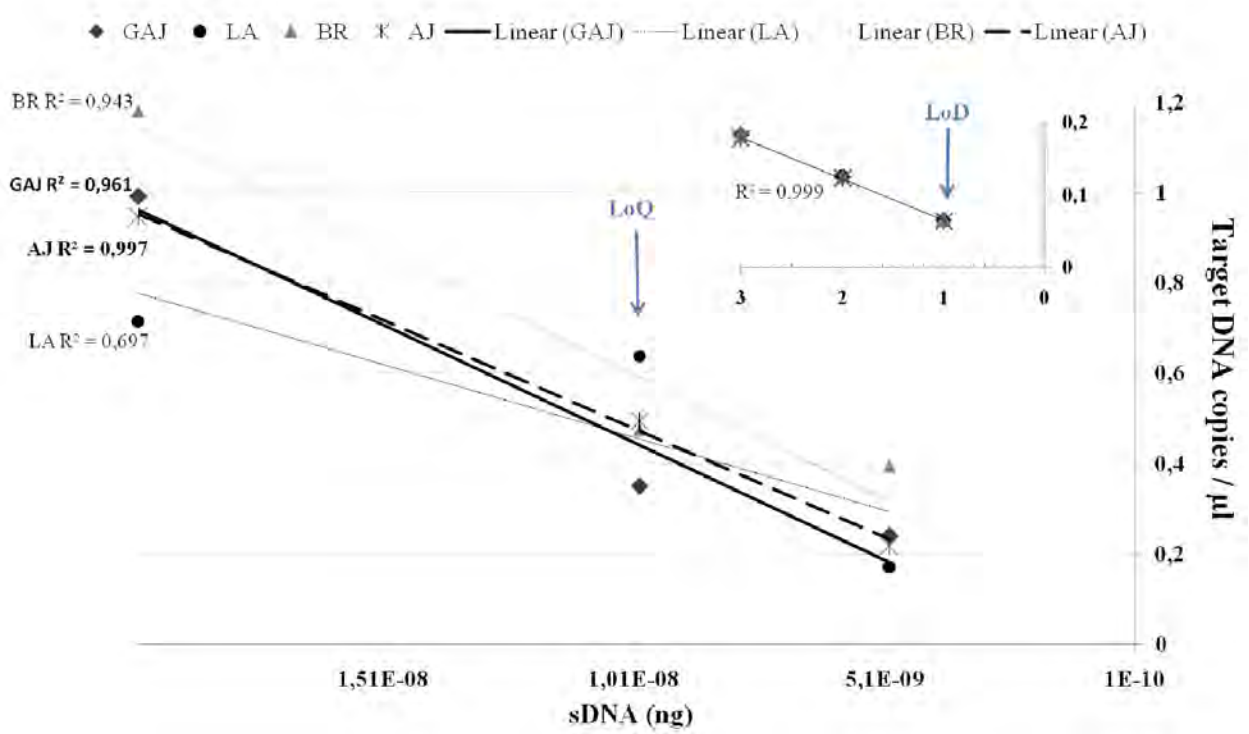
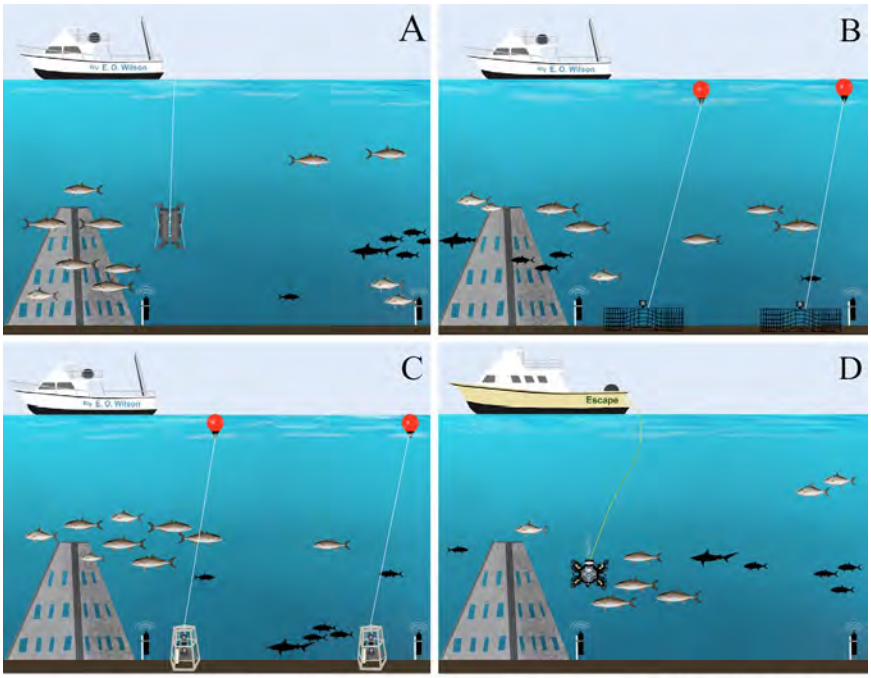


Figure 4.1. Limit of Detection (LoD) and Limit of Quantification (LoQ) of Droplet Digital PCR™ assays for *Seriola* spp. Average target DNA concentrations (copies/µL) and ddPCR efficiency of a dilution series from a concentration of 2 × 10−8 ng of synthetic DNA (sDNA) to one droplet. Data is presented for Greater Amberjack *Seriola dumerili* (GAJ), Lesser Amberjack *S. fasciata* (LA), Almaco Jack *S. rivoliana* (AJ) and Banded Rudderfish *S. zonata* (BR). Smaller inset graph in the upper right shows ddPCR efficiency on one, two and three positive droplets obtained below the LoQ at 0.25 × 10−8, 2 × 10−9 and 1 × 10−9 ng of sDNA.



Fieldwork Schedule		Day 1						Day 2						Day 3						Day 4						Day 5					
		Site 2						Site 1						Site 2						Site1						Site 2					
Gear cast	type	A	B	A	C	A	D	A	C	A	B	A	D	D	A	B	A	C	A	A	C	A	B	A	D	A	C	A	B	A	D
	number	1	2.5	1	2	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	2	1	2	1	1	1	3	1	2	1	1

Figure 4.2. Sampling scheme for environmental DNA validation and groundtruthing surveys in Alabama. Gear includes eDNA (A), chevron traps (B) and S-BRUVs (C) on one vessel, and ROV (D) on a second vessel. The different gear types were deployed over five days on two sites following the order and number of casts in the schedule below the figures. Illustration credit: Bryan Huerta-Beltrán.

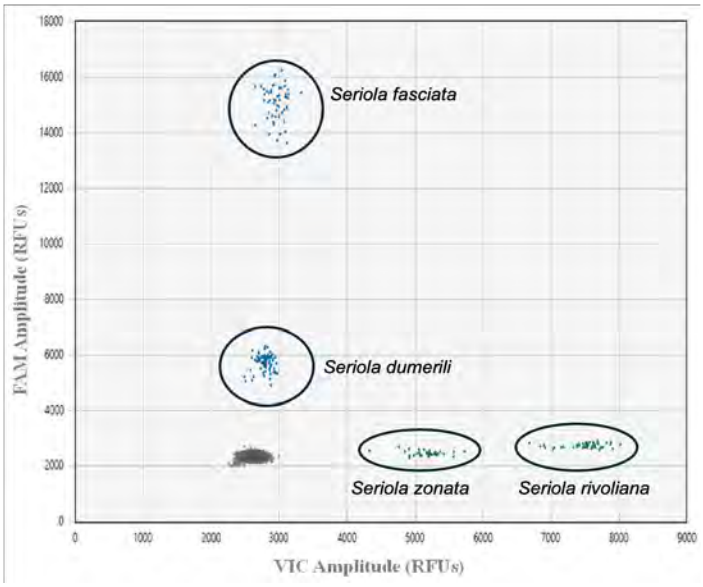


Figure 4.3. Droplet Digital PCR™ positive droplet clusters for *Seriola* spp. using 0.2 ng genomic DNA. Each droplet in the reaction was classified as negative (grey, bottom left) or positive (blue on FAM y-axis or green on VIC x-axis) for target DNA, detected using a Bio-Rad® QX200™ Droplet Reader, QuantaSoft™ software and Absolute Quantification analysis setting.

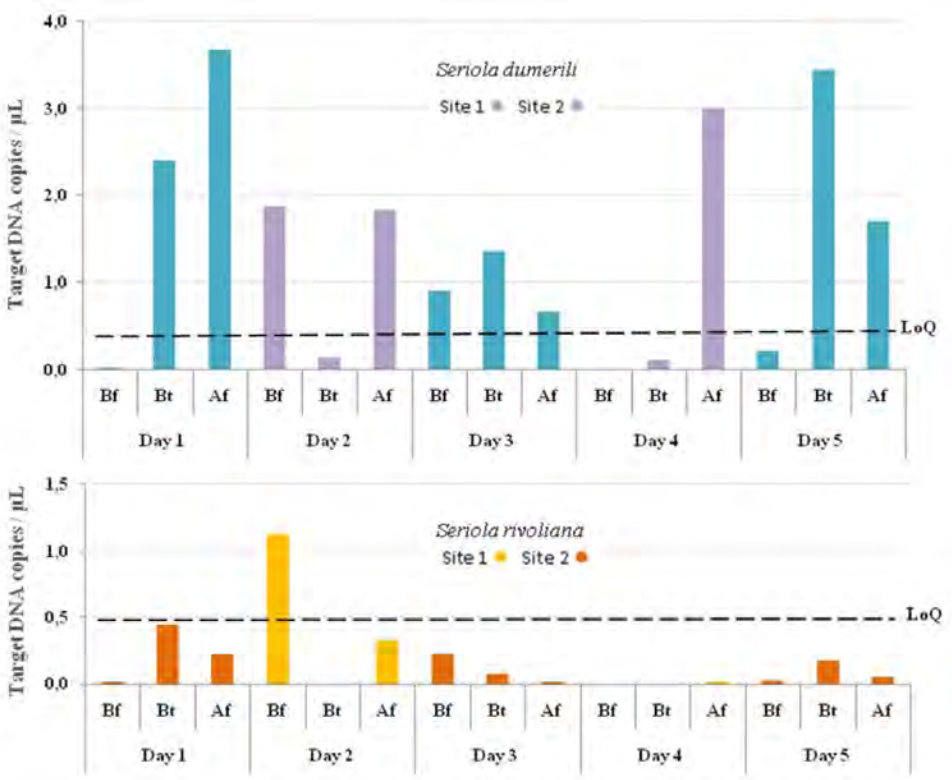


Figure 4.4. Droplet Digital PCR™ detections of Greater Amberjack, *Seriola dumerili*, and Almaco Jack, *S. rivoliana* at two sites in Alabama over five sampling days. EDNA samples were collected before (Bf), between (Bt) and after (Af) the deployment of chevron trap and S-BRUV gears (see Figure 4.2). Detections are expressed as target DNA copies/µL, and the Limit of Quantification (LoQ) are indicated by a dashed black line for each species.

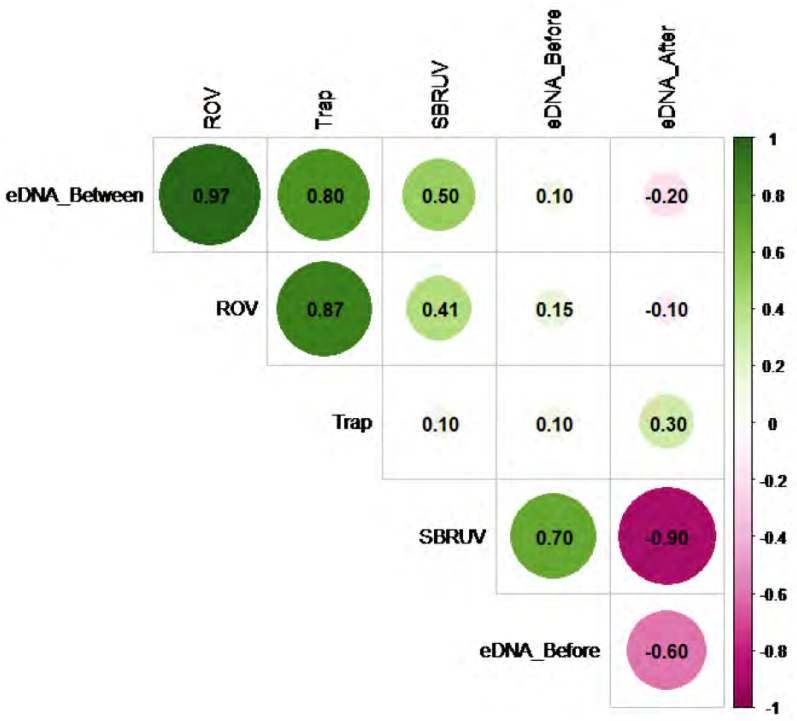


Figure 4.5. Spearman correlation coefficient (r) among concentrations of Greater Amberjack, *Seriola dumerili*, environmental DNA detected using Droplet Digital PCR™ and camera gears across five sampling days. eDNA samples were collected before, between, and after deployment of other gears. The magnitude of the correlations is depicted by color (pink=negative, green=positive) and shade. The size of the circle corresponds to the magnitude of absolute value.

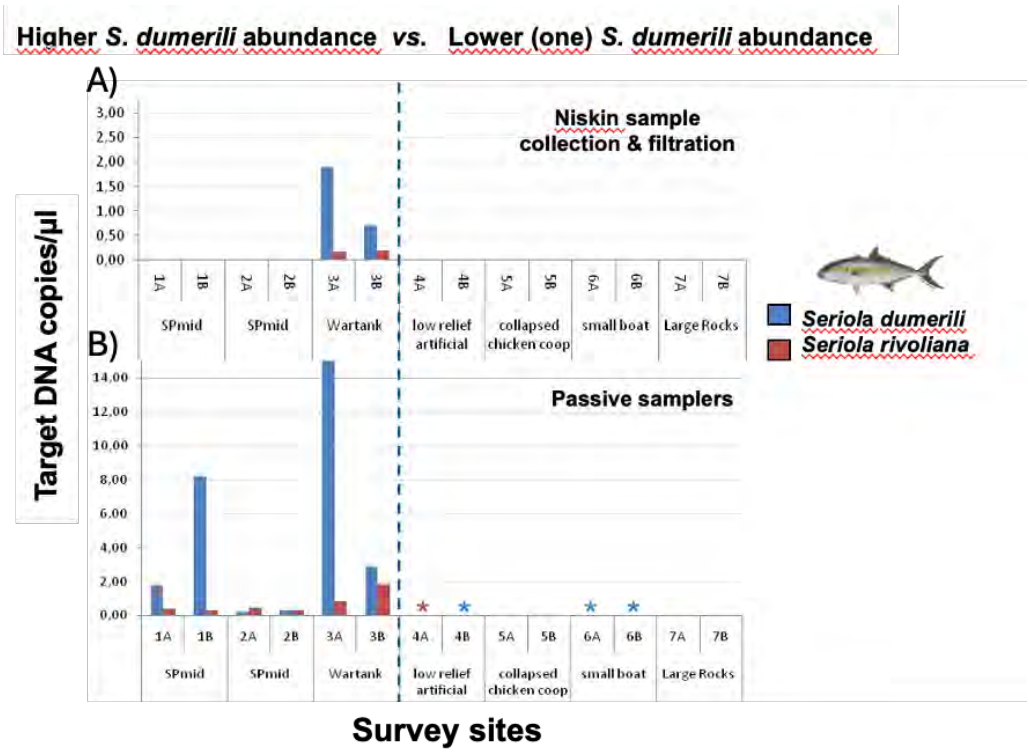


Figure 4.6. Droplet Digital PCR™ detections of *Seriola dumerili* and *S. rivoliana* at sites in Alabama with higher (Site no. 1-3) or lower (Site no. 4-7) abundance survey sites based on A) collection of 2 X 10-L water samples with filtration, and B) eDNA passive sampling sponges. Detections are expressed as target DNA copies/µL. Stars indicate valid species detections with low concentrations of target DNA not visible on the bar chart.

References

Alzohairy A. 2011. BioEdit: An important software for molecular biology. GERF Bulletin of Biosciences 2: 60-61.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Science of Molecular Biology 215: 403–410.

Bacheler N, Gillum Z, Gregalis K, Schobernd C, Schobernd Z, Teer B. 2020. Spatial Patterns in Relative Abundance and Habitat Use of Adult Gray Snapper off the Southeastern Coast of the United States. Marine and Coastal Fisheries12: 205-219. https://doi.org/10.1002/mcf2.10118.

Berry FH, Burch RK. 1979. Aspects of amberjack fisheries. Proceedings of the Gulf and Caribbean Fisheries Institute31:179–194.

Bockrath K, Maloy A, Rees C, Tuttle-Lau M, Mize E. 2022. Environmental DNA (eDNA) Best Management Practices for Project Planning, Deployment, and Application. https://doi.org/10.13140/RG.2.2.12200.96006.

Bustin S, Benes V, Garson J, Hellemans J, Huggett J, Mueller R, Nolan T, Shipley G, Vandesompele J, Wittwer C. 2009. The MIQE Guidelines: minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical chemistry 55: 611-22. https://doi.org/10.1373/clinchem.2008.112797.

Chambert T, Pilliod DS, Goldberg CS, Doi H, Takahara T. An analytical framework for estimating aquatic species density from environmental DNA. 2018. Ecology and Evolution 8: 3468–3477. https://doi.org/10.1002/ece3.3764.

Dimond J, Gathright B, Bouma J, Carson H, Sowul K. 2022. Detecting endangered pinto abalone (Haliotis kamtschatkana) using environmental DNA: Comparison of ddPCR, qPCR, and conventional diver surveys. Environmental DNA 4. https://doi.org/10.1002/edn3.351.

Dobnik D, Stebih D, Blejec A, Morisset D, Zel J. 2016. Multiplex quantification of four DNA targets in one reaction with Bio-Rad droplet digital PCR system for GMO detection. Scientific Reports 6: 35451. https://doi.org/10.1038/srep35451.

Doi H, Uchii K, Takahara T, Matsuhashi S,Yamanaka H, Minamoto T. 2015. Use of Droplet Digital PCR for Estimation of Fish Abundance and Biomass in Environmental DNA Surveys. PLoS ONE 10(3): e0122763. https://doi.org/10.1371/journal.pone.0122763

Drymon JM, Schweiss KE, Seubert EA, Lehman RN, Daly-Engel TS, Pfleger M, Phillips NM. 2021. Swimming against the flow—Environmental DNA can detect bull sharks (Carcharhinus leucas) across a dynamic deltaic interface. Frontiers in Marine Science 11(1): 22–28. https://doi.org/10.1002/ece3.7101

Galbraith JK, Bemis KE, Bemis WE, Cook HS, Wuenschel MJ. 2022. Identifications, distributions, and life history of four species of *Seriola* (Carangiformes: Carangidae) in the western North Atlantic based on contemporary and historical data. NOAA Professional Paper NMFS 22. https://doi.org/10.7755/PP.22

Goldberg CS, Pilliod DS, Arkle RS, Waits LP. 2011. Molecular detection of vertebrates in stream water: A demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. PLoS ONE 6(7): e22746. https://doi.org/10.1371/journal.pone.0022746

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution 7: 1299–1307. https://doi.org/10.1111/2041-210X.12595

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.

Hansen B, Bekkevold D, Clausen L, Nielsen E. 2017. The sceptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. Fish and Fisheries 19. https://doi.org/10.1111/faf.12286.

Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, Hellemans J, Mueller RD, Nolan T, Pfaffl MW, Shipley G, Vandesompele J, Wittwer CT, Bustin SA. 2013. The Digital MIQE Guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. Clinical Chemistry 59. https://doi.org/10.1373/clinchem.2013.206375.

Huggett JF, dMIQE Group. 2020. The Digital MIQE Guidelines Update: Minimum Information for Publication of Quantitative Digital PCR Experiments for 2020.

Hunter M, Meigs-Friend G, Ferrante JA, TakoukamKamla A, Dorazio R, Keith-Diagne L, Luna F, Lanyon J, Reid JP. 2018. Surveys of environmental DNA (eDNA): A new approach to estimate occurrence in Vulnerable manatee populations. Endangered Species Research 35. https://doi.org/10.3354/esr00880

Jerde CL. 2019. Can we manage fisheries with the inherent uncertainty from eDNA? Journal of Fish Biology 1–13. https://doi.org/10.1111/jfb.14218

Keenan SF, Switzer TS, Thompson KA, Tyler-Jedlund AJ. 2018. Comparison of Reef-Fish Assemblages between Artificial and Geologic Habitats in the Northeastern Gulf of Mexico: Implications for Fishery-Independent Surveys. American Fisheries Society Symposium 86:141–163.

Kelly RP, Lodge DM, Lee KN, Theroux S, Sepulveda AJ, Scholin CA, Craine JM, Andruszkiewicz AE, Nichols KM, Parsons KM, Goodwin KD, Gold Z, Chavez FP, Noble RT, Abbott CL, Baerwald MR, Naaum AM, Thielen PM, Simons AL, Weisberg, SB. 2023. Toward a national eDNA strategy for the United States. Environmental DNA 6(1): 1–10. https://doi.org/10.1002/edn3.432

King A, Krieg R, Weston A, Zenker A. 2022. Using eDNA to simultaneously detect the distribution of native and invasive crayfish within an entire country. Journal of Environmental Management 302. https://doi.org/113929. 10.1016/j.jenvman.2021.113929.

Kirtane A, Wieczorek D, Noji T, Baskin L, Ober C, Plosica R, Chenoweth A, Lynch K, Sassoubre L. 2021. Quantification of Environmental DNA (eDNA) shedding and decay rates for three commercially harvested fish species and comparison between eDNA detection and trawl catches. Environmental DNA 3: 1142–1155. https://doi.org/10.1002/edn3.236.

Klymus KE, Merkes CM, Allison MJ, Goldberg CS, Helbing CC, Hunter ME, Jackson CA, Lance RF, Mangan AM, Monroe EM, Piaggio AJ. 2019. Reporting the limits of detection and quantification for environmental DNA assays. Environmental DNA 2: 3–12. https://doi.org/10.1002/edn3.29.

Klymus KE, Ramos DVR, Thompson NL, Richter CA. 2020. Development and testing of species-specific quantitative PCR assays for environmental DNA applications. Journal of Visualized Experiments 165: e61825.

Kokkoris V, Vukicevich E, Richards A, Thomsen C, Hart MM. 2021. Challenges Using Droplet Digital PCR for Environmental Samples. Applied Microbiology 1: 74–88. https://doi.org/10.3390/applmicrobiol 1010007.

Lamb PD, Fonseca VG, Maxwell DL, Nnanatu CC. 2022. Systematic review and meta-analysis: Water type and temperature affect environmental DNA decay. Molecular Ecology Resources 22(7): 2494-2505. https://doi.org/10.1111/1755-0998.13627.

Langlois VS, Allison MJ, Bergman LC, To TA, Helbing CC. 2020. The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. Environmental DNA: edn3.164. https://doi.org/10.1002/edn3.164.

Ledger K, Hicks M, Hurst T, Larson W and Baetscher D. 2024, Validation of Environmental DNA for Estimating Proportional and Absolute Biomass. Environmental DN:, 6: e70030. https://doi.org/10.1002/edn3.70030

Lehman RN, Poulakis GR, Scharer RM, Schweiss KE, Hendon JM, Phillips NM (2020). An environmental DNA tool for monitoring the status of the Critically Endangered smalltooth sawfish, Pristis pectinata, in the western Atlantic. Conservation Genetics Resources, 12(4), 621–629. https://doi.org/10.1007/s12686-020-01149-5.

Lehman RN, Poulakis GR, Scharer RM, Hendon JM, Court AG, Wooley AK, Williams AM, Ajemian MJ, Hadden JP, Beal JL, McCallister MP, Phillips NM. 2022. Environmental DNA evidence of the Critically Endangered smalltooth sawfish, Pristis pectinata, in historically occupied US waters. Aquatic Conservation: Marine and Freshwater Ecosystems 32(1): 42-54.

Maes S M, Desmet S, Brys R, Sys K, Ruttink T, Maes S, Hostens K, Vansteenbrugge L, Derycke S .2024. Detection and quantification of two commercial flatfishes (Solea solea and Pleuronectes platessa) in the North Sea using environmental DNA. Environmental DNA: 6, e426. https://doi.org/10.1002/edn3.426

Mauvisseau Q, Davy-Bowker J, Bulling M, Brys R, Sabrina N, Troth C, Sweet M. 2019. Combining ddPCR and environmental DNA to improve detection capabilities of a critically endangered freshwater invertebrate. ScientificReports 9. https://doi.org/10.1038/s41598-019-50571-9.

Nathan L, Simmons M, Wegleitner B, Jerde C, Mahon A. 2014. Quantifying Environmental DNA Signals for Aquatic Invasive Species Across Multiple Detection Platforms. Environmental Science and Technology 48(21): 12800-6. https://doi.org/48.10.1021/es5034052.

Ramirez-Amaro S, Bassitta M, Picornell A, Ramon C and Terrasa B. 2022. Environmental DNA: State-ofthe-art of its application for fisheries assessment in marine environments. Frontiers in Marine Science 9:1004674. https://doi.org/10.3389/fmars.2022.1004674.

Renshaw MA, Gold JR. 2009. Genetic Tools to Identify Species of *Seriola* in the U.S. South Atlantic and Gulf of Mexico. Gulf of Mexico Science 27 (2). https://aquila.usm.edu/goms/vol27/iss2/6.

Rourke ML, Fowler AM, Hughes JM, Broadhurst MK, DiBattista JD, Fielder S, Walburn JW, Furlan EM. 2022. Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. Environmental DNA 4: 9–33. https://doi.org/10.1002/edn3.185.

Saborido-Rey F, Fernández-Otero R, Casas L, Rodríguez-Mendoza R, Pérez M, Piferrer F, Anastasiadi D, Guijarro B, Ramírez-Amaro S, Terrasa B, Picornell A, Ramon C, Eschbach E, Kasmi Y, Hanel R, Stransky C. Improving the cost-efficiency of fisheries research surveys and fish stocks assessment using next-generation genetic sequencing methods – FishGenome. Final Report. Publications Office of the European Union, 2023. https://doi.org/10.2926/803359.

Schweiss K, Lehman R, Drymon M, Phillips NM. 2019. Development of highly sensitive environmental DNA methods for the detection of Bull Sharks, Carcharhinus leucas (Müller and Henle, 1839), using Droplet Digital™ PCR. Environmental DNA 2. 2-13. https://doi.org/10.1002/edn3.39.

Shelton AO, Ramón-Laca A, Wells A, Clemons J, Chu D, Feist BE, Kelly RP, Parker-Stetter SL, Thomas R, Nichols KM, Park L. 2022. Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. Proceedings of the Royal Society B 289: 20212613.

Swart B, von der Heyden S, Bester-van der Merwe AE, Roodt-Wilding R. 2015. Molecular systematics and biogeography of the circumglobally distributed genus *Seriola* (Pisces: Carangidae). Molecular Phylogenetics and Evolution 93. https://doi.org/10.1016/j.ympev.2015.08.002.

Syngai G, Barman P, Bharali R, Dey S. 2013. BLAST: An introductory tool for students to Bioinformatics Applications. Kenean Journal of Science 2. 67-76.

Takahashi H, Kurogoushi T, Shimoyama R, Yoshikawa H. 2021. First report of natural hybridization between two yellowtails, *Seriola quinqueradiata* and S. lalandi. Ichthyological Research 68: 1-6. https://doi.org/10.1007/s10228-020-00752-8.

Takeuchi A, Iijima T, Kakuzen W, Watanabe S, Yamada Y, Okamura A, Tsukamoto K .2019. Release of eDNA by different life history stages and during spawning activities of laboratory-reared Japanese eels for interpretation of oceanic survey data. Scientific reports: 9(1), 6074.

Thomas A, Nguyen P, Howard J, Goldberg C. 2019. A self-preserving, partially biodegradable eDNA filter. Methods in Ecology and Evolution 10(8): 1136-1141. https://doi.org/10.10.1111/2041-210X.13212

Thomson-Laing G, Parai R, Kelly L, Pochon X, Newnham R, Vandergoes M, Howarth J, Wood S. 2021. Development of droplet digital Polymerase Chain Reaction assays for the detection of long-finned (Anguilla dieffenbachii) and short-finned (Anguilla australis) eels in environmental samples. PeerJ 9. e12157. https://doi.org/10.7717/peerj.12157.

von Ammon U, Averink T, Kumanan K, Brosnahan CL, Pochon X, Hutson KS and Symonds JE. 2022. An Efficient Tetraplex Surveillance Tool for Salmonid Pathogens. Frontiers in Microbiology. 13: 885585. https://doi.org/10.3389/fmicb.2022.885585.

Wood SA, Pochon X, Laroche O, von Ammon U, Adamson J, Zaiko A. 2019. A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species-specific detection in environmental DNA. Molecular Ecology Resources 19: 1407–1419.

Xia Z, Zhan A, Johansson ML, DeRoy E, Haffner GD, MacIsaac HJ. 2021. Screening marker sensitivity: Optimizing eDNA-based rare species detection. Diversity and Distribution 27: 1981–1988.

Yao M, Zhang S, Lu Q, Chen X, Zhang S, Kong Y, Zhao J. 2022. Fishing for fish environmental DNA (eDNA): ecological applications, methodological considerations, surveying designs, and ways forward. Molecular Ecology 31. https://doi.org/10.1111/mec.16659.

Yates MC, Fraser DJ, Derry AM. 2019. Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. Environmental DNA;1:5–13. https://doi.org/10.1002/ed

5. Exploitation

Matt Catalano

Summary

Conventional and acoustic tag returns were used to estimate average annual exploitation, apparent natural mortality, and tag shedding rate using a multi-state mark-recapture model. Tagging took place from May 2022 to March 2024. Over 654 fish were tagged with conventional dart tags and 443 were tagged with both dart tag and internal acoustic transmitter tags across the western Gulf, eastern Gulf and South Atlantic regions (Table 5.1). Half of these were double dart-tagged.

We used a multi-state mark recapture model with both conventional tag returns and acoustic detections to estimate parameters of interest. The model was stratified by tagging region (ATL, EGOM, WGOM) and included nine different biological states for conventionally tagged fish and fourteen states for acoustically tagged fish. We estimated monthly encounter rates by sector (recreational and commercial) for conventional tag returns and monthly crew-specific acoustic detection probabilities for acoustic tag returns. The model assumed post release mortality rates of 10% for the recreational sector and 20% for the commercial sector (from SEDAR).

Results

Returns of conventional tags came from the same region in which the fish were tagged, except for 1 fish that was tagged in Alabama and recaptured in eastern Louisiana (Table 5.1). Movement distances between tagging and recapture of dart-tagged fish were generally low in the Gulf, but several long-distance mark-recapture pairs occurred in the South Atlantic, including one fish that was tagged in North Carolina on Aug. 8, 2022 and recaptured in Placencia, Belize on Jun. 7, 2023 (Figure 5.1)

Model estimates:

1. A sigmoidal length-based retention probability consistent with length limits (Figure 5.2)
2. Relatively flat and uncertain length-based vulnerability to capture (Figure 5.3)
3. Highly variable, crew-specific acoustic detection probabilities (Figure 5.4)
4. Average annual exploitation rates (South Atlantic: 0.09, eastern Gulf of Mexico: 0.06, and western Gulf of Mexico: 0.06) lower than SEDAR estimates. Note, these include harvest plus dead discards
5. Apparent natural mortality estimate is high (0.80 yr^{-1} [+/- 0.07]), likely due to emigration. Note that this number will potentially decline as more detections accumulate across the acoustic receiver arrays.
6. Tag shedding rates were modest with an initial rate of 0.06 (+/- 0.04), and a chronic monthly rate of 0.01 (+/- 0.01)
7. Model-based acoustic tagging mortality was high, but similar to external estimates at 0.23 (+/- 0.03). Note, the model treats all unknown-fate fish as mortalities.

Tables and Figures

Table 5.1. Regional total dart tagged and dart + acoustic tagged fish numbers and dart tag returns.

Region	**Number Dart-Tagged (only)**	**Number Dart + Acoustic Tagged**	**ATL Dart-tag Returns**	**EGOM Dart-tag Returns**	**WGOM Dart-tag Returns**
ATL	157	139	40		
EGOM	285	121		75	1
WGOM	212	183			61

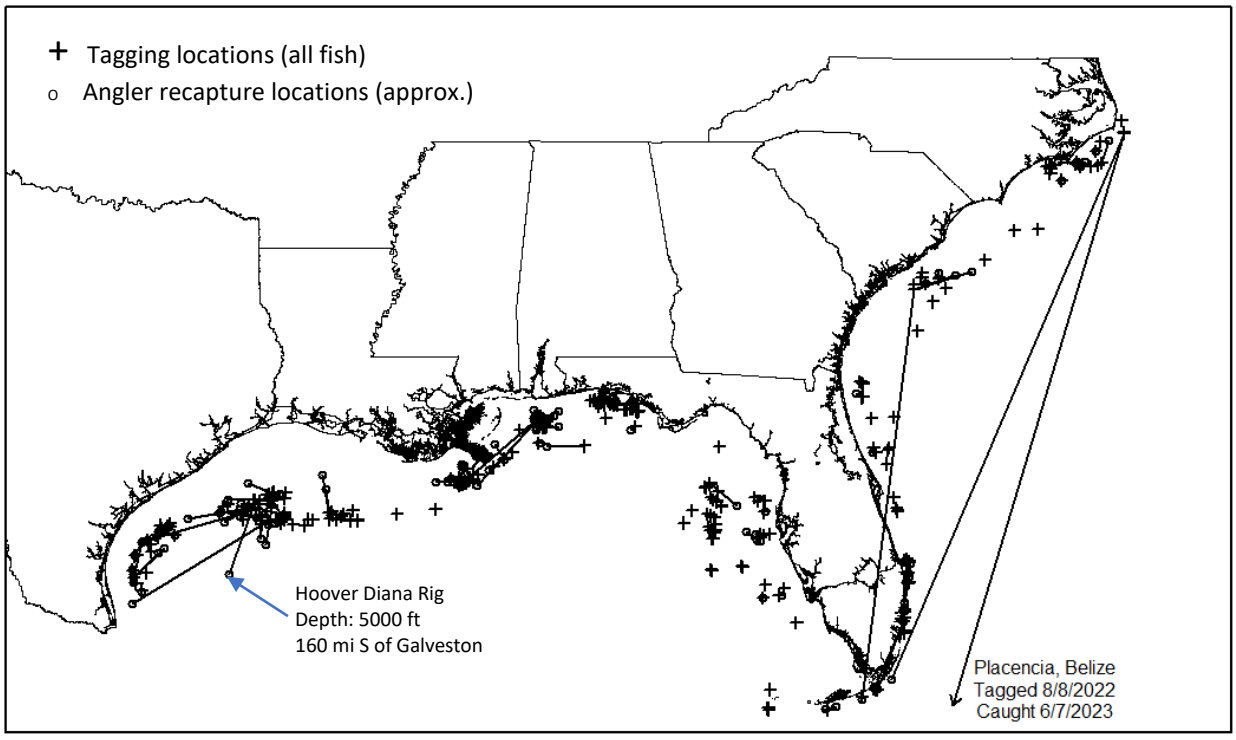


Figure 5.1. Tagging and angler recapture locations (approx.) for all fish.

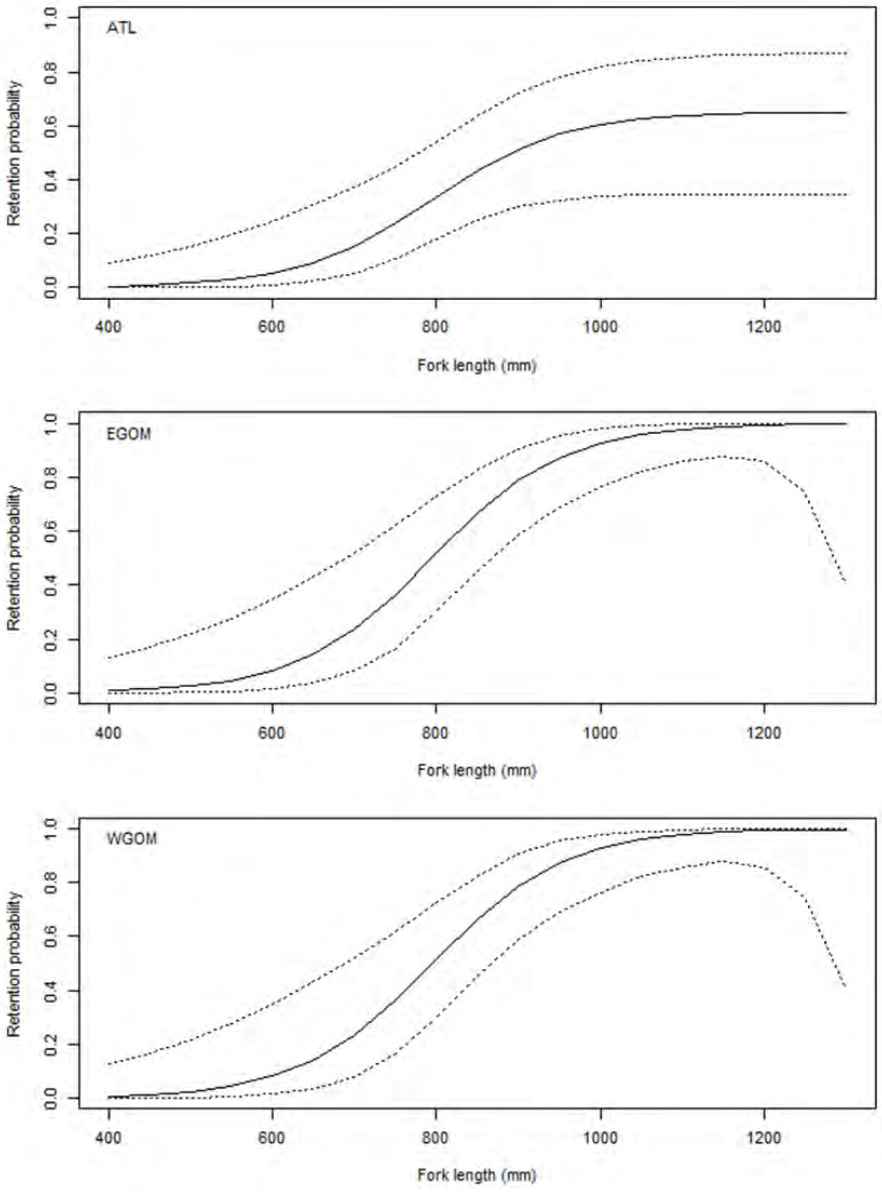


Figure 5.2: Model estimated length-based retention.

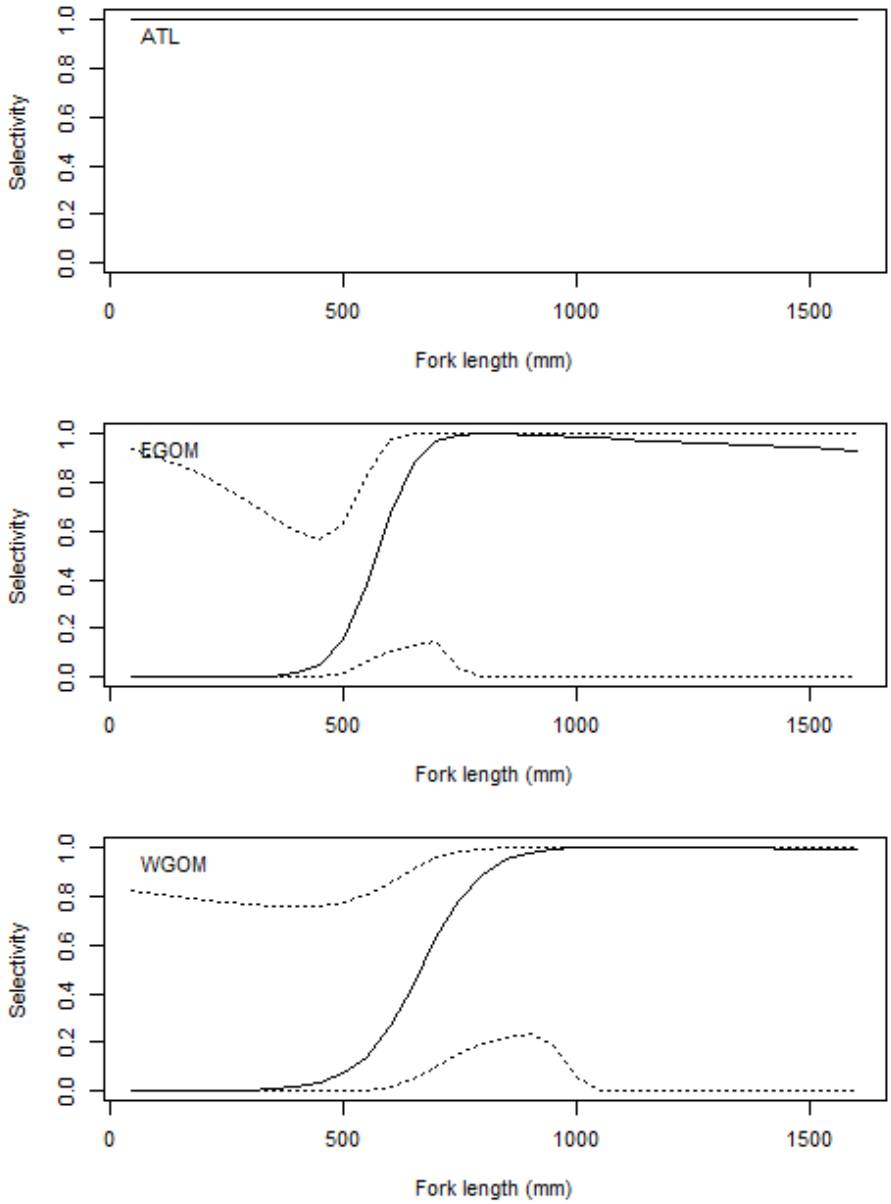


Figure 5.3. Estimated length-based selectivity

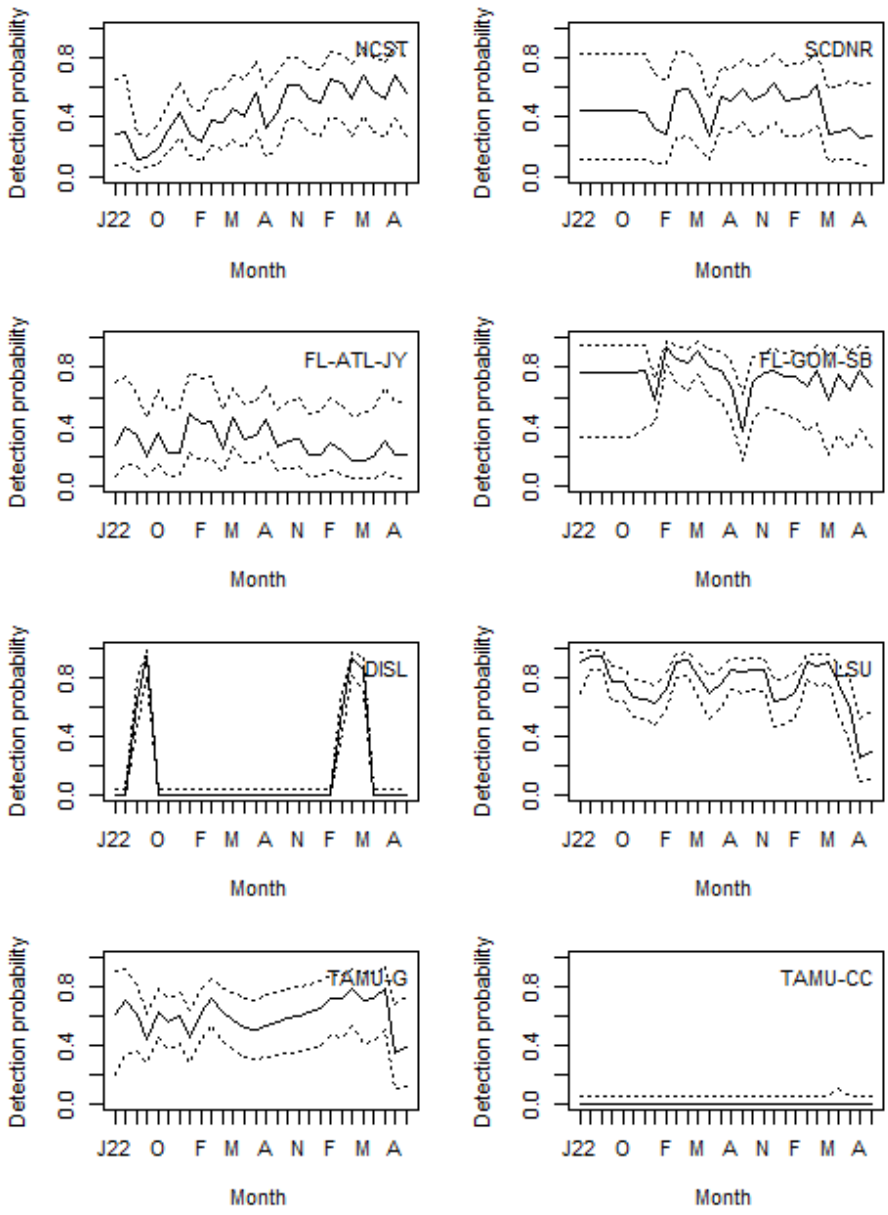


Figure 5.4. Crew specific acoustic detection probabilities

6. Movement ecology

Movement, habitat use, and connectivity of greater amberjack (*Seriola dumerili*) in the northwestern Atlantic Ocean

Creed C. Branham, Jeffery A. Buckel, Matthew J. Catalano, Mark A. Albins, Crystal L. Hightower, J. Marcus Drymon, Brett J. Falterman, Brondum M. Krebs, Daniel L. Lippi, Susan K. Lowerre-Barbieri, Stephen R. Midway, Landes L. Randall, Jay R. Rooker, Mischa K. Schultz, Matthew K. Streich, Ryan Tharpe, Aaron M. Watson, R. J. David Wells, Justin Yost, Joy M. Young, Sean P. Powers, Michael A. Dance

Introduction

Organismal movement and connectivity are critical components of marine population ecology, influencing gene flow, resilience, and population boundaries (Kerr et al., 2010; Frisk et al., 2014; Perkinson et al., 2019; Lédée et al., 2021). Movements of marine organisms occur at multiple spatial and temporal scales (e.g. local movements to cross-basin migrations) and life stages (Dance and Rooker 2015); parsing these movements is imperative to developing effective management strategies and understanding the impact of changing ecosystems (Allen & Singh, 2016). Localized movements reveal patterns of habitat use, residency, species-interaction, and response to changes in environmental conditions (Alos et al., 2012; Rooker et al. 2018, Bacheler et al., 2021), while large-scale movements contribute to connectivity between populations (Frisk et al., 2014; Rooker et al., 2016). Collectively, both broad movement behaviors and fine-scale habitat use drive population ecology, especially when movement behaviors vary within a population (Allen & Singh, 2016). Unfortunately, our understanding of movement at relevant spatial and temporal scales is lacking for most marine taxa, including species of significant economic and ecological interest, limiting our ability to effectively conserve and manage marine organisms and the ecosystems they live in (Ogburn et al., 2017; Matley et al., 2022).

The advent and recent growth of electronic tagging technology has dramatically increased our ability to monitor animal movement in marine ecosystems (Hussey et al. 2015). Passive acoustic telemetry has grown exponentially over the past two decades (Matley et al., 2022) and can be used to assess movement behaviors at a variety of spatial scales (e.g. Piraino & Szedlmeyer, 2014; Gandra, et al., 2020, Lowerre-Barbierri et al., 2021). Strategic acoustic receiver array design coupled with collaborative data-sharing networks (e.g. Ocean Tracking Network, OTN) allows for assessment of both habitat scale residency/movements and long-distance migratory behaviors that may cross population or geographic boundaries (Lowerre-Barbieri et al., 2025). The combination of passive hydrophones, data-sharing networks, and long battery life (6- 10 y) for acoustic transmitters relieves previous limitations in temporal duration, allowing for assessment of interannual variability in movement phenology (Espinoza et al., 2021; Lowerre-Barbieri et al., 2021; Jacoby & Piper, 2023). Long-term monitoring is especially applicable for species that may demonstrate divergent movement strategies within a population, and has been used widely to assess movement behavior, describe population structure, and identify areas of importance across multiple taxa (Lédée et al., 2021; Lowerre-Barbieri et al., 2021; Lowerre-Barbieri et al., 2025). Nevertheless, large-scale and long-term monitoring with acoustic telemetry is relatively new, and the majority of research has focused on charismatic species (e.g., large sharks), while many fishery species still lack movement data at scales needed for effective management.

In this study, we use acoustic telemetry to evaluate habitat use and connectivity of greater amberjack, *Seriola dumerili*, in U.S. waters of the northwestern Atlantic. The greater amberjack is a large, reef-associated fish that is globally distributed in sub-tropical and tropical waters of the world's oceans (Schultz et al., 2025). The species is of considerable economic and ecological concern, supporting valuable commercial and recreational fisheries throughout its range. Two populations are recognized in the U.S. Atlantic, one in the Gulf of America (also known as Gulf of Mexico; hereafter Gulf) and one along the southeastern U.S. Atlantic coast. Both populations experienced historical overfishing in U.S. waters, and while the Atlantic population is now recovered, the Gulf population remains overfished and experiencing overfishing (SEDAR 70). Despite the economic importance of the species, very little is known about greater amberjack movement, habitat use, and population connectivity throughout their range (McClellan & Cummings, 1997; Murie et al., 2013; Tone et al., 2021). The goal of this study was to characterize long-term movement patterns of greater amberjack in the U.S. Gulf and Atlantic to accomplish three primary objectives: 1) Evaluate habitat use and residency, 2) Identify and contrast patterns of dispersal and seasonal movement among three regions (Western Gulf, Eastern Gulf, and Atlantic), and 3) Assess connectivity and the potential for exchange and mixing among the three regions. In addition, the study aims to provide insight into the movement ecology of greater amberjack needed by fisheries managers to assess the potential influence of movement on fisheries-independent monitoring efforts and to better define population structure.

Methods

Study Area and Telemetry Arrays

The study area included shelf waters of the southeastern United States, from the United States-Mexico border to the North Carolina-Virginia border (~36.5 N, 74.6 W), and was divided into 3 regions : the Western Gulf, Eastern Gulf, and Atlantic. Acoustic receivers (VR2Tx and VR2AR, 69 kHz, Innovasea Systems Inc.) were deployed along the continental shelf in each region at natural reefs, artificial structures, or unconsolidated bottom, and were retrieved annually or biannually for maintenance and data downloads throughout the study period (June 2021-July 2025; Supplemental Figure 6.1). Additional detection data were provided from two collaborative acoustic telemetry networks: the Florida Atlantic Coast Telemetry Array (FACT) and Integrated Tracking of Aquatic Animals in the Gulf (iTAG).

Tagging

Greater amberjack (n = 498) were tagged across the study area from June 2021 to August 2024 using coded acoustic transmitters (Table 6.1; Figure 6.1; primarily V16-4H, 69 kHz, nominal delay = 90s, battery life ~ 6 y, Innovasea Systems Inc.). A subset of tags were equipped with pressure sensors to capture information on amberjack depth use (n = 71; V16P-4H, 69 kHz, Innovasea Systems Inc.). Fish were captured with hook and line and surgically implanted with transmitters. Each fish was also tagged externally with one or two yellow and red nylon dart tags in the musculature under the dorsal fin. Each external tag was inscribed with a unique identification number, a telephone number to report the tag, and an indication that the tag could be returned for a $250 reward.

Habitat Use

Patterns of residency by greater amberjack were compared across study regions and habitat types. All receivers in Alabama and Mississippi were deployed for short intervals around calibration studies (August 2022 and April-May 2024) and were excluded from analysis due to this short duration. Acoustic detections were first aggregated to a daily timestep, with an observation for each unique fish, date, and

location combination in the dataset. Next, any detections that occurred in the first three days after tagging were removed for each individual to account for changes in behavior due to tagging. Continuous detection periods were then calculated to compare patterns of residency by Greater Amberjack. The continuous detection period is defined as the number of days that an individual fish was detected on a site starting when the fish was first detected at that site and ending if the fish left for three or more days (Everett et al, 2020). Continuous detection periods were compared both between region (Atlantic, Eastern Gulf Western Gulf) and habitat type (natural, artificial) using a two-way ANOVA with a subsequent Tukey Honest Significant Differences Test to examine differences in groups ($\alpha = 0.05$). To examine seasonal patterns of habitat use, we modelled the probability of detection via Generalized Additive Mixed Model (GAMM) for each region x habitat combination. Data were aggregated to a daily scale for each tagged individual; for each possible habitat type (i.e. natural, artificial), fish were given a 1 on days they were detected at that habitat type and a 0 for days they were not, from the first day of detection to the last day that fish was detected in the study. This binary variable was used as the response in the GAMMs, with day of year as a fixed effect and individual fish included as a random intercept term to account for potential variation in individual behavior. Models were developed with a binomial distribution using a logit link function and cubic cyclic regression splines. A separate model was constructed for each possible habitat and region combination (i.e. six models total). Greater amberjack depth use was characterized by using detections from depth-coded acoustic transmitters deployed across all study regions. Individuals detected fewer than five days were excluded from the depth use analysis. All data analysis was conducted in R and all GAMM modelling was conducted using the package *mgcv* (R Core Team 2025; Wood, 2011).

Movement and Connectivity

Biological and temporal factors influencing the movement behavior of greater amberjack in each study region were assessed using a two stage hurdle modeling approach with a GAMM framework. The first model assessed the probability of an individual fish moving using a GAMM with a “logit” link function. First, the movement data were coded as binary, with 1 for a movement and 0 for a nonmovement for each date and individual combination. A nonmovement is defined as an individual being detected on the same site as they were within the prior three days. A movement is defined as either a fish being absent from a site for at least three days, being detected moving from one site to another, or a fish arriving to a site from which they were absent for at least three days. This binary variable was used as the response in the GAMM model, with day of year, lunar phase, and size of fish (fork length) considered as predictors and individual fish included as a random intercept term to account for potential variation in individual behavior. Day of year and lunar phase were modelled with a cubic cyclic regression spline and fork length was modelled using penalized thin plate regression splines. One model was created for each region, using backwards selection based on minimizing AIC to remove non-significant predictors ($\alpha = 0.05$, $\Delta AIC < 2$). The second stage in the hurdle model assessed factors influencing movement magnitude using the rate (distance/day) of movement by Greater Amberjack. To obtain a daily rate of movement, the distance travelled (calculated by Haversine Formula) was divided by the number of days from the date that the fish was detected at the start point to the date the fish was detected at the end point of a movement event. Movements that occurred over more than 90 days were removed from further analysis. This daily average rate of movement was then assigned to the middle date between the start and end of the movement. We then modelled this movement rate as the response variable in a GAMM model with day of year and fish size to examine seasonal cycles in movement rates. Movement rates

were log-transformed examined for normality. These transformed movement rates did not follow a normal distribution, thus generalized additive mixed models with scaled t-distribution and an identity link were constructed using movement rate as the response and day of year and fork length as possible predictors, including individual fish as a random intercept term. However, individual was excluded as a random intercept from the Atlantic and Eastern GoM stage 2 models given low observations per individual (Atlantic = 5.4; Eastern GoM = 2.3 observations per individual; Zuur et al., 2009). Day of year was modelled using cubic cyclic regression splines, while fork length was modelled using thin plate regression splines. Backwards selection was performed to remove non-significant predictors based on minimizing AIC (α= 0.05, ΔAIC < 2).

We examined how distance moved by greater amberjack varied across spatial and temporal, gradients. For this analysis we included all acoustic detections and angler recaptures. We calculated both overall displacement and distance moved using the Haversine formula. We define displacement as the furthest straight-line distance a fish is detected from its initial tagging location. Distance moved is the straight-line distance between subsequent detections, when they occur on different receivers. Due to the uncertainty associated with the detection radius of acoustic receivers, movements less than 2 km were removed from these analyses. We used a one-way ANOVA and Tukey Honest Significant Differences Test to examine differences in mean displacement between regions.

Network analysis was used to assess seasonal movement behaviors of greater amberjack across regions. A network analysis can be used to describe trends in movement and habitat use across geographic space based on nodes (detection locations) and edges (movements; Jacoby et al., 2012). Individual acoustic receiver locations are frequently used as nodes in spatial network analysis, however, to examine broad patterns of movement over a large area we use a gridded approach to assign nodes (Jacoby et al., 2012). A 100 km grid was overlaid across the study area; receiver locations were grouped together based on which grid cell they fell within. These grid cells served as nodes in our network analysis, with movements between nodes comprising edges in the network. In addition to acoustic tag detections, we also include recapture and harvest data of acoustically tagged greater amberjack to capture movements to areas with poor acoustic receiver coverage. To assess seasonal changes in movement behavior and connectivity, we assigned seasons to movements based on the time of year they occurred (March-May: Spring, June-August: Summer, September-November: Fall, and December-February: Winter). Movements that occurred across multiple seasons were excluded from this part of the analysis. We then constructed four directed networks, one for each season, using the package igraph (Csardi & Nepusz, 2006). Community structure was identified using the Walk-Trap approach, which is appropriate for directed networks and has been successfully used to model movements of marine fishes (Finn et al., 2014). To assess population connectivity between regions regardless of season, we again employ a network analysis approach. An overall network was constructed consisting of all movements, even those that occurred across seasons. Nodes were assigned based on the same 100 km grid as above.

Results

A total of 319 of the 498 tagged Greater Amberjack were detected (64.1%) at receiver sites (n = 159) from Texas to North Carolina (Figure 6.1), with approximately 7 million detections recorded during the study period (Table 6.1). In the Western Gulf 139 of the 226 tagged greater amberjack were detected (61.5%), in the Eastern Gulf 98 of the 123 tagged fish were detected (79.7%), and in the Atlantic 82 of the 149 tagged fish were detected (55.0%). Most fish were detected within the study area. However,

some fish moved outside of the boundaries, one individual was detected moving from North Carolina to the Bahamas and another individual tagged in North Carolina was recaptured, but not detected, approximately 40 km off the coast of Belize (~2343 km movement). Overall mean size of fish tagged in the study was 945 mm, mean size was similar in the Atlantic and Eastern Gulf (828 mm and 834 mm, respectively) but was larger in the Western Gulf (1081 mm; Table 6.1).

Habitat Use

Mean residency of greater amberjack varied by region and habitat type (ANOVA; $p < 0.0001$). Mean residency was significantly lower in the Atlantic (7.76 ± 17.99 days, max = 205 days; mean ± SD) than either region of the Gulf (Tukey HSD; W Gulf x ATL: $p < 0.0001$; E Gulf x ATL: $p = 0.0005$; Figure 6.2). However, residency did not differ (Tukey HSD; $p = 0.35$) between the Eastern (21.08 ± 39.12 days, max = 293 days) and Western Gulf (16.59 ± 50.70 days, max = 632 days). Although greater amberjack residency occasionally spanned a year or more, residency was generally much lower in all regions; in the Eastern Gulf 79.8% of residency periods were thirty days or less, in the Atlantic and Western Gulf this number was even lower (94.2% and 89.6%, respectively). Some greater amberjack were only detected at either natural (n = 97) or artificial habitats (n = 91), but many were observed using both habitat types to varying degrees (n = 92; Supplemental Figure 6.2). Overall, greater amberjack displayed greater residency at artificial (20.58 ± 52.60 days) compared to natural reefs (10.78 ± 36.85 days); however, a significant interaction was detected, as habitat-specific residency patterns varied across regions (Supplemental Figure 6.3). In the Western Gulf residency was significantly higher on artificial (40.2 ± 79.1 days) compared to natural habitats (9.9 ± 36.6 days; Tukey HSD: $p < 0.0001$). Conversely in the Eastern Gulf residency was significantly higher on natural (38.4 ± 54.8 days) than artificial habitats (10.3 ± 18.0 days; Tukey HSD: $p = 0.00015$). There was no significant difference in continuous detection period in the Atlantic between habitat types (Natural: 3.1 ± 5.6 days; Artificial: 8.8 ± 19.6 days; Tukey HSD: $p = 0.76$).

There was a significant effect of day of year on probability of detection for both habitat types in the Atlantic, Eastern Gulf, and Western Gulf (GAMM, $p < 0.0001$; Figure 6.3). Probability of detection at natural habitats was highest in all three regions in the late winter/spring and lowest in the fall months (September-October). In the Atlantic and Eastern Gulf regions probability of detection at artificial habitats was lowest in the summer and early fall, but in contrast was highest during this period in the Western Gulf.

Depth use data from fish tagged with pressure sensor tags were summarized to show regional patterns in depth use (Supplemental Figure 6.4). Individuals with fewer than five days of detection data were excluded. After filtering those fish, depth use data from 44 individuals detected across 27 unique stations remained (Supplemental Tables). In the Eastern Gulf, depth coded transmitters were only detected on one site, therefore these detections were pooled with the Western Gulf. Mean depth detected in the pooled Gulf was 48.7 ± 17.1 m and 20.7 ± 15.12 m in the Atlantic (mean ± SD). Sites at which depth tags were detected varied between the Gulf and Atlantic (Gulf: mean = 92.0 m, range 42-197 m; Atlantic: 38.5 m, 18-150 m; Supplemental Tables). However, one individual (not depth tagged) was recaptured at a floating oil and gas platform off of Texas in 1461 m of water, indicating that the species likely inhabits greater depths than those captured by depth coded transmitters. Mean depth was generally in the lower third of the water column, with some exceptions (Supplemental Tables), and tended to be closer to the bottom in the Atlantic and at shallower sites.

Movement and Connectivity

Movement behavior was modelled using a two-stage GAMM for each region. Probability of a movement event was modelled in the first stage of this approach. In the Atlantic region, the final model based on backwards selection included day of year as a fixed effect and individual as a random intercept term (DE = 26.2%). Movement probability was highest during the summer months (June, July) and lowest in the fall (October; Figure 6.4). In the Eastern Gulf, the final movement probability model included day of year, size, and moon phase as fixed effects and individual as a random intercept term (DE = 20.5%). Movement probability was highest in the late summer and early fall (July-October) and lowest in late winter (February). Movement probability was also positively associated with fish size and was highest just before the full moon (Figure 6.4). In the Western Gulf the final movement probability model included day of year and fish size as fixed effects and individual as a random intercept term (DE = 28.9%). There were two peaks in movement probability, one in the spring (March-April) and a smaller peak in the fall (September-October). Additionally, movement probability was highest in fish >1000mm but decreased at the largest sizes observed (Figure 6.4).

The second stage of the hurdle model examined movement magnitude by greater amberjack. In the Atlantic, the final stage 2 model included day of year and fish size as fixed effects (DE = 23.8%). Movement magnitude was greatest from late fall and early spring (November-March) and lowest in the summer (June-August). Movement magnitude was highest in fish ~1100 mm and lowest in the smallest and largest fish tagged (Figure 6.5). In the Eastern Gulf, the final movement magnitude model included only day of year as a fixed effect (DE = 10.7%). Movement magnitude was greatest in the summer (July-August) and lowest in the late winter (February; Figure 6.5). In the Western Gulf the final model included day of year and fish size as fixed effects and individual as a random intercept (DE = 40.0%). Movement magnitude was greatest in the spring (March-April) and lowest in the fall (September-October) and increased with fish size.

Greater amberjack displacement varied by region (ANOVA; $p < 0.0001$). Among fish that moved, greater displacement from tagging sites was observed in the Atlantic (mean = 376.6 km, max = 2343.4 km) relative to both Gulf regions (Tukey HSD; $p < 0.0001$), while no difference in displacement was observed between the Western Gulf (mean = 35.4 km, max = 322.4 km) and the Eastern Gulf (mean = 24.2 km, max = 192.8 km; Supplemental Figure 6.5) (Tukey HSD; $p = 0.97$).

Exchange among the three primary regions was low. No fish tagged in the Atlantic were detected in either of the Gulf regions, and no fish tagged in the Western Gulf were detected in either the Atlantic or Eastern Gulf regions. One acoustically tagged fish released in Alabama (0.8% of fish tagged in the Eastern Gulf) was recaptured and reported by a recreational angler just west of the Mississippi River Delta in Louisiana (Figure 6.6). Additionally, two fish tagged in the Atlantic region (1.3% of fish tagged in this region) moved into the Caribbean. Interstate movements were observed by 33 individuals across the study area by fish tagged in each state except Florida (6.6% of total fish tagged). Twenty-three greater amberjack tagged in North Carolina were detected or recaptured outside of the state (27.8% of fish tagged in North Carolina). Twenty of these individuals (87%) were last detected on the Atlantic Coast of Florida, one of which was detected in South Carolina along the way. Of the remaining two, one was detected in the Bahamas while the other was detected in the Florida Keys and subsequently recaptured in Belize. Eight of the greater amberjack tagged in South Carolina were detected or recaptured outside of the state (32.0% of fish tagged in SC). Seven of these were last detected on the

Atlantic Coast of Florida, one of which was detected in Georgia on the way. The remaining individual was last detected in Georgia. Some movements were observed between Texas and Louisiana. Two individuals tagged in Louisiana moved to Texas and returned to Louisiana, one of which made this movement twice (1.3% of fish tagged in Louisiana). One individual tagged in Texas was detected in Louisiana and then moved back to Texas (1.3% of fish tagged in Texas). The remaining interstate movement was the aforementioned fish tagged in Alabama that was recaptured in Louisiana (1.2% of fish tagged in Alabama).

Network analysis of amberjack movement across the study area revealed ten unique communities and a network that consisted of 76 edges and 43 nodes. In the Atlantic, three connected communities were identified, one connecting North Carolina and southeast Florida and the Florida Keys, one connecting South Carolina to southeast Florida, and a third community represented by movements within the Bahamas (Figure 6.6). Relatively little connectivity was observed between the groups in North and South Carolina. In the Eastern Gulf, two connected communities were identified on the West Florida Shelf, with a third community on the Alabama shelf. Four communities were identified in the Western Gulf: the eastern Louisiana shelf, the outer shelf of central Louisiana, the outer shelf of northeast Texas, and south Texas (Figure 6.6). Limited connectivity was observed among adjacent communities from central Louisiana to south Texas; however, none of these communities were connected with the community in eastern Louisiana. Instead, we observed limited connectivity between eastern Louisiana and the eastern Gulf region off Alabama.

Seasonal patterns in connectivity were also examined using network analysis. The spring network consisted of 19 edges spanning 14 nodes, and four distinct communities were identified (Figure 6.7). Movements in the Atlantic region comprised two communities, originating along the coast of the Carolinas and connecting in South Florida. In the Gulf there was a small community of two adjacent nodes in Western Florida and another separate community consisting of four nodes along the outer continental shelf of Louisiana and Texas. The summer network consisted of 12 edges spanning 11 nodes. The Walktrap cluster detection method identified five small communities. Three unconnected communities consisting of relatively short, localized movements were identified in the Atlantic region, one North Carolina with two nodes, one in South Carolina with three nodes, and one in southeast Florida with two nodes. In the Gulf, two connected communities were identified consisting of two nodes each were identified along the outer shelf of Louisiana and Texas (Figure 6.7). The fall network consisted of 14 edges and 14 nodes. The Walktrap cluster detection method identified four communities. The largest consisted of four nodes representing movements within North Carolina that was connected to a second community on the southeastern coast of Florida (Figure 6.7). Additionally in the Atlantic, there was a small community of three nodes representing movements within South Carolina. A fourth community consisting of 4 nodes was detected for Louisiana and Texas, again consisting of movements along the continental shelf. Finally, the winter network comprised 23 edges and 18 nodes and was the most complex of the seasonal movements. Cluster detection identified five connected communities. In the Atlantic, the largest comprised ten nodes of interconnect movements between the east coast of Florida, South Carolina, and North Carolina connected to a community of two nodes describing localized movements in North Carolina; one additional community was identified of movements within the Bahamas. In the Gulf , two disconnected communities consisting of local movements between two nodes were observed, one in Louisiana and one in western Florida. (Figure 6.7).

Discussion

Habitat Use

Overall, greater amberjack residency was variable among individuals, regions, and habitats. Mean residency was generally lower than typically observed in other reef fishes (e.g. snappers, jacks; Green & Starr, 2011; Piraino & Szedlmeyer, 2014; Bacheler et al., 2019; Gandra et al., 2020; Bacheler et al., 2021) However, residency periods varied by individual amberjack and ranged from a single day to over one year on a site, suggesting some individuals are more site attached than others. Regionally, lower residency in the Atlantic could reflect more variable environmental conditions in the Atlantic relative to the Gulf that might drive greater amberjack to move more frequently (Mammel et al., 2022). Alternatively, low residency in the Atlantic could result from high hardbottom habitat continuity along the continental shelf, allowing greater amberjack to move more easily between habitat patches (Ojeda et al., 2004; Caldwell & Gergel, 2013). Studies of residency behavior by reef fishes in the Gulf have largely focused on artificial habitats (e.g. Piraino & Szedlmeyer, 2014; Everett et al., 2020), and our results present a comparison of residency behaviors at both natural and artificial habitats. The finding that greater amberjack residency was higher on artificial habitat could reflect a preference for artificial habitats (Paxton et al., 2020). More likely, artificial habitats (e.g. petroleum platforms, artificial reefs) function as isolated habitat patches with high densities of prey, encouraging higher site-specific residency. In contrast, natural reefs are typically more broadly distributed across the seafloor, facilitating more frequent movement between patches and reduced residency at specific locations (Caldwell & Gergel, 2013).

Seasonal changes in habitat use by fishes are widely recognized across taxa and ecosystems, and may be related to changes in environmental conditions, forage availability, or reproductive behavior (e.g. Luo et al., 2009; Kock et al., 2013; Everett et al., 2020). In all regions, greater amberjack were most likely to be detected on natural habitats in the late winter and early spring, coinciding with the known spawning period for the species (Harris et al., 2007). Greater amberjack spawning behavior has been observed on natural habitat throughout their range in the northwestern Atlantic but is yet to be described on artificial structures (Graham and Castellanos 2005, Farmer et al., 2017). However, we observed a similar relationship in both the Atlantic and Eastern Gulf at artificial habitat, indicating that spawning may be occurring on these structures as well. Alternatively, these seasonal trends of habitat use could be related to foraging, as availability of prey resources often vary through the year (Gandra et al., 2020; Lovell et al., 2024)

Depth is a key environmental gradient that influences the distribution and habitat preferences of marine organisms and is particularly important to structuring reef fish communities (Garner et al., 2019).Greater amberjack demonstrated variable depth use patterns, but were generally found on the continental shelf (depth < 150 m), consistent with past studies in the Gulf (Murie et al., 2013; Jackson et al., 2018; Boyle et al., 2022). The use of shallower depths in the Atlantic relative to the Gulf may reflect regional differences in salinity or distribution of structured habitat across the shelf. The Mississippi River plume influences salinity regimes across the inner shelf of the northwestern Gulf (where most of the Western Gulf depth tags were deployed), and it is likely that greater amberjack avoid shallow nearshore habitats in this region that are characterized by turbid and low salinity conditions (Da Silva & Castelao 2018; Mammel et al., 2023). Likewise, much of the reef habitat in the Gulf occurs at the shelf edge, while habitat along the Atlantic is distributed across shallower regions of the shelf (Ojeda et al., 2004). It

is also possible that the observed pattern simply reflects the distribution of receivers deployed in each region. While our detection data supports the idea that greater amberjack primarily inhabit the continental shelf, it is important to note that the vast majority of acoustic telemetry infrastructure in both regions is in shelf waters. Indeed, the recapture reported in this study at a floating platform in 1461 m of water and international crossings to the Bahamas and Belize indicate that greater amberjack in the northwestern Atlantic are capable of traversing open ocean environments with much greater bottom depths than previously observed in U.S. waters (e.g. Johnston et al., 2022; Bacheler et al., 2022). Greater amberjack have been observed in open ocean environments in other regions of the world exhibiting behavior similar to pelagic fishes in near-surface waters, making foraging dives and diel vertical migrations (Tone et al., 2021; Tone et al., 2022). Our data suggests greater amberjack in the northwestern Atlantic traverse open water environments and may inhabit structure in much deeper water than previously reported, supporting the notion that the species inhabits both reef and pelagic ecosystems (Tone et al., 2022). Alternative telemetry methods, such as satellite tags or internal archival tags, would potentially elucidate these behaviors not captured by acoustic telemetry.

Movement and Connectivity

Animal movements are influenced by external factors and internal states that affect navigation, motion capacity, and ultimately willingness to change locations. In the current study, movement behavior of greater amberjack was reflective of both seasonal and size specific cues. Temporal patterns in movement are often influenced by seasonal shifts in temperature regimes, foraging opportunities, and/or reproductive cues (e.g. Barbour et al., 2014; Griffiths et al., 2018). The likelihood of movement for greater amberjack varied temporally and between regions, possibly reflecting changes in environmental conditions or forage over space and time (Champion et al., 2018; Mammel et al., 2022). However, rate of movement generally peaked in winter/spring during the spawning season (i.e. spring; Harris et al., 2007), coinciding with peaks in probability of detection at both natural and artificial sites. Paradoxically, the eastern Gulf experienced the lowest movement (likelihood and rate) during this spring window, however, those fish were tagged in late January (i.e. once the spawning season has started) and thus movements may represent post-spawn dispersal. Similarly, the probability of movement and rate of movement generally increased with fish size, which is perhaps not surprising given that motion capacity increases and predation risk decreases with size. While larger fish may move more as part of local foraging behavior to maintain higher energetic demands, changes to internal states (e.g. maturity) may also initiate ontogenetic habitat shifts or spawning migrations (Dunlop et al., 2015; Espinoza et al., 2015; Lowerre-Barbierri et al., 2021). Given the timing, increased movement with size likely also reflects the fact that we would only expect larger, mature individuals to make spawning movements.

The Atlantic region was characterized by high within-region connectivity from the Carolinas to Southeast Florida, driven by directed north-south movements during winter and spring. Cape Hatteras, North Carolina represents a well-known biogeographic boundary between the South Atlantic and Mid-Atlantic Bights (Fautin et al., 2010) near the northern limit of distribution for the species (Schultz et al., 2025). Thus, directed southern migrations in the Atlantic along this latitudinal gradient might be expected during winter to avoid suboptimal temperatures (Lédée et al., 2015; Secor et al., 2020; Mammel et al., 2022). In addition, the peak intensity of these long movements in late winter and spring coincided with the spawning season, which is consistent with amberjack spawning migrations of similar magnitude in the western Pacific (Tone et al., 2021). Indeed, there is growing evidence of greater amberjack spawning aggregations in the Florida Keys (Harris et al., 2007; Lowerre-Barbieri et al., 2021), and it seems likely

that the consistent seasonal movement from the Carolinas to the Florida Keys over the three-year study period represented spawning migrations to these aggregation sites. Amberjack spawning and oocyte maturation is influenced by increasing photoperiod and temperature regime (Tone et al., 2021), and migration to warmer waters of the Florida Keys may provide the temperature regime necessary for reproduction. It is also worth noting that spawning in the Keys likely provides additional oceanographic benefits for larval dispersal, as the Gulf Stream facilitates northward dispersal along the Atlantic coast.

Greater amberjack migrating from the Carolinas were rarely detected between the Carolinas and SE Florida despite substantial receiver coverage along the shelf, suggesting fish may not follow the coastline. Instead, individuals may move along the shelf edge or take a more direct route over open water. Although our depth data supports previous studies that greater amberjack in the SE USA generally associate with shelf habitats (Garner et al., 2019; Bacheler et al., 2022), tag recaptures at deepwater platforms and in the Bahamas and Belize indicate that they are capable crossing open ocean environments, a behavior observed elsewhere in their range (Tone et al., 2021). Interestingly, there were no observations of amberjack returning northward once they were detected in Southeast Florida across our three-year study. It is unclear whether these individuals remained near presumed spawning locations to avoid energetic costs of a return migration, were subjected to increased mortality (e.g., predation or fishing), or simply returned to different sites off the Carolinas without receiver coverage. The latter seems unlikely, given the large sample size and length of the study, but evidence of northward movements has been observed in Greater Amberjack tagged in the Florida Keys (Lowerre-Barbieri, pers com).

Connectivity within both Gulf regions was characterized by smaller spatial networks and localized movement (primarily within-state). Similar to the Atlantic, peaks in movement and connectivity aligned with the spawning window, supporting the notion that reproduction is a major driver of amberjack movement (Tone et al. 2021). The area of highest connectivity in the Gulf was along the outer shelf of the western Gulf, where numerous natural reef banks and artificial structures (e.g., oil and gas platforms) may serve as “stepping stones” for movement (Schulze et al. 2020). These habitats have been identified as areas of high productivity for amberjacks and other species (Gallaway et al., 2021; Sanchez et al., 2023) and the banks are believed to be important spawning sites for a variety of reef fishes (Grüss et al. 2018). In contrast to the Atlantic, movement patterns were less consistent and directed migratory behavior was not observed in the Gulf, suggesting that spawning movements may be more localized towards specific habitats and spawning sites. Outer shelf waters of the Gulf remain relatively warm throughout the year as evidenced by the fact that several banks support coral communities, which may preclude the need for amberjack to migrate to find suitable temperature regimes for spawning. Similar patterns were described in the western Pacific, with individuals from more northern latitudes making spawning migrations, while more southernly distributed individuals did not migrate prior to spawning (Tone et al. 2021). Alternatively, it is possible that the lack of migratory behavior in the Gulf simply reflects the absence of acoustic telemetry infrastructure in Mexican waters of the southern Gulf that prevented detection of migrating individuals. However, we believe this is unlikely given that many individuals were tracked for several years in the northern Gulf and similar migrations were not observed in the eastern Gulf where receiver infrastructure in south Florida is substantial. In general, greater amberjack in the Gulf rarely crossed state boundaries; however, a few individuals were detected/recaptured up to 322 km from their tagging location, suggesting larger movements do occur. Infrequent, long movements have been described for greater amberjack in the Gulf (McClellan and

Cummings 1997; Ingram & Patterson, 2001; Murie et al., 2011) and may be examples of wandering behavior that facilitates adult dispersal as a mechanism of gene flow, a strategy which is described among fishes in the marine environment (e.g. Cermeño et al., 2015, Schiavon et al., 2025).

Connectivity between populations influences population dynamics and therefore is a critical component of stock delineation and fisheries management (Cadrin 2020; Lédée et al., 2021). We found that exchange between regions was low, with some evidence of movement between the eastern and western Gulf, but no movements between the Gulf and Atlantic. Interestingly, there were a couple individuals tagged in the Atlantic that moved both across international boundaries and into basins outside of the study area (i.e. movements to Belize, Bahamas), demonstrating the species is capable of inter-basin movement. Similar movements to the south have been described in previous tagging studies, with greater amberjack tagged in the United States moving to the Bahamas, Jamaica, and Mexico (McClellan and Cummings 1997; Murie et al., 2011). Drivers for these international movements are unclear, but it is possible movements could be the result of roaming behavior or directed movements to alternative spawning aggregations that are known to occur in those areas (Graham & Castellanos, 2005; Heyman & Kjerve, 2008). Regarding exchange between study regions, our results support findings from previous genetics and conventional tagging studies that found low (<2%) exchange between the Atlantic and Gulf (McClellan and Cummings 1997; Murie et al., 2011; Hargrove et al., 2018). While our study observed no adult movement between these basins, limited connectivity may still exist through larval transport to the Atlantic via larvae spawned in the Gulf and to the Gulf via Atlantic fish spawning in the Caribbean (Qian et al., 2015). Limited exchange between the eastern and western Gulf (with no western individuals moving east), suggests some sub-structuring likely exists on either side of the Mississippi River Delta, which is consistent with other species in the Gulf (e.g., red snapper, king mackerel; Dance and Rooker 2019, Banks et al. 2024) however, data from both this and concurrent studies (Lowerre-Barbierri pers com) clearly demonstrate that movements across this boundary do occur.

The results of this study present the most comprehensive assessment of greater amberjack movement to date and highlight how amberjack movement and habitat use varies across spatial and temporal scales in U.S. waters. Habitat use and residency patterns were region- and habitat-specific, reflecting differences in the composition and distribution of habitat, with lower residency in areas of more continuous available habitat and higher residency at more isolated structures in the Gulf. Clear temporal trends in residency and movement were common to each region with movements peaking during winter/spring and coinciding with reproductive timing. Similarly, regional differences in movement patterns were tied to differing reproductive strategies between the Atlantic and Gulf to optimize temperature regimes for spawning and oceanography for larval transport. Greater amberjack in the Atlantic undergo lengthy migrations to warmer waters of the Florida Keys, while Gulf amberjack moved shorter distances to specific habitats (e.g., natural reefs) during the spawning season. Despite large movements by greater amberjack in the Atlantic, we found no evidence of connectivity between the Atlantic and Gulf basins and limited exchange between the eastern and western Gulf and across state lines. Collectively these patterns support potential sub-structuring within the Gulf stock. Given distinct movement behaviors across U.S. waters, spatially explicit management strategies that account for regional variations in connectivity, movement, and residency may be beneficial for greater amberjack fisheries management.

References

Allen, A. M. & Singh, N. J. (2016). Linking movement ecology with wildlife management and conservation. *Frontiers in Ecology and Evolution 3*(155). doi: 10.3389/fevo.2015.00155

Alos, J., Palmer, M., Arlinghaus, R. (2012). Consistent selection towards low activity phenotypes when catchability depends on encounters among human predators and fish. *PloS ONE 7*(10), e480030. https://doi.org/10.1371/journal.pone.0048030

Bacheler, N. M., Shertzer, K. W., Runde, B. J., Rudershausen, P. J., & Buckel, J. A. (2021). Environmental conditions, diel period, and fish size influence the horizontal and vertical movements of red snapper. *Scientific Reports, 11*(9580). https://doi.org/10.1038/s41598-021-88806-3

Bacheler, N. M., Gregalis, K. C., Gillum, Z. D., Pickett, E. P., Schobernd, C. M., Schobernd, Z. H., & Teer, B. Z. (2022). Using stationary video data to infer relative abundance and distribution of four *Seriola* species along the southeast United States Atlantic coast. *Fisheries Research*, *249*, Article 106238. https://doi.org/10.1016/j.fishres.2022.106238

Banks K. G., Streich M. K., Stunz, G. W. (2024) Age, growth, and mortality of King Mackerel in the western Gulf of Mexico, Marine and Coastal Fisheries, 16: e210278, https://doi.org/10.1002/mcf2.10278

Barbour, A. B., Adams, A. J., & Lorenzen, K. (2014). Size-based, seasonal, and multidirectional movements of an estuarine fish species in a habitat mosaic. *Marine Ecology Progress Series 507*, 263-276. doi: 10.3354/meps10837

Boyle, K. S., Hightower, C. L., & Powers, S. P. (2022). Diel patterns of depth use and swimming activity of post-release greater amberjack (*Seriola dumerili*) in the northern Gulf of Mexico. *Environmental Biology of Fishes*, *106*, 491–518. https://doi.org/10.1007/s10641-022-01385-8

Cadrin, S. X. (2020). Defining spatial structure for fishery stock assessment. *Fisheries Research 221,* Article 105397. https://doi.org/10.1016/j.fishres.2019.105397

Caldwell, I. R. & Gergel, S. E. (2013). Thresholds in seascape connectivity: influence of mobility, habitat distribution, and current strength on fish movement. *Landscape Ecology 28*, 1937-1948. DOI 10.1007/s10980-013-9930-9

Cermeño, P., Quílez-Badia, G., Ospina-Alvarez, A., Sainz-Trápaga, Boustany, A. M., Seitz, A. C., Tudela, S., & Block, B. A. (2015). Electronic tagging of Atlantic Bluefin Tuna (*Thunnus thynnus,* L.) reveals habitat use and behaviors in the Mediterranean Sea. *PLoS ONE 10*(2), Article e0116628. doi:10.1371/journal.pone.0116638

Champion, C., Hobday, A. J., Zhang, X., Pecl, G. T., & Tracey, S. R. (2018). Changing windows of opportunity: past and future climate-driven shifts in temporal persistence of kingfish (*Seriola lalandi*) oceanographic habitat within south-eastern Australian bioregions. *Marine and Freshwater Research 70*(1), 33-42. https://doi.org/10.1071/MF17387

Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems,* 1695. https://igraph.org.

Da Silva, C. E. & Castelao, R. M. (2018). Mississippi River plume variability in the Gulf of Mexico from SMAP and MODIS-Aqua Observations. *JGR Oceans 123*(9), 6620-6638.

https://doi.org/10.1029/2018JC014159Digital Object Identifier (DOI)

Dance, M. A. & Rooker, J. R. (2015). Habitat- and bay-scale connectivity of sympatric fishes in an estuarine nursery. *Estuarine, Coastal and Shelf Science 167B,* 447-457. https://doi.org/10.1016/j.ecss.2015.10.025

Dance M. A., & Rooker J. R. (2019) Cross-shelf habitat shifts by red snapper (Lutjanus campechanus) in the Gulf of Mexico. *PLoS ONE* 14(3): e0213506. https://doi.org/10.1371/journal.pone.0213506

Dunlop, S. W., Mann, B. Q., Cowley, P. D., Murray, T. S., & Maggs, J. Q. (2015). Movement patterns of *Lichia amia* (Teleostei: Carangidae): Results from a long-term cooperative tagging project in South Africa. *African Zoology 50*(3), 249-257. https://doi.org/10.1080/15627020.2015.1058724

Espinoza, M., Lédée, E. J. I., Simpfendorfer, C. A., Tobin, A. J., & Heupel, M. R. 2015. Contrasting movements and connectivity of reef associated sharks using acoustic telemetry: implications for management. Ecological Applications 25(8): 2101-2118. https://doi.org/10.1890/14-2293.1

Espinoza, M., Lédée, E. J. I., Smoothey, A. F., Heupel, M. R., Peddemors, V. M., Tobin, A. J., & Simpfendorfer, C. A.(2021). Intra-specific variation in movement and habitat connectivity of a mobile predator revealed by acoustic telemetry and network analyses.*Marine Biology 168*, Article 80. https://doi.org/10.1007/s00227-021-03886-z

Everett, A. G., Szedlmeyer, S. T., & Gallaway, B. J. (2020). Movement patterns of red snapper *Lutjanus campechanus* based on acoustic telemetry around oil and gas platforms in the northern Gulf of Mexico. *Marine Ecology Progress Series 649*, 155-173. https://doi.org/10.3354/meps13448

Farmer, N. A., Heyman, W. D., Karnauskas, M., Kobara, S., Smart, T. I., Ballenger, J. C., Reichert, M. M. J., Wyanski, D. M., Tishler, M. S., Lindeman, K. C., Lowerre-Barbieri, S. K., Switzer, T. S., Solomon, J. J., McCain, K., Marhefka, M., & Sedberry, G. R. 2017. Timing and locations of reef fish spawning off the southeastern United States. *PLoS ONE*, 12(3): 1-35. https://doi.org/10.1371/journal.pone.0172968

Finn, J. T., Brownscombe, J. W., Haak, C. R., Cooke, S. J., Cormier, R., Gagne, T., & Danylchuk, A. J. (2014). Applying network methods to acoustic telemetry data: Modeling the movements of tropical marine fishes. *Ecological Modelling 293*, 139-140. http://dx.doi.org/10.1016/j.ecolmodel.2013.12.014

Frisk, M. G., Jordaan, A., Miller, T. J. (2014). Moving beyond the current paradigm in marine population connectivity: are adults the missing link? *Fish and Fisheries 15*, 242-254. https://doi.org/10.1111/faf.12014Digital Object Identifier (DOI)

Gallaway, B. J., Raborn, S. W., McCain, K. A., Beyea, R. T., Dufault, C., Heyman, W., Putnam, N. F., & Egerton, J. (2021). Absolute abundance estimates for red snapper *Lutjanus campechanus*, greater amberjack *Seriola dumerili* and other federally-managed fish on offshore petroleum platforms in the Gulf of Mexico. *North American Journal of Fisheries Management, 41(6)*. https://doi.org/10.1002/nafm.10678

Gandra, M., Afonso, P., & Fontes, J. 2020. Intra- and interspecific associations in two predatory fishes at a shallow seamount. Marine Progress Ecology Series, 654: 93-107. https://doi.org/10.3354/meps13501

Garner, S. B., Boswell, K. M., Lewis, J. P., Tarnecki, J. H., & Patterson, W. F. (2019). Effect of reef morphology and depth on fish community and trophic structure in the northcentral Gulf of Mexico. *Estuarine, Coastal and Shelf Science*, *230*, 1–12. https://doi.org/10.1016/j.ecss.2019.106423

Graham, R. T. & Castellanos, D. W. 2005. Courtship and spawning behaviors of carangid species in Belize. Fisheries Bulletin, 103:426-432.

Green, K. M., & Starr, R. M. (2011). Movements of small adult black rockfish: Implications for the design of MPAs. *Marine Ecology Progress Series*, *444*, 219–230. https://doi.org/10.3354/meps

Griffiths, C. A., Patterson, T. A., Blanchard, J. L., Righton, D. A., Wright, S. R., Pitchford, J. W., & Blackwell, P. G. (2018). Scaling marine fish movement behavior from individuals to populations. *Ecology and Evolution 9,* 7031-7043. DOI: 10.1002/ece3.4223

Grüss, A., Biggs, C., Heyman, W. D., Erisman, B. (2018). Prioritizing monitoring and conservation efforts for fish spawning aggregations in the U.S. Gulf of Mexico. *Scientific Reports 8*: Article 8473. DOI:10.1038/s41598-018-26898-0

Hargrove, J. S., Murie, D. J., Parkyn, D. C., Saarinen, E. V., & Austin, J. D. 2018. Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico. Genetica 146: 393-402. https://doi.org/10.1007/s10709-018-0031-1

Harris, P. J., Wyanski, D. M., White, D. B., and Mikell, P. P. 2007. Age, Growth, and Reproduction of Greater Amberjack off the Southeastern U.S. Atlantic Coast. Transactions of the American Fisheries Society 136:1534-1545. https://doi.org/10.1577/T06-113.1

Heyman, W. D., & Kjerve, B. (2008). Characterization of transient multi-species reef fish spawning aggregations at Gladden Spit, Belize. *Bulletin of Marine Science*, *83*(3), 531–551.

Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., Harcourt, R. G., Holland, K. N., Iverson, S. J., Kocik, J. F., Flemming, J. E. M., & Whoriskey, F. G. (2015). Aquatic animal telemetry: A panoramic window into the underwater world. *Science 348*(6240), Article 1255642. 10.1126/science.1255642

Ingram, G. W., Jr. & Patterson, W. F., III. (2001). Movement patterns of red snapper (Lutjanus campechanus), greater amberjack (*Seriola Dumerili*), and gray triggerfish (Balistes capriscus) in the Gulf of Mexico and the utility of marine reserves as management tools. *Proceedings of the Gulf and Caribbean Fisheries Institute 52,* 686-699.

Jackson, L. S., Drymon, J. M., Nelson, T. R., & Powers, S. P. (2018). Biotelemetry based estimates of greater amberjack (*Seriola dumerili*) post-release mortality in the northern Gulf of Mexico. *Fisheries Research*, *208*, 239–246. https://doi.org/10.1016/j.fishres.2018.07.017

Jacoby, D. M. P., Brooks, E. J., Croft, D. P., & Sims, D. W. (2012). Developing a deeper understanding of animal movements and spatial dynamics through novel application of network analysis. *Methods in Ecology and Evolution 3*, 574-583. doi: 10.1111/j.2041-210X.2012.00187.x

Johnston, M. A., Nuttall, M. F., Hickerson, E. L., O'Connell, K., Blakeway, R. D., Embesi, J. A., MacMillan, J., Peter, D., & Schmal, G. P. (2022). Characterizing the biological community before and after partial removal of an offshore gas platform in the northwestern Gulf of Mexico. *Environmental Management*, *70*, 1078–1092. https://doi.org/10.1007/s00267-022-01714-8

Kerr, L. A., Cadrin, S. X., & Secor, D. H. (2010). The role of spatial dynamics in the stability, resilience, and productivity of an estuarine fish population. *Ecological Applications 20*(2), 497-597. https://doi.org/10.1890/08-1382.1

Kock, A., O'Riain, M. J., Mauff, K., Meÿer, M., Kotze, D., & Griffiths, C. (2013). Residency, habitat use, and sexual segregation of white sharks, *Carcharodon carcharias* in False Bay, South Africa. *PLoS ONE 8*(1), Article e55048. doi:10.1371/journal.pone.0055048

Jacoby, D. M. P. & Piper, A. T. (2023). What acoustic telemetry can and cannot tell us about fish biology. *Journal of Fish Biology*, *106*, 1260-1284. 10.1111/jfb.15588

Lédée, E. J. I., Heupel, M. R., Taylor, M. D., et al. (2021) Continental-scale acoustic telemetry and network analysis reveal new insights into stock structure. *Fish and Fisheries* 22, 987–1005. https://doi.org/10.1111/faf.12565

Lédée, E. J. I., Heupel, M. R., Tobin, A. J., & Simpfendorfer, C. A. (2015). Movements and space use of giant trevally in coral reef habitats and the importance of environmental drivers. *Animal Biotelemetry*, *3*, Article 6. DOI 10.1186/s40317-015-0024-0

Lovell, M. S., Polito, M. J., Schuster, J. A., Shallow, E. E., Janosik, A. M., Falterman, B. J., Dance, M. A. (2024). Seasonal variability in the feeding ecology of an oceanic predator. *Scientific Reports 14*, Article 17353. https://doi.org/10.1038/s41598-024-63557-z

Lowerre-Barbieri, S. K., Friess, C., Griffin, L. P., Morley, D., Skomal, G. B., Bickford, J. W., Hammerschlag, N., Rider, M. J., Smukall, M. J., van Zinnicq Bergmann, M. P. M., Guttridge, T. L., Kroetz, A. M., Grubbs, R. D., Gervasi, C. L., Rehage, J. S., Poulakis, G. R., Bassos-Hull, K., Gardiner, J. M., Casselberry, G. A., … Brownscombe, J. W. (2021). Movescapes and eco-evolutionary movement strategies in marine fish: Assessing a connectivity hotspot. *Fish and Fisheries 22*, 1321–1344. https://doi.org/10.1111/faf.12589Digital Object Identifier (DOI)

Lowerre-Barbieri, S. K., Wall, K., Friess, C., Keenan, S., Lembke, C., Tarnecki, J., Williams-Grove, L. J., & Patterson, W. F. (2025). Movement traits important to conservation and fisheries management: and example with red snapper. *Scientific Reports 15*, 4614. https://doi.org/10.1038/s41598-025-86892-1

Mammel, M., Lee, M.-A., Naimullah, M., Liao, C.-H., Wang, Y.-C., & Semedi, B. 2023. Habitat changes and catch rate variability for greater amberjack in the Taiwan Strait: the effects of El-Nino-southern oscillation events. Frontiers in Marine Science 10:1024669. https://doi.org/10.3389/fmars.2023.1024669

Mammel, M., Naimullah, M., Vaygan, A. H., Hsu, J., Lee, M.-A., Wu, J.-H., Chen, Y.-C., & Lan, K.-W. 2022. Variability in the spatiotemporal distribution patterns of greater amberjack in response to environmental factors in the Taiwan Strait using remote sensing data. Remote Sensing 14:2932. https://doi.org/10.3390/rs14122932

Matley, J. K., Klinard, N. V., Martins, A. P. B., Aarestrup, K., Aspillaga, E., Cooke, S. J., Cowley, P. D., Heupel, M. R., Lowe, C. G., Lowerre-Barbieri, S. K., Mitamura, H., Moore, J., Simpfendorfer, C. A., Stokesbury, M. J. W., Taylor, M. D., Thorstad, E. B., Vandergoot, C. S., & Fisk, A. T. (2022). Global trends in aquatic animal tracking with acoustic telemetry. *Trends in Ecology & Evolution 37*(1), 79-94. https://doi.org/10.1016/j.tree.2021.09.001

McClellan, D. B. & Cummings, N. J. 1997. Preliminary analysis of tag and recapture data of greater amberjack, *Seriola dumerili*, in the southeastern United States. Proceedings of the Gulf Coast Fisheries Institute 49: 25-45.

Murie, D. J., Parkyn, D. C., & Austin, J. 2013. Seasonal movement and mixing rates of greater amberjack in the Gulf of Mexico and assessment of exchange with the South Atlantic spawning stock. SEDAR33-DW12 1-46.

Ogburn, M. D., Harrison, A.-L., Whoriskey, F. G., Cooke, S. J., Mills Flemming, J. E., & Torres, L. G. (2017). Addressing challenges in the application of animal movement ecology to aquatic conservation and management. *Frontiers in Marine Science 4*(70). 10.3389/fmars.2017.00070

Ojeda, G. Y., Gayes, P. T., Dolah, R. F. V., & Schwab, W. C. (2004). Spatially quantitative seafloor habitat mapping: example from the northern South Carolina inner continental shelf. *Estuarine, Coastal, and Shelf Science 59*(3):399-416. https://doi.org/10.1016/j.ecss.2003.09.012

Qian, H., Li, Y., He, R., Eggleston, D. B. (2015). Connectivity in the Intra-American Seas and implications for potential larval transport. *Coral Reefs 34*, 403-417. DOI 10.1007/s00338-014-1244-0.

Paxton, A. B., Newton, E. A., Adler, A. M., Van Hoeck, R. V., Iverson Jr., E. S., Taylor, J. C., Peterson, C. H., & Silliman, B. R. (2020). Artificial habitats host elevated densities of large reef-associated predators. *PLoS ONE*, *15*(9), Article e0237374. https://doi.org/10.1371/journal. pone.0237374

Perkinson, M., Darden, T., Jamison, M., Walker, M. J., Denson, M. R., Franks, J., Hendon, R., Musick, S., Orbensen, E. S. (2019). Evaluation of the Stock Structure of Cobia (Rachycentron canadum) In the Southeastern United States By Using Dart-Tag and Genetics Data. *Fishery Bulletin, 117*(3), 220-233. doi:10.7755/FB.117.3.9

Piraino, M. N. & Szedlmayer, S. T. (2014). Fine-scale movements and home ranges of red snapper around artificial reefs in the northern Gulf of Mexico. *Transactions of the American Fisheries Society, 143(*4): 988-998. https://doi.org/10.1080/00028487.2014.901249

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rooker, J. R., Dance, M. A., Wells, R. J. D., Quigg, A., Hill, R. L., Appeldoorn, R. S., Ferreira, B. P., Boswell, K. M., Sanchez, P. J., Moulton, D. L., Kitchens, L. L., Rookers, G. J., & Aschenbrenner, A. (2018).

Seascape connectivity and the influence of predation risk on the movement of fishes inhabiting a back-reef ecosystem. *Ecosphere 9*(4), Article e02200. https://doi.org/10.1002/ecs2.2200

Rooker, J. R., Wells, R. J. D., Itano, D. G., Thorrold, S. R., & Lee, J. M. (2016). Natal origin and population connectivity of bigeye and yellowfin tuna in the Pacific Ocean. *Fisheries Oceanography 25*(3), 277-291. doi:10.1111/fog.12154

Sanchez, P. J., Dance, M. A., Kraus, R. T., Hill, R. L., Rooker, J. R. (2023). Fish community characterization of mid-shelf and shelf-edge mesophotic coral ecosystems in the expanded Flower Gardens Banks National Marine Sanctuary. *Bulletin of Marine Science 99*(1), 41-49. https://doi.org/10.5343/bms.2022.0014

Schiavon, L., Desvignes, T., Ronco, F., Matschiner, M., Gastaldi, M., Koppetsch, T., Schrøder-Nielsen, A., Postlethwait, J. H., Prestanti, A., Stranci, F., Ceballos, S. G., Mark, F. C., Lucassen, M., Riginella, E., La Mesa, M., & Papetti, C. (2025). Sex differentiation and long-distance gene flow in the elusive Antarctic fish *Aethotaxis mitopteryx*. *Ecology and Evolution 15*, Article e71847. https://doi.org/10.1002/ece3.71847

Schultz, M. K., Dance, M. A., Powers, S., & Midway, S. R. (2025). Global review of Greater Amberjack (*Seriola dumerili*) life history. *Reviews in Fisheries Science & Aquaculture*, 1-22. https://doi.org/10.1080/23308249.2025.2595023

Schulze, A., Erdner, D. L., Grimes, C. J., Holstein, D. M., & Miglietta, M. P. (2020). Artificial reefs in the northern Gulf of Mexico: Community ecology amid the "Ocean Sprawl." *Frontiers in Marine Science 7*: Article 447. https://doi.org/10.3389/fmars.2020.00447

Secor, D. H., O'Brien, M. H. P., Gahagan, B. I., Watterson, J. C., Fox, D. A. 2020. Differential migration in Chesapeake Bay striped bass. PLoS ONE 15(5):e0233103. https://doi.org/10.1371/journal.pone.0233103

SEDAR, 2020. SEDAR 70 – Gulf of Mexico Greater Amberjack (Southeast Data, Assessment, and Review). SEDAR, North Charleston, SC.

Tone, K., Chiang, W. C., Yeh, H. M., Hsiao, S. T., Li, C. H., Komeyama, K., Kudo, K., Hasegawa, T., Sakamoto, T., Nakamura, I., Sakakura, Y., & Kawabe, R. (2022). Two-way habitat use between reefs and open ocean in adult greater amberjack: Evidence from biologging data. *Marine Ecology Progress Series*, *699*, 135–151. https://doi.org/10.3354/meps

Tone, K., Nakamura, Y., Chiang, W. C., Yeh, H. M., Hsiao, S. T., Li, C. H., Komeyama, K., Tomisaki, M., Hasegawa, T., Sakamoto, T., Nakamura, I., Sakakura, Y., & Kawabe, R. 2021. Migration and spawning behavior of the greater amberjack *Seriola dumerili* in eastern Taiwan. Fisheries Oceanography 1:1-18. https://doi.org/10.1111/fog.12559Digital Object Identifier (DOI)

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer.

Tables

Table 6.1: Summary of the number of fish tagged, detected, along with their sizes and number of detections presented by state and region. Detections from Alabama/Mississippi are given as number of positions generated from acoustic receiver VPS arrays, rather than raw detection counts.

Region/State	*Tag #*	*Depth Tags*	*Tags Detected*	*Detections*	*Days Detected*	*Mean Days Detected*	*Mean Detection Duration (days)*	*Mean Fork Length (mm)*	*Minimum Fork Length (mm)*	*Maximum Fork Length (mm)*
Atlantic	**149**	**18**	**82**	**1632780**	**4385**	**53.2**	**125**	**827.6**	**500**	**1300**
North Carolina	79	7	43	1115857	2444	58.8	136.8	839.3	620	1140
South Carolina	25	4	19	471482	1188	62.5	141.6	800.1	511	1030
East Florida	45	7	20	70370	861	33.7	83.8	822.6	500	1300
Eastern Gulf	**123**	**23**	**98**	**1775303**	**6306**	**64.3**	**92.3**	**834.3**	**430**	**1189**
West Florida	42	2	39	1405090	4036	103.5	172.6	871.5	600	1189
Alabama/Mississippi	81	21	59	370213	2270	38.5	39.3	814.6	430	1025
Western Gulf	**226**	**40**	**139**	**3711163**	**22594**	**162.5**	**278.9**	**1080.5**	**464**	**1525**
Louisiana	151	40	114	2828781	16957	148.8	268.7	1088.1	571	1525
Texas	75	0	25	882382	5637	225.3	325.6	1065.3	464	1435
Total	**498**	**71**	**319**	**7119246**	**33285**	**104.4**	**182**	**944.7**	**430**	**1525**

Figures

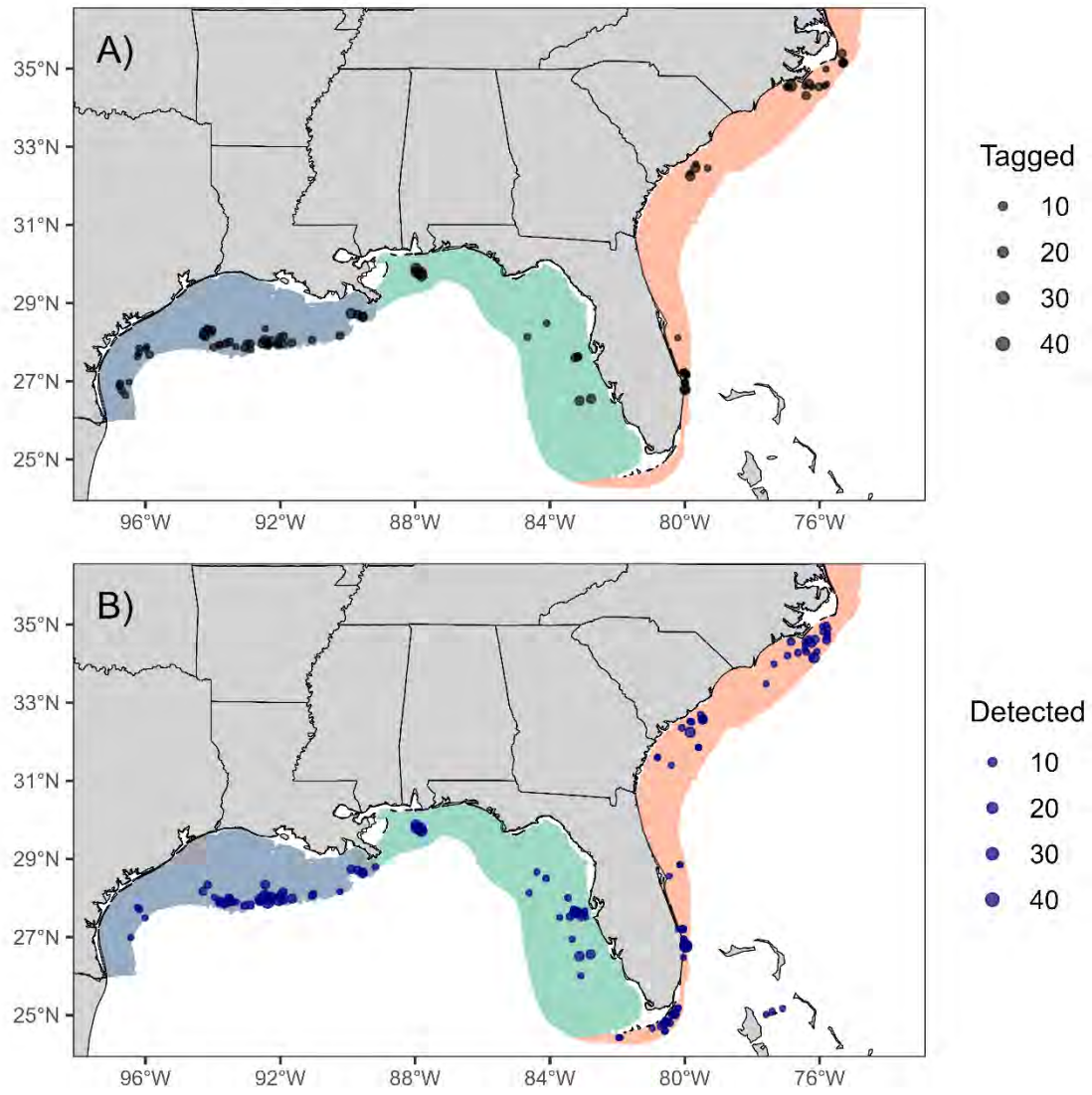


Figure 6.1: Map of tagging (Figure 6.1A) and detection (Figure 6.1B) locations, with points sized according to the number of fish tagged or detected at that location. Project regions are shown by colored polygons, with the western Gulf in blue, eastern Gulf in green, and the Atlantic in peach.

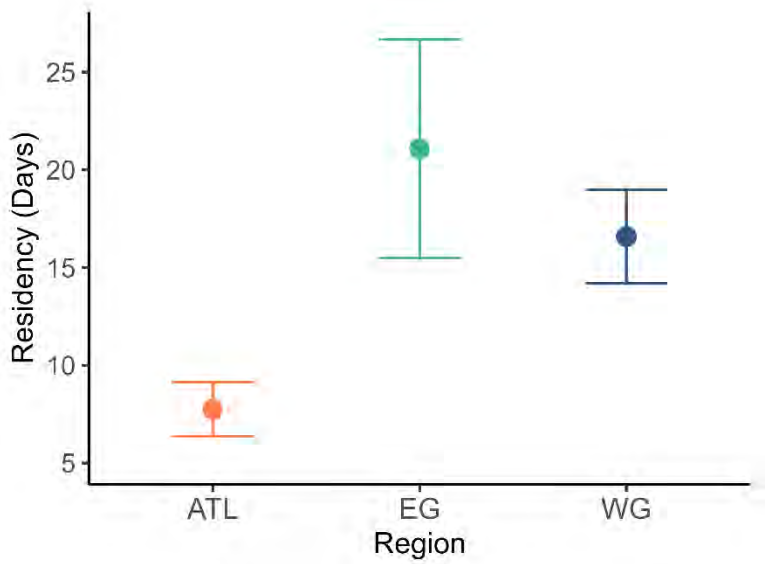


Figure 6.2: Mean continuous detection period at a given site for acoustically tagged Greater Amberjack by region. Error bars indicate associated 95% confidence intervals. Abbreviations: ATL = Atlantic, EG = Eastern Gulf, WG = Western Gulf.

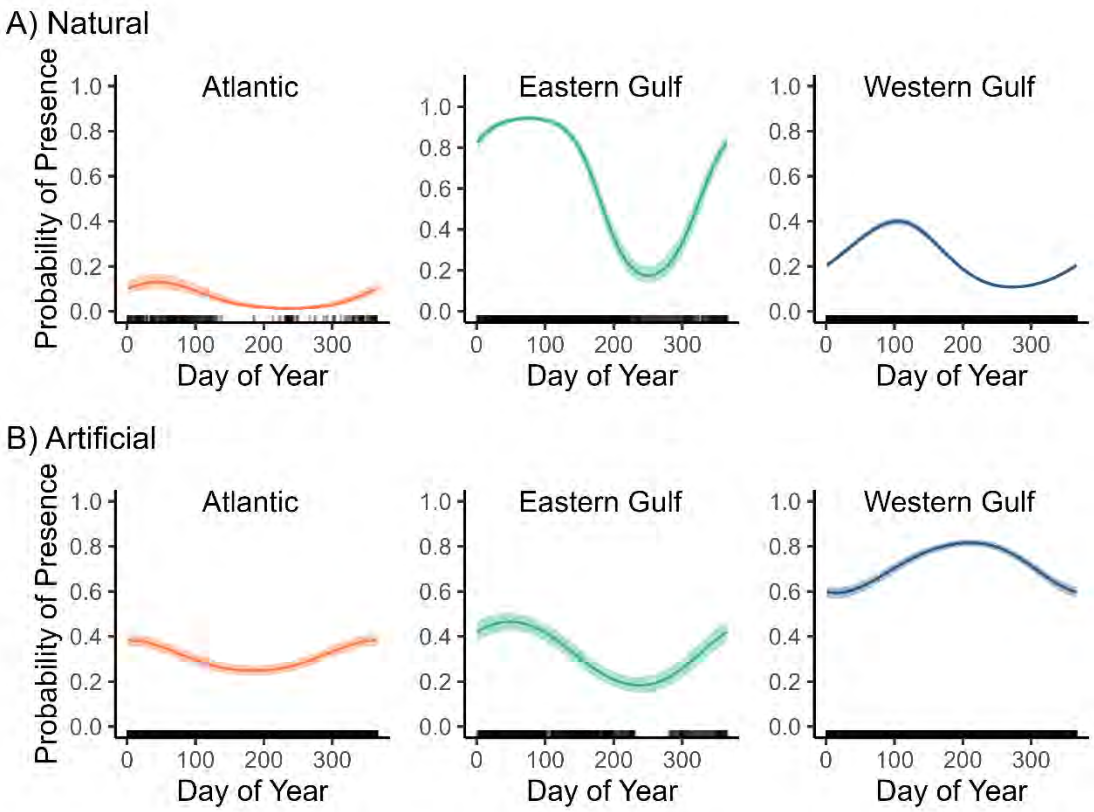


Figure 6.3: Smooth estimates from generalized additive mixed models showing the likelihood of detecting a tagged greater amberjack on either natural (row A) or artificial (row B) habitat throughout the year. Shaded areas are 95% confidence intervals.

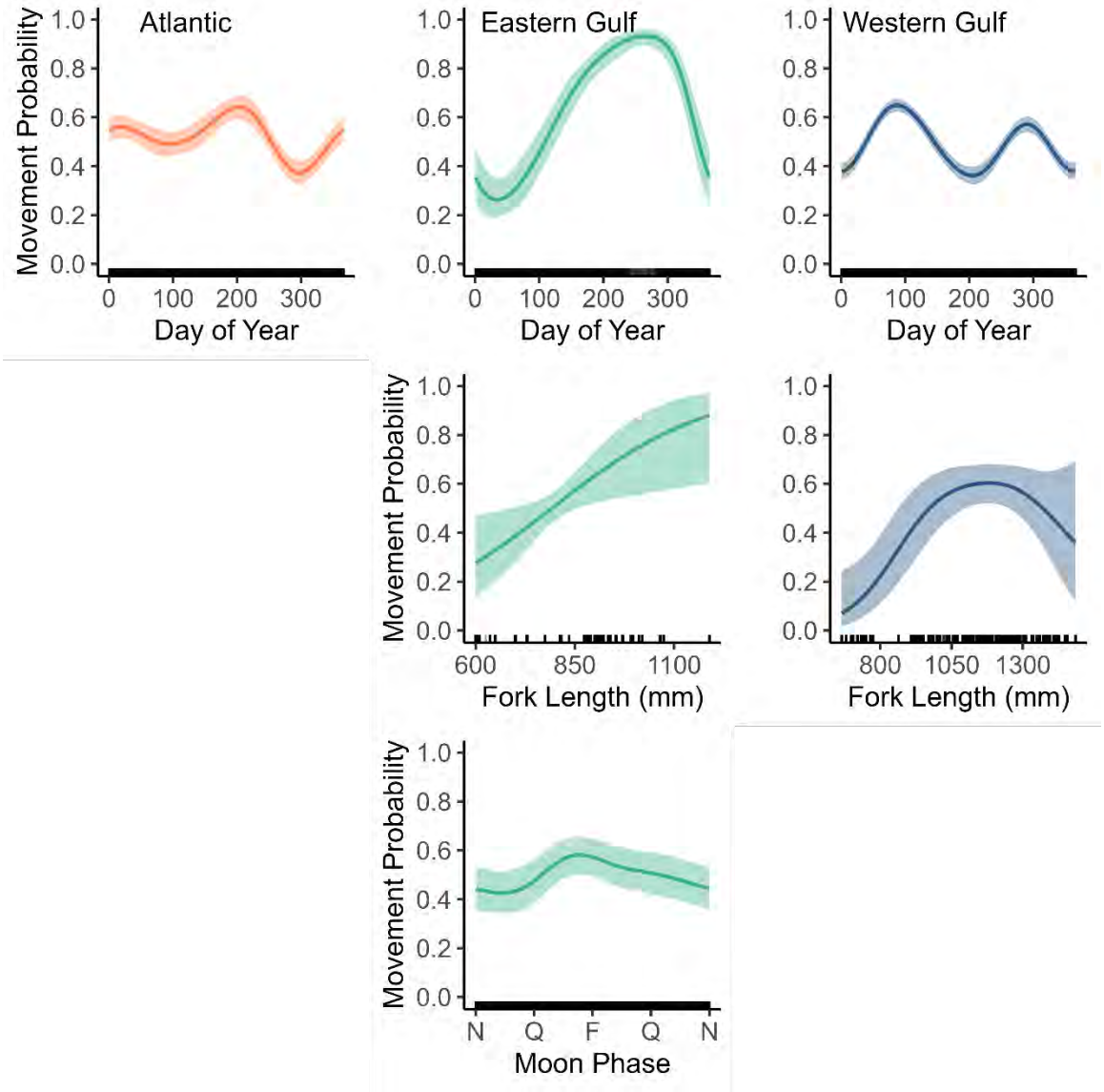


Figure 6.4: Smooth estimates from region-specific generalized additive mixed models, showing the influence of significant (⟨ = 0.05) predictor variables on the probability of amberjack movement. Shaded areas are 95% confidence intervals.

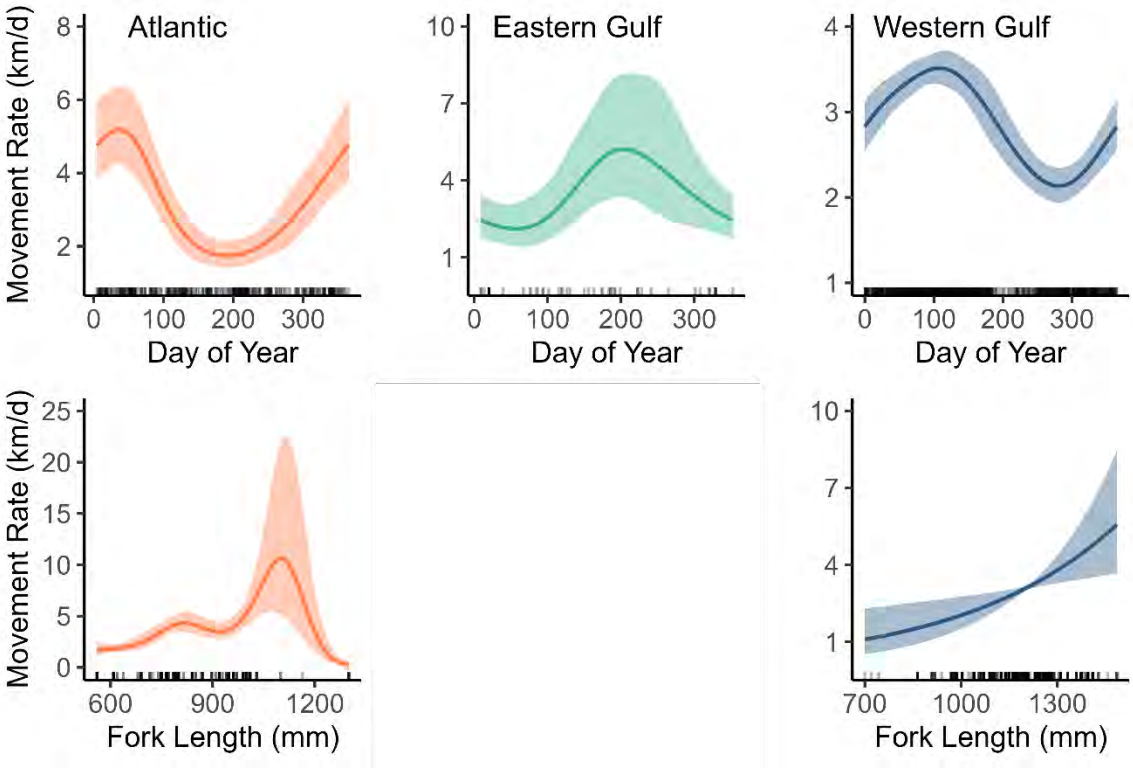


Figure 6.5: Smooth estimates from generalized additive mixed modelling. Each model shows the influence of significant (⟨ = 0.05) predictor variables on the rate of movement by Greater Amberjack. Shaded areas are 95% confidence intervals.

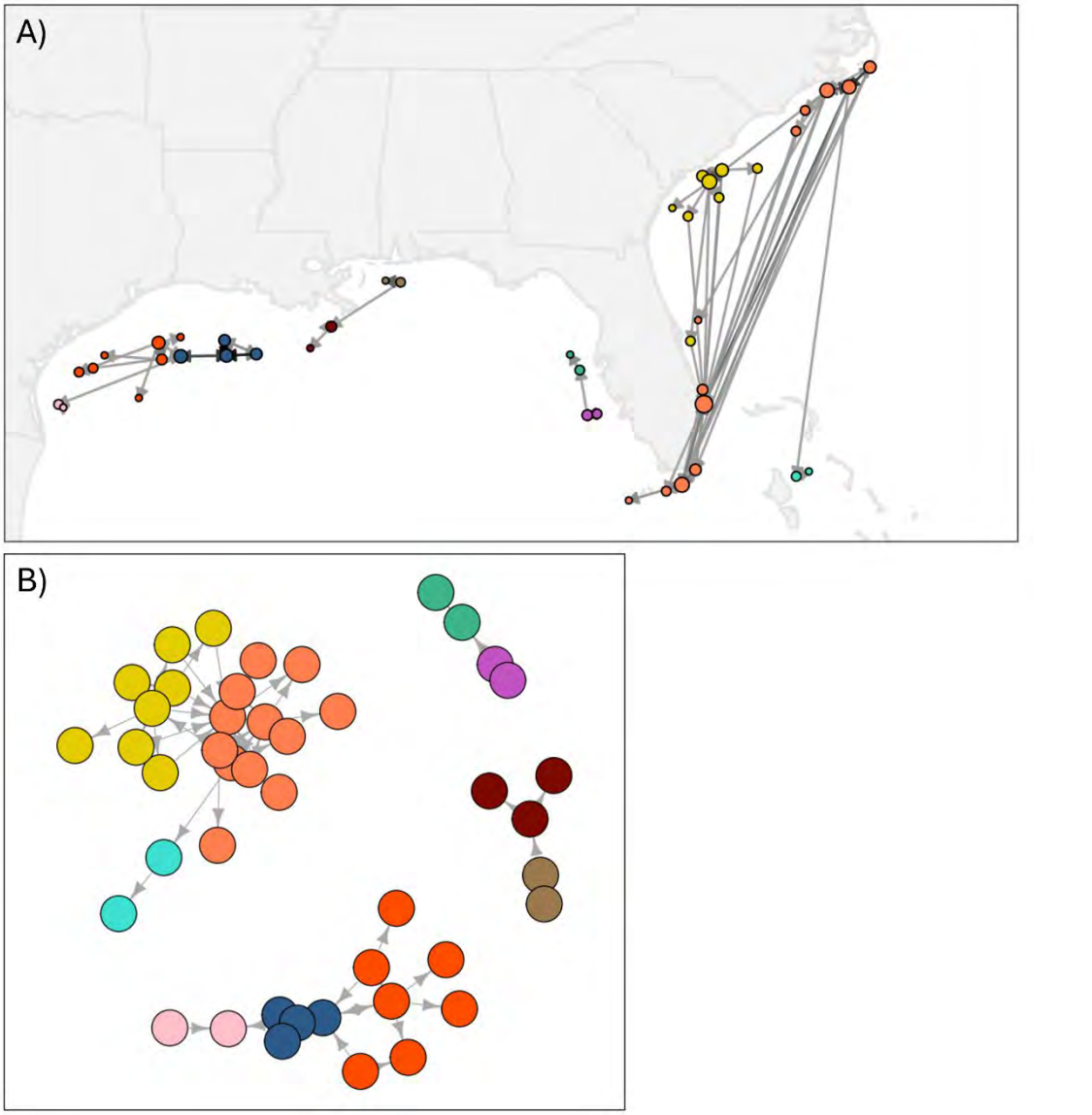


Figure 6.6: A) Network based on 100 km gridded nodes, plotting with spatial reference the respective nodes and edges including all movements of tagged Greater Amberjack. Degree centrality of each node is indicated by point size, edge weight is depicted by line color. Node colors are indicative of distinct communities identified in the cluster analysis for each network. B) The same network, plotted without spatial reference using Fruchterman-Reingold algorithm.

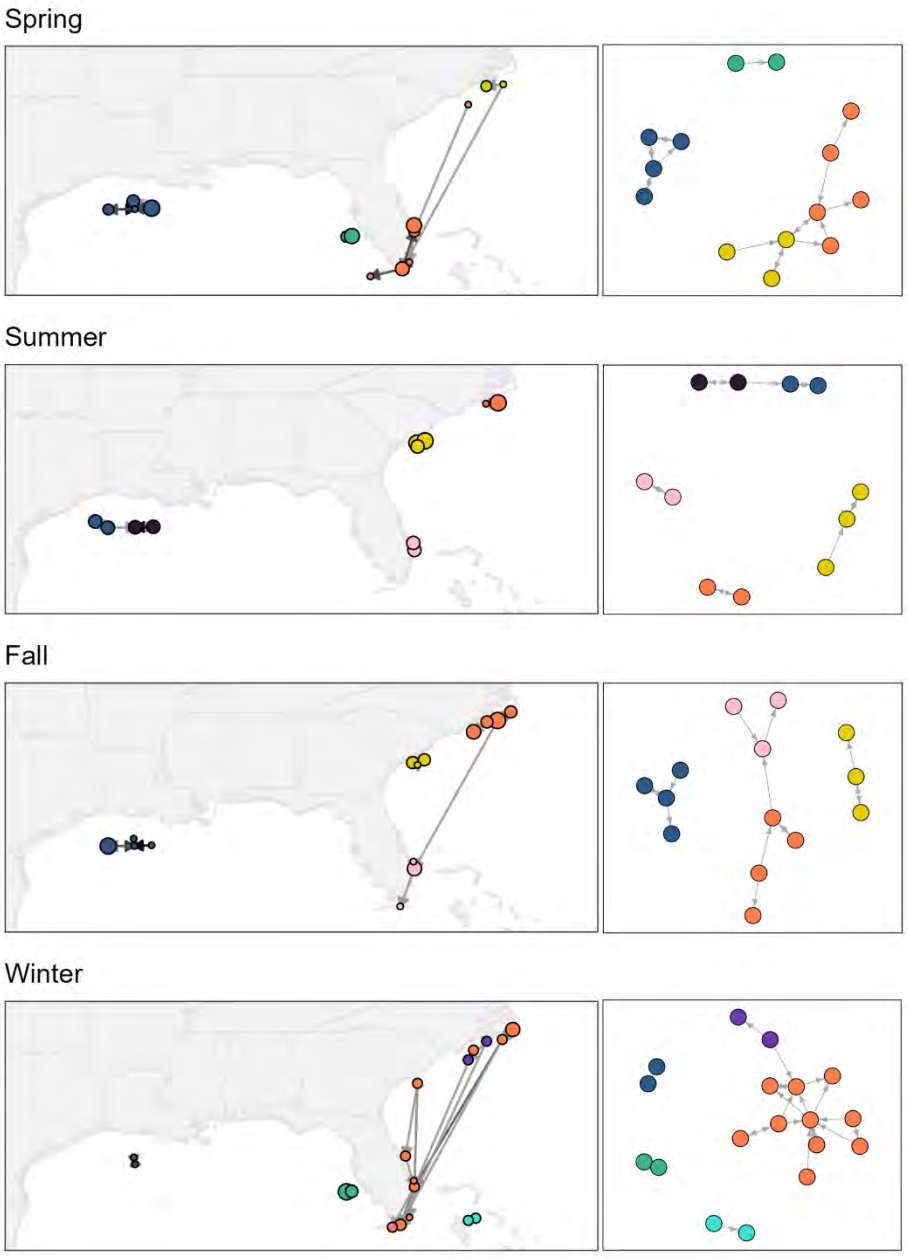
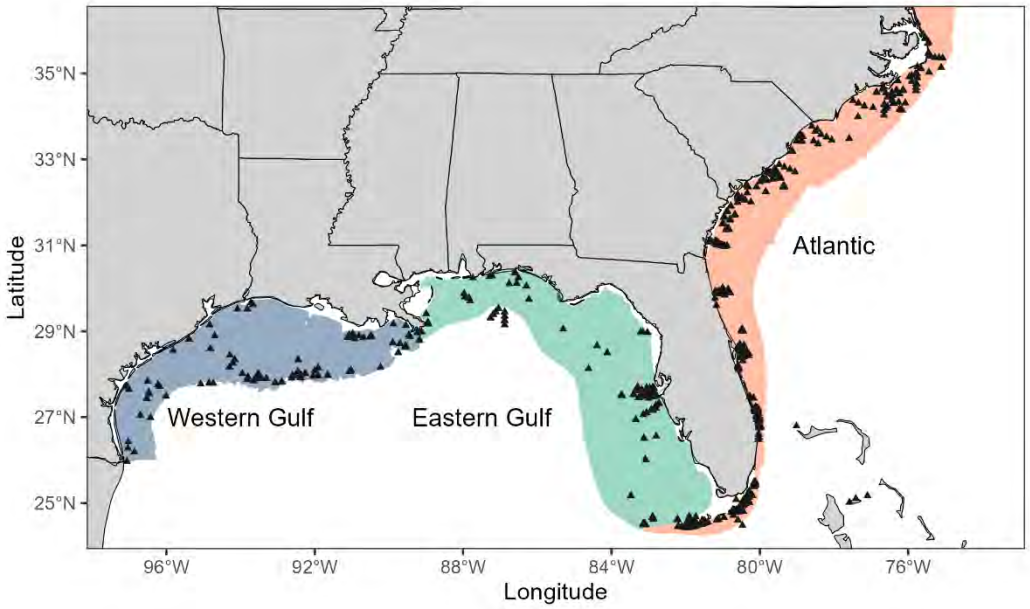
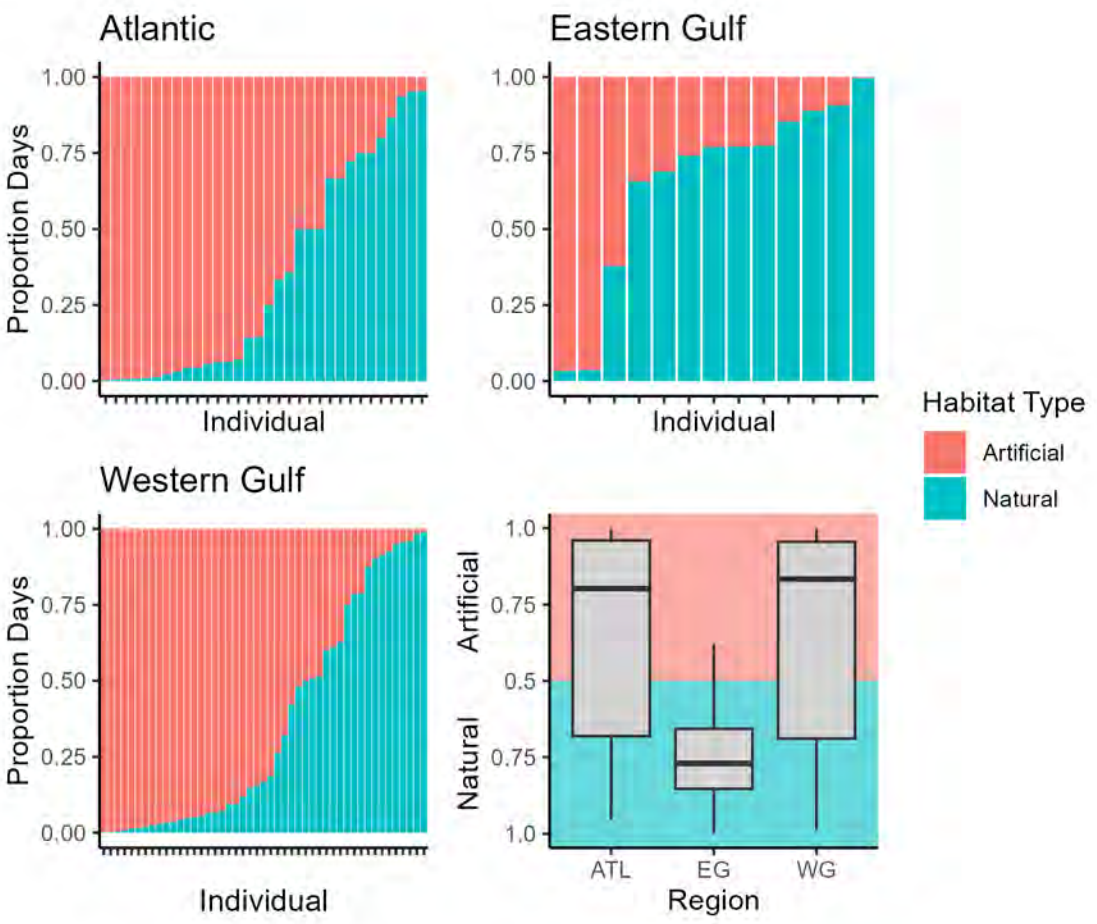


Figure 6.7: Seasonal networks based on 100 km gridded nodes, plotted with the respective nodes and edges. Degree centrality of each node is indicated by point size, edge weight is depicted by line color. Node colors are indicative of distinct communities identified in the cluster analysis for each network. Left panel plots are plotted with spatial reference, right panel plots are shown without spatial reference using the Fruchterman-Reingold algorithm.

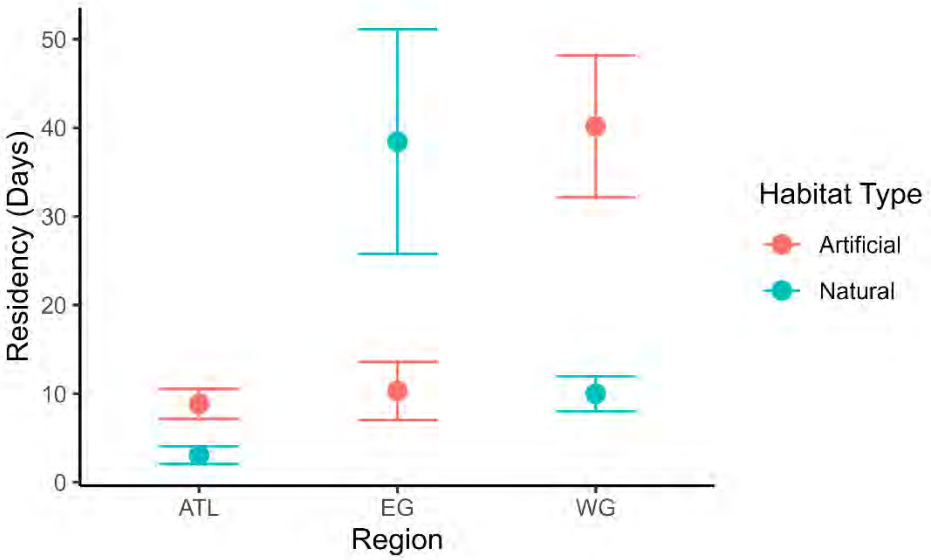
Supplemental Figures



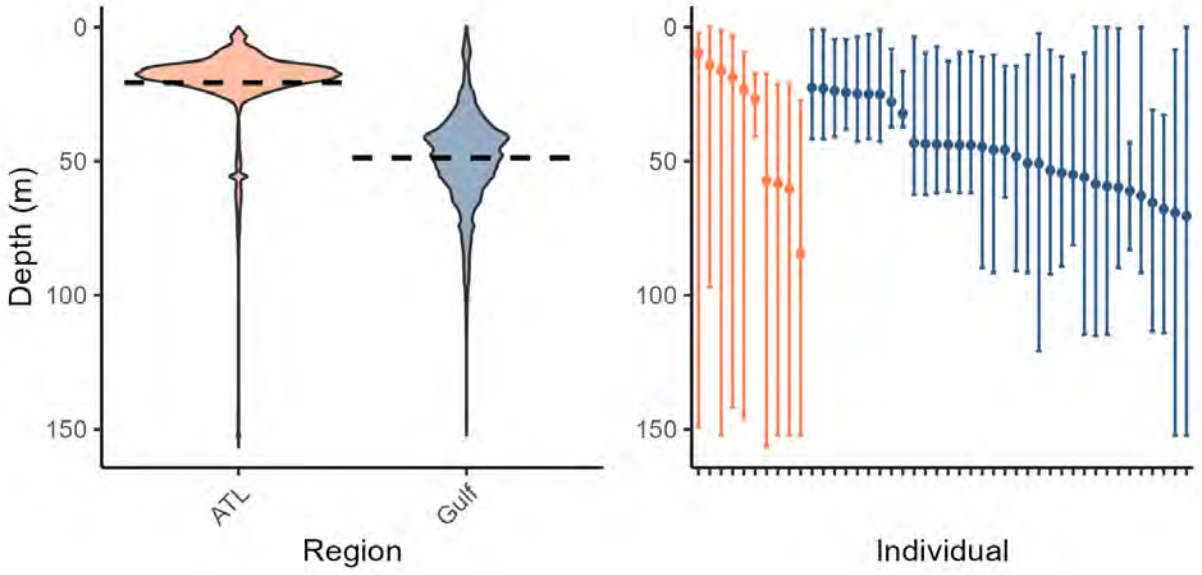
Supplemental Figure 6.1: Map of acoustic receivers deployed throughout the study region, both as part of the current project and those deployed by outside researchers in the FACT and iTAG networks. Project regions are shown by colored polygons, with the western Gulf in blue, eastern Gulf in green, and the Atlantic in peach.



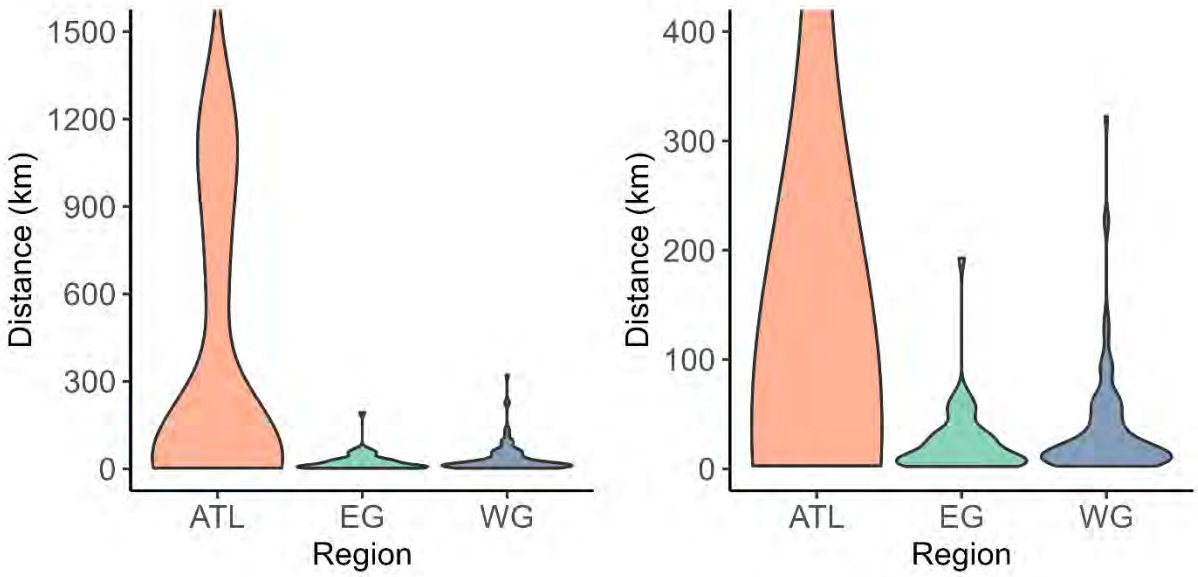
Supplemental Figure 6.2: Stacked bar plot of all individuals that were detected on both natural and artificial habitats. Each bar represents an individual, showing the proportion of time that fish spent in each habitat type. Bottom right: Box and whisker plots depicting pooled habitat use by region, the horizontal line represents the median ratio of time spent between natural and artificial habitats, the box indicates the interquartile range (IQR) and the whiskers values 1.5 times the IQR.



Supplemental Figure 6.4: Mean continuous detection period at a given site for acoustically tagged Greater Amberjack by region and habitat type. Error bars indicate associated 95% confidence intervals. Abbreviations: ATL = Atlantic, EG = Eastern Gulf, WG = Western Gulf.



Supplemental Figure 6.5: Depth use by greater amberjack. Left Panel: Violin plots showing the distribution of depth occupied by greater amberjack in the Atlantic and pooled Gulf regions. Dashed line indicates overall mean depth. Right Panel: Individual depth use by greater amberjack, fish tagged in the Atlantic are shown in peach, those tagged in the Gulf are shown in blue. Points indicate that individual's mean depth, bars show maximum and minimum depths detected.



Supplemental Figure 6.6: Violin plot showing the frequency of displacements observed by Greater Amberjack. The left panel shows all displacements (except one in the Atlantic, 2433.0 km; excluded for readability), while the right is adjusted to show relationships where density was highest. Abbreviations: ATL = Atlantic, EG = Eastern Gulf, WG = Western Gulf.

Supplemental Table: Results from Tukey's Honest Significant Differences Test comparing residency period between regions and habitat types. Pairwise comparisons denoted with an asterisk indicate significant differences (⟨ = 0.05).

Pairwise Comparison	Difference	95 % CI	Adjusted p
Eastern Gulf-Atlantic*	13.32	8.27	0.00048
Western Gulf - Atlantic*	8.83	4.62	0.00002
Western Gulf- Eastern Gulf	-4.49	7.66	0.35485
Atlantic: Natural - Artificial	-5.79	12.28	0.75927
Eastern Gulf: Natural - Artificial*	28.14	18.19	0.00015
Western Gulf: Natural - Artificial*	-30.18	7.09	<0.00001
Natural - Artificial*	-13.36	3.38	<0.00001

Supplemental Table: Summary of sites where greater amberjack depth tags were detected.

Region	Habitat Type	Latitude	Longitude	Tags Detected	Detections	Days of Detection	Bottom Depth (m)	Mean Depth ± SD (m)	Dist Off Bottom (m)
ATL	Artificial	26.78	-80.00	1	3	1	29	24.3 ± 0.5	4.7
ATL	Artificial	34.28	-76.64	1	7996	24	28	19.6 ± 5.1	8.4
ATL	Artificial	26.48	-80.04	1	2	1	39	37.0 ± 1.7	2.0
ATL	Artificial	32.59	-79.48	1	35	2	21	16.6 ± 2.1	4.4
ATL	Artificial	32.59	-79.45	1	5	1	21	18.2 ± 4.8	2.8
ATL	Artificial	32.56	-79.45	1	10	1	21	17.8 ± 1.2	3.2
ATL	Artificial	32.56	-79.48	1	104	4	21	17.7 ± 2.1	3.3
ATL	Artificial	26.79	-80.00	3	98	10	32	25.1 ± 4.5	6.9
ATL	Artificial	26.77	-79.97	8	292	35	150	142.9 ± 12.4	7.1
ATL	Artificial	26.77	-79.97	8	348	39	148	143.5 ± 10.3	4.5
ATL	Artificial	27.20	-80.05	1	465	11	Unknown	26.4 ± 3.3	-
ATL	Artificial	32.25	-79.84	3	86431	250	23	17.8 ± 4.0	5.2
ATL	Artificial	34.56	-76.85	2	10029	39	18	8.8 ± 4.5	9.2
ATL	Artificial	26.79	-80.01	1	2	1	Unknown	28.2 ± 2.6	-
ATL	Artificial	26.79	-80.01	2	31	2	Unknown	28.2 ± 3.6	-
ATL	Natural	26.79	-80.02	1	1	1	Unknown	20.0	-
ATL	Artificial	26.79	-80.00	2	37	6	29	22.4 ± 3.7	6.6
ATL	Natural	27.20	-80.05	1	240	11	Unknown	27.6 ± 3.0	-
ATL	Natural	27.20	-80.06	1	3	2	Unknown	28.2 ± 2.4	-
ATL	Artificial	26.75	-80.01	2	2	2	27	26.4 ± 1.3	0.6
ATL	Unknown	26.79	-79.99	4	8006	121	85	58.5 ± 11.2	26.5
EG	Artificial	29.77	-87.83	9	50304	37	42	24.6 ± 7.9	17.4
WG	Artificial	28.73	-89.90	7	625155	251	56	43.9 ± 7.3	12.1
WG	Natural	28.63	-89.55	9	142359	488	64	58.1 ± 8.7	5.9
WG	Artificial	28.79	-89.18	2	34568	1	150	72.9 ± 24.6	77.1
WG	Artificial	28.71	-89.70	6	109131	259	86	50.8 ± 12.4	35.2
WG	Artificial	28.66	-89.55	8	233120	567	107	56.7 ± 27.1	50.3

7. Age and Growth

Variability in growth of greater amberjack (*Seriola dumerili*) in the U.S. Gulf and Atlantic

Mischa K. Schultz[1], Michael A. Dance[1], Justin McDonald[2], Ashley Pacicco[3], Robert Allman[3], Andy Ostrowski[4], Creed Branham[1], Brett Falterman[5], Mark Albins[2], Sean Powers[2] and Stephen R. Midway[1]

[1]Department of Oceanography and Coastal Science, Louisiana State University, Baton Rouge, LA. Mschu31@lsu.edu

[2]Department of Marine Sciences, University of South Alabama, Mobile, AL.

[3]Southeast Fisheries Science Center, NOAA Fisheries, 3500 Delwood Beach Road, Panama City, FL 32408

[4]NOAA Fisheries, U.S. Department of Commerce, 101 Pivers Island Rd. Beaufort, NC 28516

[5]Fisheries Research Support LLC, Mandeville, LA 70448

INTRODUCTION

Somatic growth, or an increase in body size over time, is an important process that influences fecundity, natural mortality, and biomass production (Haddon, 2011; Lorenzen, 1996; Punt & Hilborn, 1997). Length-at-age data derived from growth models inform estimates of age structure and associated age-related reproductive output and mortality (Punt & Hilborn, 1997). Therefore, understanding growth dynamics is critical for determining stock status and fishery-associated population impacts. Growth varies widely among individuals within fish populations, and this demographic plasticity makes growth difficult to characterize across populations. At the individual level, body size affects maturation schedules (Stearns & Koella, 1986), and natural mortality (e.g., Gislason et al., 2010; Pauly, 1980; Then et al., 2014), linking growth directly to fitness and life-history strategy. Growth rate is often strongly correlated to natural mortality: smaller, faster-growing individuals experience higher mortality, whereas larger, slower-growing individuals experience lower mortality risk (Gislason et al., 2010; Houde, 2002; Pauly, 1980). Consequently, variability in growth can influence population productivity and resilience. Despite increasing evidence of variability in somatic growth (Gertseva et al., 2017; Midway et al., 2015), stock assessment models commonly assume homogenous demographic processes (Cadrin, 2020). Models that assume uniform growth across space or sex risk systematically misrepresent population structure and productivity when biologically meaningful differences exist (Lorenzen, 2016; Su et al., 2011; Wang et al., 2005). Because reference points derived from length-at-age data are used to adjust harvest levels and define population structure (Hilborn & Walters, 1992), understanding variation in somatic growth and mortality within fish populations is essential for effective management.

Growth rates in fish vary across a species' geographic range or between sexes, and can be attributed to variation in temperature, food availability, genetics, and density dependence (Croll et al., 2023; Erisman et al., 2021; Kotowych et al., 2023; Lorenzen, 2016). For example, fish growth rates are positively correlated with temperature (Gislason et al., 2010), an association demonstrated by changes in growth across geographical gradients (Adams et al., 2018; Denit & Sponaugle, 2004; Williams et al., 2012). Spatial heterogeneity in the growth of marine fishes has been documented across reef patches, latitudinal gradients, and regional habitats, likely driven by the complex interaction of these environmental factors (Adams et al., 2018; Allman et al., 2018; Gertseva et al., 2017). In addition to spatial variation, differences in growth between sexes can shape population productivity. Sexual size dimorphism (SSD) is common among fishes and reflects trade-offs between fecundity and reproductive strategy (Parker, 1992). Female-biased SSD, in which females attain larger body sizes than

males, is observed in over 80% of fish species (Pauly, 2019) and is recognized in species assessed using modern stock assessment models (Banks et al., 2024; Pacicco et al., 2021). Natural mortality and vulnerability to fishing will differ between males and females of species that have sexually dimorphic growth rates (e.g., Maunder & Wong, 2011). To make management decisions, it is important to capture this variation in the management process. While statistical approaches that account for variability within and among groups have improved our understanding of population structure in other marine fish, many species remain understudied, and current growth estimates may be insufficient to fully capture the spatial or sex-specific patterns that influence population dynamics.

Greater amberjack is a federally managed species in the U.S. that supports recreational fisheries in both the U.S. Atlantic (Atlantic) and Gulf of America (Gulf; SEDAR, 2020a; SEDAR, 2020b). Following increased commercial harvest in the 1980s, both stocks were overfished by the 1990s. Management regulations implemented in the 1990s have included decreasing the daily bag limit to one fish, increasing minimum size limits, and enacting seasonal closures (SEDAR, 2014). As a result, the Atlantic stock has recovered, as the most recent stock assessment indicates it is neither overfished nor undergoing overfishing (SEDAR, 2020a). Unfortunately, despite regulatory efforts, the most recent stock assessment conducted for the Gulf stock confirms that it remains overfished and subject to overfishing (SEDAR, 2020b). The lack of recovery in the Gulf greater amberjack stock raises questions about stock productivity and the biological factors influencing the difference in recovery between the stocks. Growth estimates for greater amberjack remain uncertain because earlier studies relied primarily on fisheries-dependent sampling, lacking fish below fishery minimum size limits, and did not examine spatial variation in growth among subpopulations (Harris et al., 2007; Thompson et al., 1999). As a result, current estimates may not accurately reflect variability in growth within the defined unit stocks. Because natural mortality is commonly parameterized as a function of growth rate (Pauly, 1980; Then et al., 2014), uncertainty or bias in growth estimates can propagate directly into natural mortality and influence perceptions of stock productivity. Estimating natural mortality alongside updated growth parameters is therefore critical to improving assessment models and understanding differences in recovery potential between the Gulf and Atlantic stocks.

Here, we estimated ages of greater amberjack from annual growth increments in sagittal otoliths to estimate subregional, sex-, and stock-specific growth parameters in the Gulf and Atlantic. Because growth in greater amberjack has traditionally been described using von Bertalanffy growth function (VBGF) parameters, we adopted the same approach to maintain comparability. Given that other species with similar life histories exhibit within-stock variation in growth (e.g., Banks et al., 2024; DeVries & Grimes, 1997; Saari et al., 2014) and that greater amberjack may exhibit sexual dimorphism in growth (Harris et al., 2007; Murie & Parkyn, 2008), our primary objective was to develop age-growth relationships that contrast subregional growth variation and assess sex-specific differences in growth. To minimize confounding effects, our models incorporated both sex and subregion simultaneously. Growth parameters derived from the models were then used to empirically estimate natural mortality for greater amberjack for both sex and subregion.

METHODS

Study site and samples

Greater amberjack sagittal otoliths were sampled from 2014 – 2024 in the Gulf (n = 2,101) and the Atlantic (n = 724) from fishery-dependent and fisheries-independent sources. Samples were collected from recreational fishery port samplers throughout the Gulf, as well as provided by the National Marine Fisheries Service (NMFS) Panama City and Beaufort Laboratories. Samples were collected from Corpus Christi, TX to Collier County, FL

with most samples collected off the coast of Louisiana. Samples were further divided into three subregions (western Gulf, eastern Gulf, and Atlantic) to assess subregional differences in growth parameters. Samples from 2014 – 2018 were previously used in the SEDAR assessment. The three subregions reflected current stock boundaries between the Gulf and Atlantic at the Florida Keys (Gulf vs. Atlantic; Hargrove et al., 2018), and a split between eastern and western Gulf at the Mississippi River based on divergence in demographics for other species (DeVries & Grimes, 1997; Saari et al., 2014). Samples were provided as records of aged fish, otolith samples or prepared otolith slides. Samples were collected via hook and line during ongoing fishery-independent surveys. Fishery-dependent samples were obtained during dockside intercepts of commercial and charter vessels. In the Atlantic, fishery-dependent samples were collected by port samplers from North Carolina to south Florida. Across both regions, fishery-dependent samples were collected via rod and reel (electric, manual, hydraulic), longline, handline, and speargun, while fisheries-independent samples were collected via rod and reel (electric, manual, hydraulic), short bottom longline, and vertical longline. This multifaceted sampling method allowed for collection of samples in both stocks that were evenly distributed across size classes. Data collected with each fish included date of capture, capture location, type of fishery and body metrics such as total length (mm, TL), fork length (mm, FL), whole weight (g), and sex, whenever possible. Sex and weight were not recorded during dockside sampling when fish were brought back to the dock "gutted", without gonads or other internal organs. Fork length was used for all subsequent analyses.

Otolith preparation

Sagittal otoliths were rinsed with distilled water, stored dry, and then processed for aging following the methodology of Murie and Parkyn (2008). Whole otoliths were weighed to the nearest 0.001 g, and the left otolith was embedded in EpoThin EPOXY Resin. Right otoliths were used if the left was unavailable. Otoliths were sectioned through the core into three 0.5 mm sections that were then mounted to a glass slide with crystal bond. A liquid cover slip was applied over the sections to improve visibility of annuli.

Age determination

Otolith sections were viewed using a combination of compound or zoom stereomicroscope. An annulus was identified under transmitted light as an opaque zone followed by a translucent zone. Annuli were counted from the core along the primary and secondary axes. Additionally, edge types were quantified as 1, opaque zone at edge; 2, translucent space < 1/3 of the previous increment; 3, translucent space > 1/3 and < 2/3 of the previous increment; and 4, translucent space > 2/3 of the previous increment (Murie & Parkyn, 2008). All fish were assigned an age class based on number of annuli, edge type, and harvest date. Otoliths were read by two independent readers, with each otolith section read twice by each reader without knowledge of length, sex, or date of harvest of any specimen. A readability code was given to each otolith from A – F, where F is an unreadable otolith. Otoliths with differing ages between readers were read again by both readers simultaneously to assign an agreed upon age. Otoliths that were unreadable or an age could not be agreed upon were discarded from further analyses. Fish measuring less than 300 mm FL were assigned to age 0 without reading, based on capture date and biological birth date of April 1st, as well as by the length-at-age of age 0 fish (Wells & Rooker, 2004). To assign fish to year-class, a calendar age was calculated. Annuli counts were advanced by 1 year if an edge type of a 3 or 4 was assigned to a greater amberjack harvested from January 1 to June 30. Calendar age was equal to the annuli counts if the edge type was assigned as 1 or 2. If a fish was collected between July 1 and December 31, the calendar age equaled the number of annuli. Analysis of variance was used to determine whether there were significant differences in length at age between the western Gulf, eastern Gulf, and Atlantic for the most common age classes.

Growth models

The von Bertalanffy growth model (VBGM) was fit to length-at-age data for greater amberjack using Bayesian methods. Growth model parameters were estimated in JAGS (Plummer, 2003) via the rjags R package (v4-17; Plummer et al., 2025). To evaluate the influence of growth model structure on parameter estimates, we fit eight candidate models that varied in two dimensions: (i) the treatment of early growth (length-at-birth, L0, or theoretical age-at-zero-length, t_o), and (ii) the functional form of the growth curve. This sensitivity framework builds on the multi-model approach of Smart and Grammer (2021), who demonstrated that reparametrizing the VBGM in terms of L0 (length-at-birth) rather than t_o produces a more biologically interpretable parameter with a direct link to observable species traits. The candidate models were: (M1) conventional VBGM with t0 fixed at −1.2 for all subregions, based on initial estimation from a pooled VBGM. This model represents the most parameterization-constrained baseline; (M2) conventional VBGM with t0 estimated independently for each subregion; (M3) an L0 parameterization with L0 estimated independently in each subregion; (M4) a hierarchical L0 parameterization in which subregion-specific L0 values were drawn from a shared distribution, allowing partial pooling across subregions; (M5) a fixed L0 parameterization in which L0 was constrained to 150 mm — the smallest length observed in the data — for all subregions, and (M6) VBGM with L0 fixed at the published length-at-hatch for greater amberjack, providing the most biologically literal anchor on early growth. Given the fast early growth period for greater amberjack, we fit an additional two models that constitute a shape comparison within the hierarchical L0 framework: (M7) a Gompertz growth model and (M8) a logistic growth model fit with the same hierarchical structure as M3. For M1 and M2, the conventional VBGM was used:

$$L_{i,t} = L_{\infty,i}\left(1 - e^{-K(t-t_{0,i})}\right)$$

Where $L_{i,t}$ is the predicted length (FL, mm) at age t in subregion i, $L_{\infty,i}$ is the predicted asymptotic length for stock i, K_i is the growth rate for stock i,and $t_{0,i}$ is the theoretical age when length would be zero for stock i. For models M2–M4, the VBGM was reparametrized in terms of L0 following Smart and Grammer (2021):

$$L_{a,s} = L_{\infty,s}\left(L_{\infty,s} - L_{0,s}\right)e^{-K_s \cdot a}$$

where L0 replaces $t_{o,s}$ and has a direct biological interpretation as length-at-birth. This reparameterization allows the inclusion of biologically meaningful prior information on early growth and ensures that fitted curves pass through plausible length values at age 0. The Gompertz growth model was parameterized as:

$$L_{a,s} = L_{0,s} \cdot e^{\ln(L_{\infty,s}/L_{0,s})(1-e^{-K_s \cdot a})}$$

and the Logistic growth model as:

$$L_{a,s} = \frac{L_{\infty,s} \cdot L_{0,s} \cdot e^{-K_s \cdot a}}{L_{\infty,s} + L_{0,s}(e^{K_s \cdot a} - 1)}$$

For all models, priors were selected to be consistent with published applications of growth models in marine fishes (Smart & Grammer, 2021), and to reflect realistic biological bounds observed in the data. Informative priors for L∞ were specified as Normal distributions informed by a range of published estimates of asymptotic lengths for greater amberjack ($L_{\infty,i} \sim \mathrm{Normal}(1200, 0.0001)$; Harris et al., 2007; Murie & Parkyn, 2008). Priors for K were defined as weakly informative Uniform distributions ($K_i \sim \mathrm{Uniform}(0,5)$), providing flexibility that allowed the model to closely fit around known lengths but fluctuate around K to estimate growth rates primarily from the data (Smart & Grammer, 2021). For M2, t0 was assigned a weakly informative Normal prior centered at –1.2 with sufficient variance to allow data-driven estimation in each subregion. For M3, L0 was assigned an

independent weakly informative Normal prior in each subregion, truncated to positive values. For M4, subregion-specific L0 values were drawn from a shared Normal distribution with a weakly informative hyperprior on the population mean, enabling information from subregions with abundant age-0 observations to inform L0 estimates in subregions where such observations were sparse. For M5, L0 was fixed at 150 mm to provide a hard anchor at the smallest observed length and represent the most restrictive constraint on early growth. The residual standard deviation σ was assigned a weakly informative Uniform (0, 100) prior in all models.

Parameter estimation was conducted using the Markov Chain Monte Carlo (MCMC) method, implemented in JAGS with 100,000 iterations from four chains and a burn-in period of 50,000 iterations for each chain. The model was run until convergence was reached. Final posterior distributions were assessed for convergence both visually with a diagnostic plot and with the Gelman-Rubin test where any values <1.1 indicated convergence. Posterior distributions were summarized to obtain mean parameter estimates and 95% credible intervals (CRI). Models were compared using leave-one-out information criterion (LOOIC) calculated with the loo R package, with the best-supported model identified by the lowest LOOIC value and LOOIC weights (LOOICw) used to quantify relative support among candidates.

Modeling was conducted in two stages to first capture broad subregional patterns and then resolve finer-scale differences. In the first stage, a subregional model was fit to the full dataset (n = 2,823) with sexes pooled to obtain growth curves for each subregion using the hierarchical L_o VBGM described above (M3). In the second stage, a subset of individuals (n = 1,066) with known sex was used to estimate sex-specific growth. Due to limited sample sizes of sexed individuals in the western Gulf, the two Gulf subregions were pooled, and growth was modeled with a single grouping variable with four levels (M:Gulf, F:Gulf, M:Atlantic, F:Atlantic). The sex-specific model retained the hierarchical L_o structure of M3 but applied it to the region level (Gulf, Atlantic) rather than to sex_region, because length-at-birth is not expected to vary between sexes prior to maturation. Specifically, $L\infty$ and K were estimated separately for each of the four sex_region groups, while L0 was shared across sexes within each region and drawn from a shared Normal hyperdistribution across the two regions. To anchor early growth, the 61 age-0 individuals in the dataset were retained and incorporated through a two-pathway likelihood. Sexed fish (age > 0) contributed to the full VBGM with sex_region-specific $L\infty$ and K and region-shared L0, while unsexed age-0 fish contributed only to the region-shared L0 estimates (since at age 0 the VBGM reduces exactly to L0). Both pathways shared a common residual standard deviation σ. This structure allowed the age-0 fish from a single region to anchor L0 for that region directly, with the other region's L0 informed through the hierarchical hyperdistribution. Priors and MCMC settings were identical to those used in the first-stage M3 model. For both stages, growth parameters were considered significantly different if the 95% CRI did not overlap. For the sex-specific model, sexual dimorphism in asymptotic length was further evaluated by computing the posterior probability that female $L\infty$ exceeded male $L\infty$ within each region ($P(L\infty,F > L\infty,M)$) directly from the joint posterior draws.

Mortality

Both age and growth-based natural mortality (M) estimates outlined by Then et al. (2014) were applied to amberjack length-at-age data. First, the updated Hoenig equation (Hoenig, 1983) using maximum age (tmax):

$$M = 4.899_{t_{max}}^{-0.916}$$

And then the modified Pauly (1980) that uses the growth parameters L∞ and K:

$$M = 8.87\ K^{0.73} {L_\infty}^{-0.33}$$

To get an estimate of M across all age classes, the age-based M estimate from (Then et al., 2014) and the growth parameters for each stock were scaled across age-classes using the methodology following Lorenzen et al. (2022) adjusted for L0 with the equation:

$$M(a) = M_{L\infty}(1 - \left(1 - \frac{L_0}{L_\infty}\right) e^{-K a})^c$$

Where a is age; a0 = age at length 0; M(a) is the natural mortality rate at age; ML∞ is the natural mortality rate at asymptotic length L∞; K is the von Bertalanffy growth rate; c is the allometric scaling of mortality with body length = -1 (Lorenzen, 2022).

RESULTS

Samples

Sagittal otoliths were sampled from a total of 2,899 greater amberjack from 2014 – 2024 from fishery dependent (~74%) and fishery-independent sources (~26%). Sample distributions varied regionally, with slightly more individuals collected from the western Gulf (n = 1,111) than the eastern Gulf (n = 1,048) and the smallest number of samples coming from the Atlantic (n = 737). Three samples did not have an associated capture location and therefore were discarded from further analyses. Age estimates were obtained for 2,826 individuals (western Gulf = 1,087; eastern Gulf = 1,015; Atlantic = 724), spanning age classes 0 to 15 years (Table 7.1). The oldest individuals observed were 14 years in the Gulf and 15 years in the Atlantic. Most individuals (87%) were ≤5 years old, and the most observed ages were 2 (15.3%), 3 (27.9%), and 4 (24.9%). Significant differences in mean lengths were observed among subregions for ages 2–5 (ANOVA; $p < 0.01$), with the largest differences observed at ages 3 and 4 ($F = 153.0$, $p < 0.001$; $F = 56.8$, $p < 0.001$). Fish from the western Gulf were larger than both the Atlantic and eastern Gulf at all ages where differences were detected (2–5; $p < 0.05$), with the greatest differences observed between eastern Gulf and western Gulf at ages 2 and 3 (age 2: 128 mm, $p < 0.001$; age 3: 140 mm, $p < 0.001$). No significant differences were detected for ages 1, or 6–8 ($p > 0.05$). Sex was determined for 1065 individuals (37.8%), comprising 305 males and 551 females in the Gulf, and 101 males and 108 females in the Atlantic. Only 61 individuals from the western Gulf were sexed, therefore comparisons of sex ratios were between the pooled Gulf samples and the Atlantic. The observed sex ratio (1:1.81) in the Gulf differed from the expected ratio of 1:1 (chi-square test; $p < 0.001$); however, the observed sex ration in the Atlantic (1:1.06) did not ($p = 0.679$). Visual examination of sex ratios across fork length using a linear model indicated that in the Gulf, the proportion of females increased with length, whereas in the Atlantic, sex ratios remained relatively constant across lengths (S1). These trends were consistent with the overall chi-square results, highlighting size-specific variation in the Gulf but not in the Atlantic. The observed length distribution for all greater amberjack (combined sexes) in the sample in the Gulf (n = 2,102), ranged from 138 to 1,492 mm FL, whereas in the Atlantic FL ranged from 211 to 1,520 mm (Table 7.1).

Growth rates

Among the eight candidate models, the three VBGM models that estimated early growth from the data had LOOIC values within 0.4 units of one another (LOOICw = 0.297, 0.354, and 0.349, respectively; Table 7.2). Constraining L0 to the smallest observed length (M4) reduced predictive performance (ΔLOOIC = 32.04) and

constraining it to biological length-at-hatch (6) reduced it substantially (ΔLOOIC = 300.64). The Gompertz (M7) and Logistic (M8) shapes were also outperformed by the VBGM (ΔLOOIC = 41.91 and 103.15, respectively), and the fixed-t_0 baseline performed worst overall (OG, ΔLOOIC = 399.55). The hierarchical L0 VBGM (M3) was selected as the preferred model, as it produced estimates equivalent to the LOOIC-best model while explicitly accommodating the unequal age-0 sampling across subregions. Although model fit differed, regional differences in growth was consistent across all eight candidate models for the eastern Gulf but varied for the Atlantic and western Gulf depending on model structure (Supplementary Table 7.1, Supplementary Figure 7.2). Across all eight models, the eastern Gulf consistently showed the slowest growth rate and the largest asymptotic length, with 95% CRIs for K and L∞ that did not overlap those of either the Atlantic or the western Gulf. In contrast, the Atlantic and western Gulf were generally not distinguishable from one another. However, subregional differences were greater using the Gompertz and Logistic models (Supplementary Table 7.2). In both of these models, the 95% CRIs for K did not overlap among any of the three subregions, separating all three regions on growth rate. For all models, the growth curves for the eastern Gulf were the most variable at older ages (Supplementary Figure 7.2).

Using the hierarchical L0 model (based on LOOIC), greater amberjack growth parameters (L∞ and K) differed among subregions (Table 7.3, Figure 7.1). Growth was slowest in the eastern Gulf (K = 0.29, 95% CRI = 0.25–0.33), where fish also attained the largest asymptotic length (L∞ = 1,261 mm FL, 95% CRI = 1,211–1,317). Growth in the western Gulf was fastest (K = 0.61, 95% CRI = 0.57–0.65), with a smaller asymptotic length than the eastern Gulf (L∞ = 1,117 mm FL, 95% CRI = 1,103–1,131). Growth in the Atlantic (K = 0.51, 95% CRI = 0.46–0.57) was similar to the western Gulf, and fish reached the smallest asymptotic lengths (1,091 mm FL, 95% CRI = 1,069–1,115) (Figure 7.1). Greater amberjack in the Atlantic and western Gulf attained asymptotic length ~2 years (~22%) faster than those in the eastern Gulf (in ~ 6 vs. 8 years). Estimated L0 ranged from 177 mm FL (95% CRI = 151–203) in the western Gulf to 275 mm FL (95% CRI = 232–315) in the eastern Gulf, with an intermediate value in the Atlantic (203 mm FL, 95% CRI = 133–269).

Sexual dimorphism in growth varied by region (Table 7.3, Figure 7.2). Females in the Gulf reached the largest asymptotic lengths (L∞ = 1,263 mm FL, 95% CRI = 1,217–1,312) and exhibited the slowest growth rates (K = 0.32, 95% CRI = 0.29–0.35). Males in the Gulf grew faster than Gulf females (K = 0.40, 95% CRI = 0.35–0.45) but reached smaller asymptotic lengths (L∞ = 1,119 mm FL, 95% CRI = 1,068–1,172). The posterior difference was 143 mm (95% CRI = 74–210) and the probability that female L∞ exceeded male L∞ (P = 1.000), indicating strong sexual dimorphism in asymptotic length. In the Atlantic, growth rates were faster for both sexes (Males: K = 0.49, 95% CRI = 0.39–0.60; Females: K = 0.46, 95% CRI = 0.37–0.55), and asymptotic lengths were smaller than in the Gulf (Males: 1,107 mm FL, 95% CRI = 1,048–1,174; Females: 1,165 mm FL, 95% CRI = 1,099–1,239). The posterior difference was 58.0 mm with a 95% CRI spanning zero (−32.7–149.9) and a posterior probability of P = 0.896, indicating there was no differences in asymptotic length between males and females in the Atlantic. Across both regions, female asymptotic length exceeded male asymptotic length on average, but this difference was credible only in the Gulf.

Differences in VBGF parameter estimates among subregions and sexes were reflected by variation in predicted length-at-age (Figure 7.3). Both male and female greater amberjack from the Atlantic were larger at age-1, age-3, and age-5 than males and females from the Gulf, reflecting faster growth rates of individuals in the Atlantic. However, from the pooled subregional model, greater amberjack in the western Gulf were larger than fish in the eastern Gulf and the Atlantic until age-10, with fish from the eastern Gulf reaching larger lengths after age-10. Females in the Gulf were larger than males at age-5 and age-10, and larger than other group at age-10, consistent with the larger asymptotic lengths of female greater amberjack in the Gulf.

Mortality

Baseline M from the modified Hoenig from Then et al. (2014) was 0.41 (tmax= 15) in the Atlantic and 0.44 in the Gulf (tmax= 13). However, we know that the maximum recorded age for a greater amberjack is tmax= 19 (Murie et al., 2020) which would give a baseline M = 0.33. The modified Pauly growth-based mortality equation resulted in an estimate of 0.37 and 0.44 for females and males in the Gulf, respectively. Estimates of M were higher for both males (0.52) and females (0.49) in the Atlantic. Lorenzon-M scaled across age-classes for the Atlantic and Gulf showed a similar trend (Figure 7.4).

DISCUSSION

Growth differences in greater amberjack from this study support the scientific evidence for stock delineation between the Gulf and the Atlantic (Harris et al., 2007; Murie et al., 2020; Murie & Parkyn, 2008), with potential for substructuring in the Gulf based on demographic differences. Greater amberjack exhibited significant regional and sex-specific differences in growth. Individuals from the Gulf attained larger asymptotic lengths compared to those from the Atlantic. Within the Gulf, eastern Gulf fish reached larger asymptotic lengths, but grew slower than in the western Gulf, reflecting divergent growth trajectories between subregions. Similar spatial variation in growth has been documented in other marine fish (DeVries & Grimes, 1997; Saari et al., 2014; Shepard et al., 2008), suggesting environmental or ecological drivers of fish growth between the Gulf and Atlantic may be widespread.

Because greater amberjack are sampled almost entirely as sub-adults and adults, length-at-age estimates are sensitive to assumptions about length-at-birth and the early growth trajectory (Smart & Grammer, 2021). We evaluated this directly by fitting a sensitivity gradient of models ranging from data-driven early growth (free t_o, independently estimated L0, and hierarchical L0) to biologically literal constraints (L0 fixed at the smallest observed length, and length-at-hatch). Constraining L_o progressively degraded model fit, with the hatch-length constraint performing worst which is consistent with Smart & Grammer (2021), who showed that forcing a single-phase von Bertalanffy curve through true length-at-birth is incompatible with adult-dominated length-at-age data. However, regional differences in growth were present across every parameterization and growth curve shape examined, suggesting that these results represent observed regional differences rather than an artifact of modelling choice.

The extent of spatial variation in growth can vary by geographical area; for example, growth rates can differ among reef patches within the same system, or over wide geographic distances, such as latitudinal gradients (Adams et al., 2018; Allman, 2007; Streich et al., 2017), and may reflect environmental conditions, food availability, and length-specific mortality patterns. Our analysis suggests heterogeneous growth rates exist within the Gulf: the mean growth rate is slower, and the asymptotic length is larger in the eastern Gulf, compared to growth in the western Gulf. Notably, however, fish in the western were larger at younger ages (2–5), indicating that differences in asymptotic lengths do not necessarily translate to larger length-at-age across the life history. This distinction is consistent with a trade-off between early growth rate and asymptotic length and is important when interpreting regional growth differences. Differences in somatic growth between the eastern and western Gulf are seen with other marine species like king mackerel (DeVries & Grimes, 1997) and red snapper (Saari et al., 2014). Red snapper (Lutjanus campechanus) attain larger sizes and grow slower off the coast of Alabama and Louisiana compared to the rest of their range in the Gulf (Saari et al., 2014). There is some evidence that king mackerel reach larger asymptotic lengths in the eastern Gulf, compared to the western Gulf (DeVries & Grimes, 1997), however these are historic growth rates and may not be comparable to patterns seen in greater amberjack. The potential mechanisms for the observed demographic variation in somatic growth

include environmental differences, such as suitable habitat availability, ecosystem productivity, or community structure among subregions. Although greater amberjack are pelagic spawners with larvae dependent on Sargassum (Wells & Rooker, 2004), the dispersal rate of larvae between geographic areas may not be substantial enough to offset the local environmental or fishery effects on somatic growth.

Sexual size dimorphism is typically attributed to natural and sexual selection, where variation in reproductive success leads to differences in the optimal body size for each sex (Parker, 1992). Female-biased SSD is associated with larger body size, which favors increased fecundity. SSD is not fixed, however, and can shift with environmental and demographic conditions. In the Gulf, females grew slower but reached larger asymptotic lengths than males, consistent with fecundity selection. Female-skewed SSD is also evident in king mackerel, where Gulf females are larger than Atlantic females, as well as larger than males in both the Atlantic and Gulf (Shepard et al., 2008). In contrast, SSD was not apparent in the Atlantic, where both sexes were smaller overall and exhibited faster growth. Historically, female-biased SSD was documented in both the Gulf (Murie & Parkyn, 2008) and the Atlantic (Harris et al., 2007). Our results differ from Harris et al. (2007), who found that greater amberjack sampled off the Southeastern US Atlantic coast from 2000–2004 showed female-biased sexual size dimorphism. This divergence may reflect differences in stock status, environmental conditions, migration ecology, or some combination of factors (Blanckenhorn, 2005; Parker, 1992; Young, 2005). The Atlantic population has rebuilt from past overfishing (SEDAR, 2020a), exhibits smaller overall body size, and undergoes energetically costly north-south migrations (Branham et al., manuscript in preparation; Berry & Burch, 1979; McClellan & Cummings, 1997), potentially constraining asymptotic length and homogenizing growth trajectories between sexes. The sex-specific growth trajectory in the Gulf is likely linked to the female-skewed sex ratio (1:1.8) that is known to increase with age and size (Smith et al., 2018). Whereas in the Atlantic, balanced sex ratios (1:1.05) coincide with overlapping male and female growth trajectories.

The oldest greater amberjack observed in this study was estimated to be age 15, which represents a decrease (~4 year) in recorded longevity for greater amberjack. Differences in longevity for different studies may be an artifact of sampling as older individuals may be relatively rare in the population. Natural mortality declined rapidly from age-0 to age-3 across all stocks and sexes, with the steepest early reductions observed in the Gulf populations. By older ages (>10 years), Atlantic fish maintained higher mortality rates than Gulf fish, while females consistently exhibited slightly lower mortality than males within each stock. These small but consistent differences in natural mortality mirror the observed regional growth variation and suggest that life-history differences may contribute to stock productivity. Despite some growth differences and skewed sex ratios, greater amberjack are not currently modeled as separate sexes in SEDAR stock assessments because sex is rarely recorded in landing data (SEDAR, 2020a, 2020b). Growth results from this study, sex-specific growth and mortality should be considered in future assessment models in the Gulf if sufficient data become available.

Greater amberjack are already managed as two separate stocks by the Gulf and South Atlantic Fishery Management Councils (SEDAR, 2020a, 2020b). The observed growth differences further support this stock delineation. Genetic divergence between stocks may also play a role, alongside management differences. The exploitation histories of the Gulf and the Atlantic have been quite different, as the Gulf stock has been overfished for more than a decade. The Atlantic has a less restrictive minimum size limit (SEDAR, 2020a), which may disproportionately remove faster-growing individuals. This could bias growth estimates toward larger asymptotic lengths without representing the true growth potential of the stock. Previous age and growth estimates from the Atlantic range from asymptotic lengths of 1,105 mm FL (males) to 1,351 mm FL (females; Harris et al., 2007). The combined-sex parameters used in the South Atlantic stock assessment ($L\infty$ = 1,204 mm FL; K = 0.28) fall within the credible interval of our sex-specific model estimates. The smaller asymptotic lengths

observed here may reflect a lack of older, larger females in samples collected from 2014–2024. Estimates of greater amberjack growth in the Gulf have varied over time, with estimates of L_∞ ranging from 1,109 mm to 1,436 mm depending on the data used and modeling choice (Manooch & Potts, 1997; Murie et al., 2020; Murie & Parkyn, 2008; Thompson et al., 1999). Our results showed differences in growth between sexes and subregions within the Gulf that could explain variation among previous estimates that did not account for spatial or sex-specific effects. These findings highlight the need to consider sampling structure and demographic representation when interpreting growth dynamics and their implications for stock assessments.

Spatial and sex-specific variation in individual growth, age structure, and mortality have implications for the assessment and management of greater amberjack in the Gulf and Atlantic. Management units or stocks can be defined not only by genetic differences but also by spatial heterogeneity in demographic processes (Cadrin, 2020; Punt, 2003). Our results demonstrate clear differences in growth and asymptotic length between Gulf and Atlantic stocks, as well as between eastern and western Gulf subregions, and reveal female-biased SSD in the Gulf but not in the Atlantic. These findings support growing evidence for stock delineation between the Gulf and Atlantic, and potentially within Gulf delineation, and highlight how local environmental conditions and sex ratio dynamics interact to shape growth and mortality patterns. Importantly, current stock assessments do not account for sex-specific differences in life history in the Gulf, which may bias estimates of reproductive potential and stock resilience. Our findings provide a critical foundation to incorporate demographic heterogeneity into future assessments. A better understanding of other demographic traits like spatially explicit size-at-maturity and movement dynamics would further inform assessment and management of this species.

TABLES AND FIGURES

Table 7.1: Number (n) and percentage (%) of greater amberjack (*Seriola dumerili*) in each 100-mm fork length size-bin landed in the Gulf of America (Gulf) subregions and western Atlantic.

FL (mm)	Atlantic (ATL)		Eastern Gulf (EG)		Western Gulf (WG)		Total	
	n	%	n	%	n	%	n	%
100–199	0	0.00	0	0.00	47	4.32	47	1.66
200–299	7	0.97	5	0.49	3	0.28	15	0.53
300–399	3	0.41	0	0.00	2	0.18	5	0.18
400–499	8	1.10	24	2.36	1	0.09	33	1.17
500–599	17	2.35	38	3.74	5	0.46	60	2.12
600–699	35	4.83	115	11.33	5	0.46	155	5.48
700–799	83	11.46	230	22.66	18	1.66	331	11.71
800–899	128	17.68	217	21.38	30	2.76	375	13.27
900–999	199	27.49	203	20.00	333	30.63	735	26.01
1,000–1,099	169	23.34	126	12.41	419	38.55	714	25.27
1,100–1,199	61	8.43	41	4.04	161	14.81	263	9.31
1,200–1,299	10	1.38	10	0.99	44	4.05	64	2.26
1,300–1,399	3	0.41	4	0.39	13	1.20	20	0.71
≥1,400	1	0.14	2	0.20	6	0.55	9	0.32
Total (n)	724	100.00	1015	100.00	1087	100.00	2826	100.00

Table 7.2: LOOIC results of Bayesian growth models. Best fitting models identified in bold typeface

Model	Shape	LOOIC	SE	dLOOIC	LOOICw
M1: Fixed t0	VBGM	34261.2	115.9	399.5	0.0
M2: Free t0		33862.0	116.7	0.4	0.3
M3: Free L0		33861.7	116.7	0.0	0.4
M4: Hierarchical L0		33861.7	116.8	0.0	0.3
M5: Fixed L0=150		33893.7	117.2	32.0	0.0
M6: Fixed L0=2.9		34162.3	117.9	300.6	0.0
M7: Gompertz (Hierarchical L0)	Gompertz	33903.6	115.7	41.9	0.0
M8: Logistic (Hierarchical L0)	Logistic	33964.8	114.8	103.1	0.0

Table 7.3: Parameters from the von Bertalanffy growth model (VBGM) for all greater amberjack (*Seriola dumerili*) pooled and for sex specific growth models (L∞ = average maximum fork length [mm]; K = growth coefficient; L_0 = length at age zero.

Model	Parameter	Estimate	2.50%	97.50%
Subregional				
Atlantic	L∞	1091	1069	1115
	K	0.51	0.46	0.57
	L0	203	133	269
Eastern Gulf	L∞	1261	1211	1317
	K	0.29	0.25	0.33
	L0	275	232	315
Western Gulf	L∞	1117	1103	1131
	K	0.61	0.57	0.65
	L0	177	151	203
Sex-Region Specific				
Atlantic				
Males	L∞	1106	1048	1174
	K	0.49	0.39	0.6
Females	L∞	1164	1099	1239
	K	0.46	0.37	0.55
Shared	L0	239	168	323
Gulf of America (Gulf)				
Males	L∞	1119	1068	1172
	K	0.4	0.35	0.45
Females	L∞	1262	1217	1311
	K	0.32	0.29	0.35
Shared	L0	201	179	223

Supplementary Table 7.1: Parameters from the von Bertalanffy growth model (VBGM) for all greater amberjack (*Seriola dumerili*) subregional growth models that represent the sensitivity test for early growth influence on parameter estimates (L∞ = average maximum fork length [mm]; K = growth coefficient; L_0 = length at age zero.

Model	Subregion	Parameter	Estimate	2.50%	97.50%
M1: Fixed t0	Atlantic				
		K	0.35	0.33	0.38
		L∞	1148	1123	1175
	Eastern Gulf				
		K	0.24	0.23	0.26
		L∞	1324	1280	1370
	Western Gulf				
		K	0.35	0.33	0.37
		L∞	1211	1192	1231
M2: Free t0	Atlantic				
		K	0.51	0.45	0.57
		L∞	1092	1069	1117
		t0	-0.41	-0.63	-0.24
	Eastern Gulf				
		K	0.28	0.25	0.32
		L∞	1268	1216	1327
		t0	-0.90	-1.14	-0.68
	Western Gulf				
		K	0.61	0.57	0.65
		L∞	1117	1103	1131
		t0	-0.28	-0.34	-0.23
M3: Free L0 (indep)	Atlantic				
		K	0.52	0.46	0.58
		L0	195	127	261
		L∞	1090	1068	1114

	Eastern Gulf				
		K	0.29	0.25	0.33
		L0	276	235	316
		L∞	1263	1213	1320
	Western Gulf				
		K	0.61	0.57	0.65
		L0	175	149	201
		L∞	1117	1103	1131
M5: Fixed L0=150	Atlantic				
		K	0.55	0.52	0.58
		L∞	1082	1064	1102
	Eastern Gulf				
		K	0.38	0.36	0.41
		L∞	1182	1153	1213
	Western Gulf				
		K	0.62	0.58	0.66
		L∞	1115	1102	1129
M6: Fixed L0=2.9	Atlantic				
		K	0.64	0.61	0.68
		L∞	1063	1046	1080
	Eastern Gulf				
		K	0.49	0.46	0.51
		L∞	1121	1098	1146
	Western Gulf				
		K	0.69	0.65	0.74
		L∞	1107	1094	1119

Supplementary Table 7.2: Parameters from the von Bertalanffy growth model (VBGM) for all greater amberjack (*Seriola dumerili*) subregional growth models that represent the shape analysis for greater amberjack growth using the best fitting model from the early growth sensitivity analysis (L∞ = average maximum fork length [mm]; K = growth coefficient; L_0 = length at age zero.

Model	Subregion	Parameter	Estimate	2.50%	97.50%
M4: Von Bertalanffy	Atlantic				
		K	0.51	0.46	0.57
		L0	203	133	269
		L∞	1091	1069	1115
	Eastern Gulf				
		K	0.29	0.25	0.33
		L0	275	232	315
		L∞	1261	1211	1317
	Western Gulf				
		K	0.61	0.57	0.65
		L0	177	151	203
		L∞	1117	1103	1131
M7: Gompertz	Atlantic				
		K	0.69	0.61	0.77
		L0	289	239	339
		L∞	1068	1048	1090
	Eastern Gulf				
		K	0.42	0.37	0.47
		L0	346	314	378
		L∞	1205	1163	1252
	Western Gulf				
		K	0.90	0.84	0.96
		L0	191	165	216
		L∞	1093	1081	1104

M8: Logistic	Atlantic				
		K	0.84	0.74	0.94
		L0	345	304	389
		L∞	1057	1038	1077
	Eastern Gulf				
		K	0.54	0.49	0.60
		L0	390	363	417
		L∞	1176	1139	1217
	Western Gulf				
		K	1.25	1.17	1.35
		L0	208	184	232
		L∞	1076	1067	1086

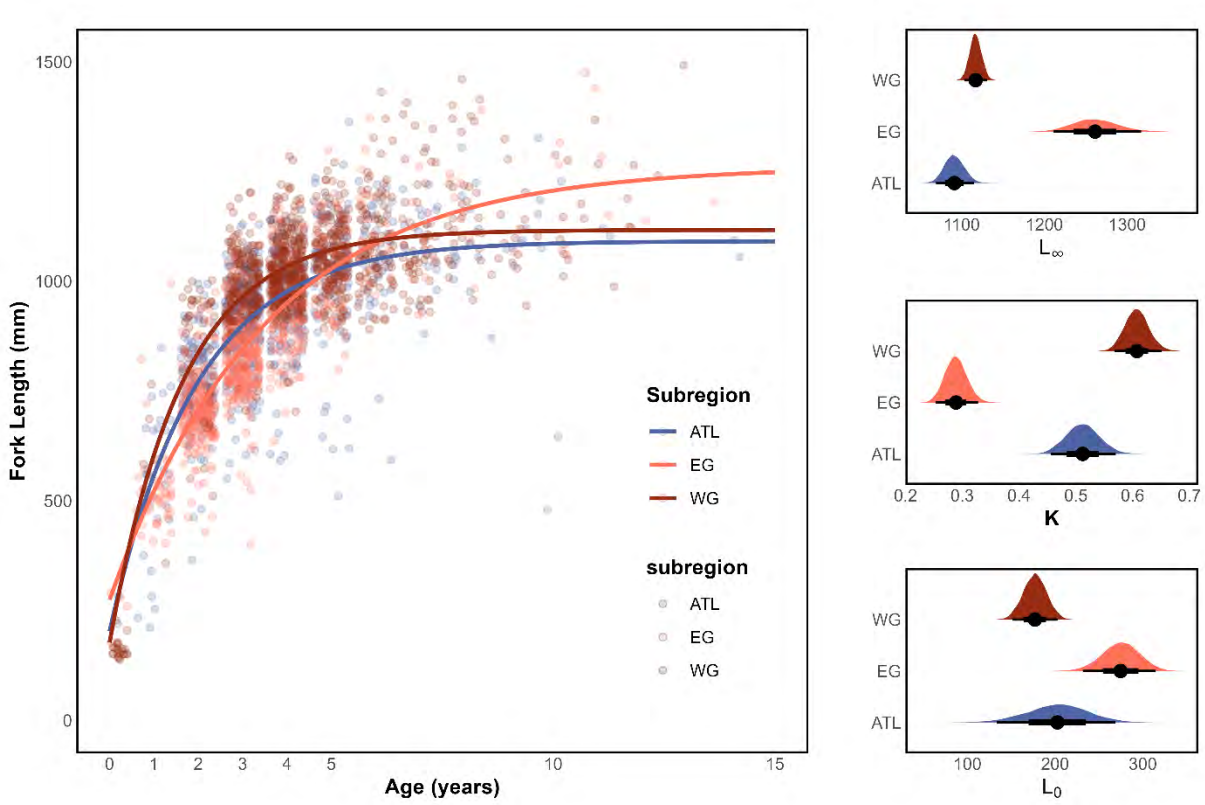


Figure 7.1: Right) Predicted Von Bertalanffy subregion-specific growth curves for greater amberjack (*Seriola dumerili*) predicted fork length as a function of calendar age for all observations (n = 2,823). Left) Posterior distributions, along with 95% and 80% credible intervals, for each growth parameter. L∞ = asymptotic length; K = growth coefficient. ATL = Atlantic; EG = Eastern Gulf; WG = Western Gulf. *

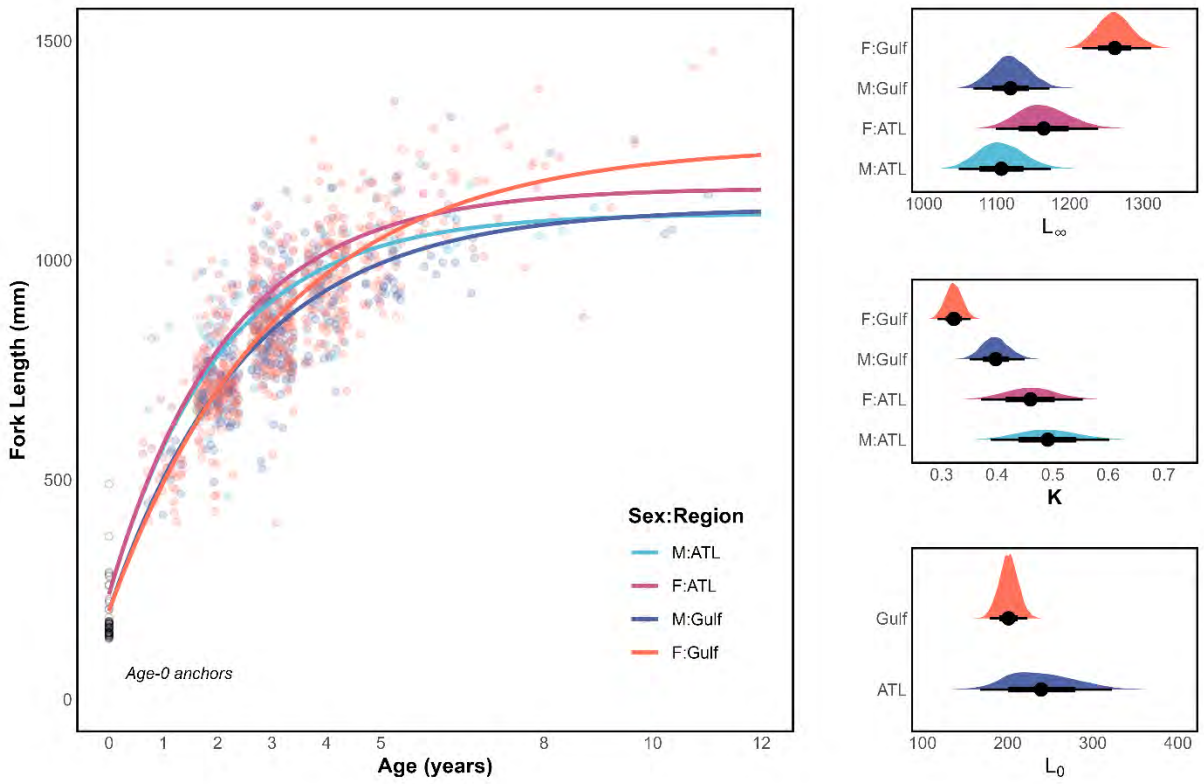


Figure 7.2: Right) Predicted Von Bertalanffy sex- and stock- specific growth curves for greater amberjack (*Seriola dumerili*) predicted fork length as a function of calendar age for all observations (n = 1,066). Left) Posterior distributions, along with 95% and 80% credible intervals, for each growth parameter. L∞ = asymptotic length; K = growth coefficient. M:ATL = Males, Atlantic; F:ATL = Females, Atlantic; M:Gulf, Males, Gulf; F:Gulf = Females, Gulf.

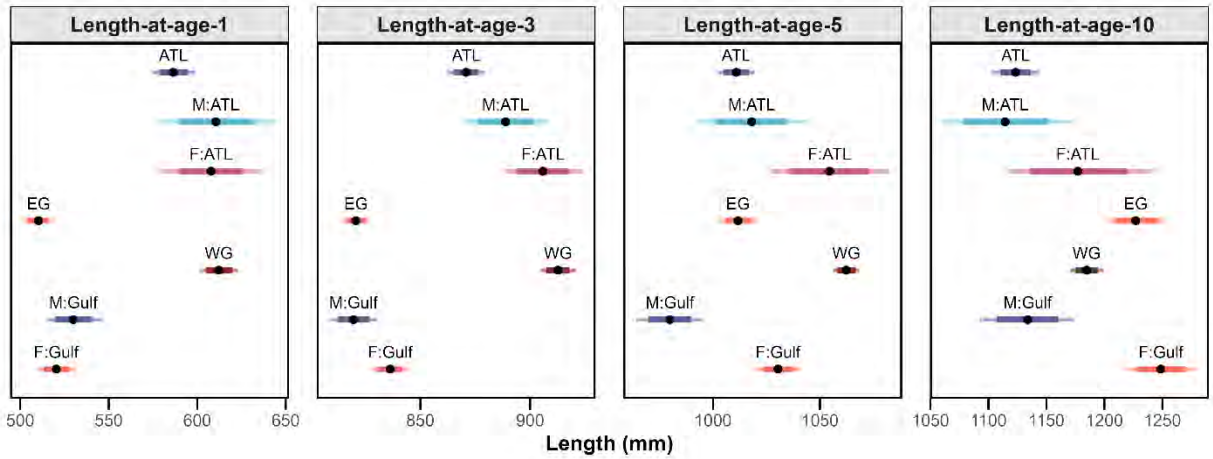


Figure 7.3: Predicted length-at-age using subregion- and sex-specific von Bertalanffy parameter estimates from greater amberjack (*Seriola dumerili*) from the Gulf of America (Gulf) and Atlantic stocks. The thicker line represents the 80% credible intervals, and the thin portion represents the 95% credible interval. ATL: Atlantic; M:ATL = Males, Atlantic; F:ATL = Females, Atlantic; EG = eastern Gulf; WG: western Gulf; M:Gulf, Males, Gulf; F:Gulf = Females, Gulf.

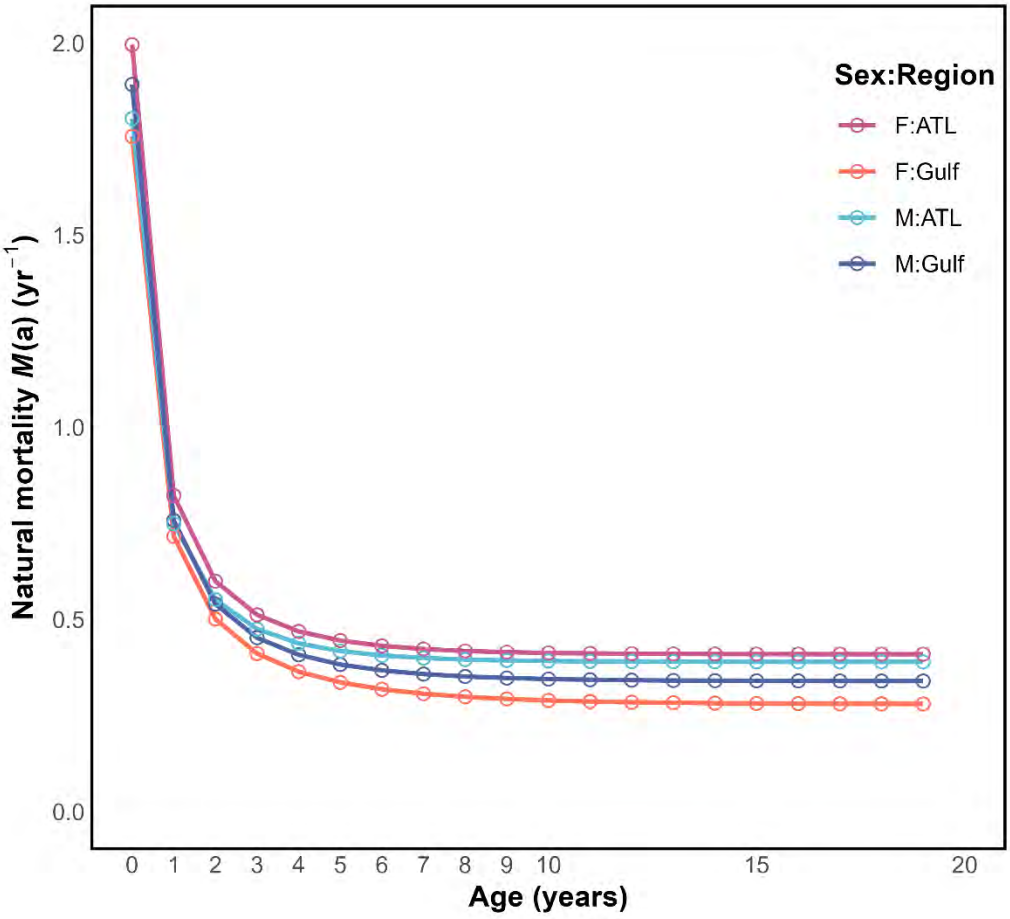
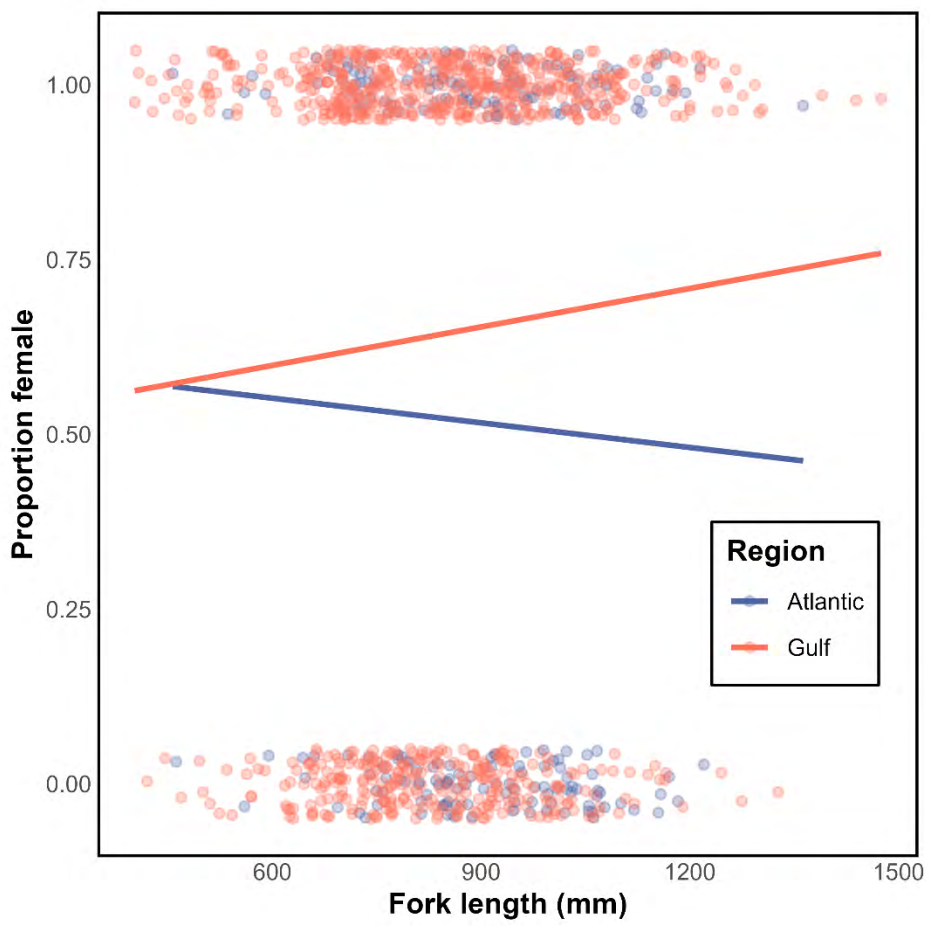
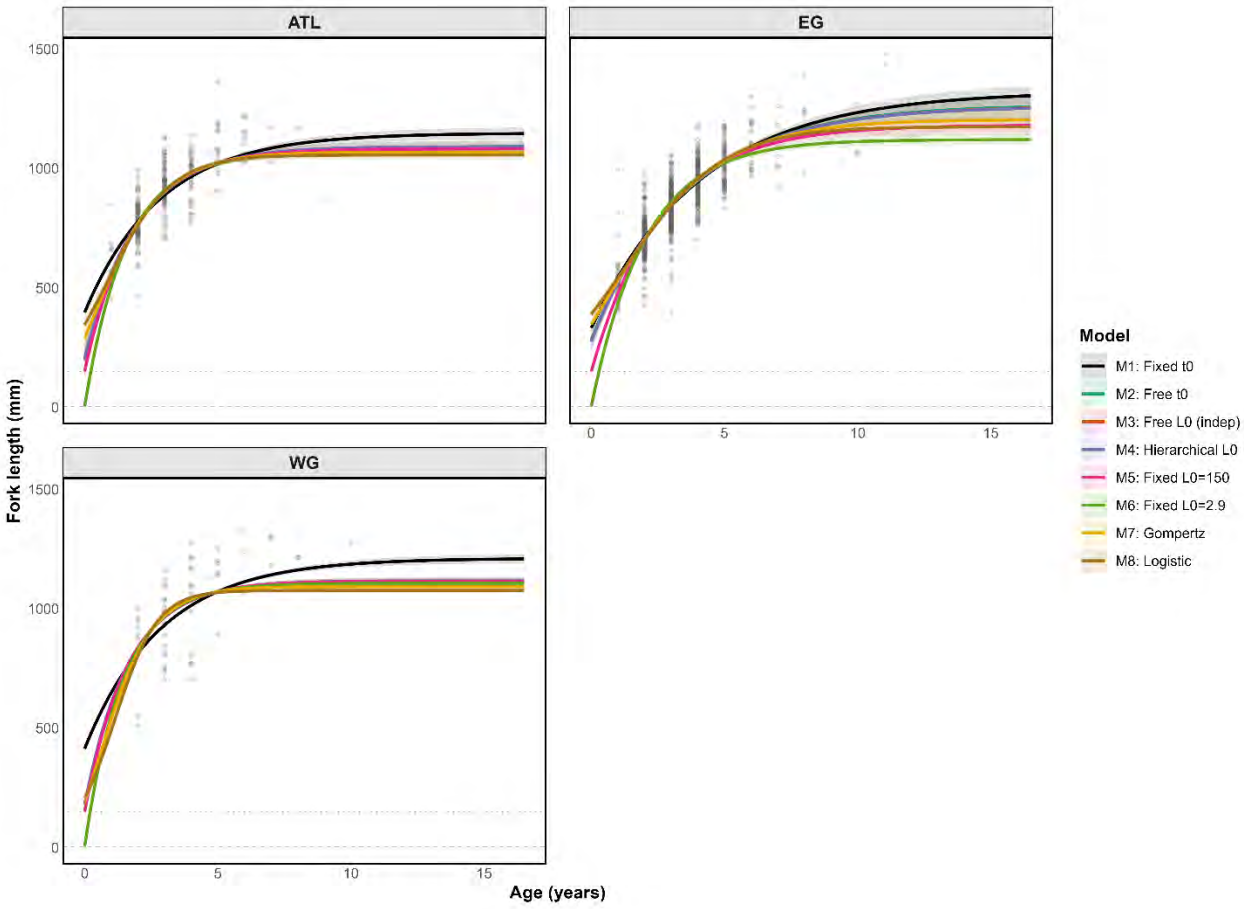


Figure 7.4: Stock- and sex- specific Lorenzen natural mortality (Lorenzen-M) estimates for greater amberjack (*Seriola dumerili*) scaled across age-classes using the growth parameters estimated from the Von Bertalanffy growth model. ATL = Atlantic; Gulf = Gulf of America.



Supplementary figure 7.1 (S1): Proportion of female greater amberjack (*Seriola dumerili*) across different lengths (fork length, mm) in the Atlantic and Gulf.



Supplementary figure 7.2 (S2): Predicted Von Bertalanffy subregion-specific growth curves for greater amberjack (*Seriola dumerili*) predicted fork length as a function of calendar age for all models.

REFERENCES

Adams, G. D., Leaf, R. T., Ballenger, J. C., Arnott, S. A., & McDonough, C. J. (2018). Spatial variability in the growth of Sheepshead (Archosargus probatocephalus) in the Southeast US: Implications for assessment and management. Fisheries Research, 206, 35-43. https://doi.org/https://doi.org/10.1016/j.fishres.2018.04.023

Allman, R. J. (2007). Small-scale spatial variation in the population structure of vermilion snapper (Rhomboplites aurorubens) from the northeast Gulf of Mexico. Fisheries Research, 88(1), 88-99. https://doi.org/https://doi.org/10.1016/j.fishres.2007.07.017

Allman, R. J., Patterson III, W. F., Fioramonti, C. L., & Pacicco, A. E. (2018). Factors affecting estimates of size at age and growth in grey triggerfish Balistes capriscus from the northern Gulf of Mexico. Journal of Fish Biology, 92(2), 386-398. https://doi.org/https://doi.org/10.1111/jfb.13518

Banks, K. G., Streich, M. K., & Stunz, G. W. (2024). Age, growth, and mortality of King Mackerel in the western Gulf of Mexico. Marine and Coastal Fisheries, 16(1), e10278. https://doi.org/https://doi.org/10.1002/mcf2.10278

Berry, F. H., & Burch, R. K. (1979). Aspects of the amberjack fisheries. Proceedings of the Gulf and Caribbean Fisheries Institute, 31, Marathon, FL.

Blanckenhorn, W. U. (2005). Behavioral causes and consequences of sexual size dimorphism. Ethology, 111(11), 977-1016.

Cadrin, S. X. (2020). Defining spatial structure for fishery stock assessment. Fisheries Research, 221. https://doi.org/10.1016/j.fishres.2019.105397

Croll, J. C., van Kooten, T., & de Roos, A. M. (2023). The consequences of density-dependent individual growth for sustainable harvesting and management of fish stocks. Fish and Fisheries, 24(3), 427-438.

Denit, K., & Sponaugle, S. (2004). Growth variation, settlement, and spawning of gray snapper across a latitudinal gradient. Transactions of the American Fisheries Society, 133(6), 1339-1355.

DeVries, D., & Grimes, C. (1997). Spatial and temporal variation in age and growth of king mackerel, Scomberomorus cavalla, 1977-1992. Fishery Bulletin, 95, 694-708.

Erisman, B. E., Reed, E. M., Román, M. J., Mascareñas-Osorio, I., van der Sleen, P., López-Sagástegui, C., Aburto-Oropeza, O., Rowell, K., & Black, B. A. (2021). Relationships among somatic growth, climate, and fisheries production in an overexploited marine fish from the Gulf of California, Mexico. Fisheries Oceanography, 30(5), 556-568. https://doi.org/https://doi.org/10.1111/fog.12537

Gertseva, V., Matson, S. E., & Cope, J. (2017). Spatial growth variability in marine fish: example from Northeast Pacific groundfish. ICES Journal of Marine Science, 74(6), 1602-1613. https://doi.org/10.1093/icesjms/fsx016

Gislason, H., Daan, N., Rice, J. C., & Pope, J. G. (2010). Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries, 11(2), 149-158.

Haddon, M. (2011). Modelling and Quantitative Methods in Fisheries (2nd ed.). Chapman and Hall/CRC. https://doi.org/https://doi.org/10.1201/9781439894170

Hargrove, J. S., Murie, D. J., Parkyn, D. C., Saarinen, E. V., & Austin, J. D. (2018). Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico. Genetica, 146(4-5), 393-402. https://doi.org/10.1007/s10709-018-0031-1

Harris, P. J., Wyanski, D. M., White, D. B., Mikell, P. P., & Eyo, P. B. (2007). Age, growth, and reproduction of greater amberjack off the aoutheastern U.S. Atlantic coast. Transactions of the American Fisheries Society, 136(6), 1534-1545. https://doi.org/10.1577/t06-113.1

Hilborn, R., & Walters, C. J. (1992). Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Springer Science & Business Media. https://doi.org/10.1007/978-1-4615-3598-0

Hoenig, J. M. (1983). Empirical use of longevity data to estimate mortality rates. Fishery Bulletin, 82(1), 898-903.

Houde, E. D. (2002). Chapter 3: Mortality. In L. A. Fuiman & R. G. Werner (Eds.), Fishery science: The unique contribution of early life stages. Blackwell Scientific Publishing.

Kotowych, N., Smalås, A., Amundsen, P.-A., & Primicerio, R. (2023). Climate warming accelerates somatic growth of an Arctic fish species in high-latitude lakes. Scientific Reports, 13(1), 16749. https://doi.org/10.1038/s41598-023-43654-1

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49(4), 627-642.

Lorenzen, K. (2016). Toward a new paradigm for growth modeling in fisheries stock assessments: Embracing plasticity and its consequences. Fisheries Research, 180, 4-22. https://doi.org/https://doi.org/10.1016/j.fishres.2016.01.006

Lorenzen, K. (2022). Size- and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. Fisheries Research, 255, 106454. https://doi.org/https://doi.org/10.1016/j.fishres.2022.106454

Lorenzen, K., Camp, E. V., & Garlock, T. M. (2022). Natural mortality and body size in fish populations. Fisheries Research, 252, 106327. https://doi.org/https://doi.org/10.1016/j.fishres.2022.106327

Manooch, C. S., III, & Potts, J. C. (1997). Age, growth, and mortality of greater amberjack, *Seriola dumerili*, from the U.S. Gulf of Mexico headboat fishery. Bulletin of Marine Science, 61(3), 671-683.

Maunder, M. N., & Wong, R. A. (2011). Approaches for estimating natural mortality: Application to summer flounder (Paralichthys dentatus) in the U.S. mid-Atlantic. Fisheries Research, 111(1), 92-99. https://doi.org/https://doi.org/10.1016/j.fishres.2011.06.016

McClellan, D. B., & Cummings, N. J. (1997). Preliminary analysis of tag recapture data of Greater Amberjack, *Seriola dumerili*, in the southeastern United States. Proceedings of the Gulf and Caribbean Fisheries Institute, 49, Miami, Florida.

Midway, S. R., Wagner, T., Arnott, S. A., Biondo, P., Martinez-Andrade, F., & Wadsworth, T. F. (2015). Spatial and temporal variability in growth of southern flounder (Paralichthys lethostigma). Fisheries Research, 167, 323-332. https://doi.org/https://doi.org/10.1016/j.fishres.2015.03.009

Murie, D., Parkyn, D., Smith, G., Allman, R., Pacicco, A., Carroll, J., & Smith, N. (2020). Gulf of Mexico greater amberjack (*Seriola dumerili*) growth model for SEDAR 70 operational Assessment (SEDAR70-WP-10).

Murie, D. J., & Parkyn, D. C. (2008). Age, growth and sex maturity of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico (SEDAR33-RD13).

Pacicco, A. E., Allman, R. J., Lang, E. T., Murie, D. J., Falterman, B. J., Ahrens, R., & Walter III, J. F. (2021). Age and growth of yellowfin tuna in the U.S. Gulf of Mexico and western Atlantic. Marine and Coastal Fisheries, 13(4), 345-361. https://doi.org/https://doi.org/10.1002/mcf2.10158

Parker, G. A. (1992). The evolution of sexual size dimorphism in fish. Journal of Fish Biology, 41(sB), 1-20. https://doi.org/https://doi.org/10.1111/j.1095-8649.1992.tb03864.x

Pauly, D. (1980). On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. Journal du Conseil, 39(2), 175-192. https://doi.org/10.1093/icesjms/39.2.175

Pauly, D. (2019). Female Fish Grow Bigger – Let's Deal with It. Trends in Ecology & Evolution, 34(3), 181-182. https://doi.org/https://doi.org/10.1016/j.tree.2018.12.007

Plummer, M. (2003, 20-22 March 2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna.

Plummer, M., Stukalov, A., & Denwood, M. (2025). rjags: Bayesian Graphical Models using MCMC. In (Version R package version 4-17)

Punt, A. E. (2003). The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. Fisheries Research, 65(1-3), 391-409.

Punt, A. E., & Hilborn, R. A. Y. (1997). Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7(1), 35-63. https://doi.org/10.1023/A:1018419207494

Saari, C., Cowan, J. H., & Boswell, K. (2014). Regional differences in the age and growth of red snapper (Lutjanus campechanus) in the U.S. Gulf of Mexico. Fishery Bulletin, 112, 261-273. https://doi.org/10.7755/FB.112.4.3

SEDAR. (2014). SEDAR 33 – Gulf of Mexico Greater Amberjack Stock Assessment Report.

SEDAR. (2020a). SEDAR 59– South Atlantic Greater Amberjack Stock Assessment Report. http://sedarweb.org/sedar-59

SEDAR. (2020b). SEDAR 70 – Gulf of Mexico Greater Amberjack Stock Assessment Report.

Shepard, K., Patterson III, W. F., DeVries, D. A., & Palmer, C. (2008). Age and growth and stock mixing in Gulf of Mexico and Atlantic king mackerel (Scomberomorus cavalla). Proceedings of the Gulf and Caribbean Fisheries Institute, Volume 60,

Smart, J. J., & Grammer, G. L. (2021). Modernising fish and shark growth curves with Bayesian length-at-age models. PLoS ONE, 16(2), e0246734. https://doi.org/10.1371/journal.pone.0246734

Smith, G. H., Murie, D. J., & Parkyn, D. C. (2018). Effects of sex-specific fishing mortality on sex ratio and population dynamics of Gulf of Mexico greater amberjack. Fisheries Research, 208, 219-228. https://doi.org/10.1016/j.fishres.2018.07.011

Stearns, S. C., & Koella, J. C. (1986). The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. Evolution, 40(5), 893-913.

Streich, M. K., Ajemian, M. J., Wetz, J. J., Williams, J. A., Shipley, J. B., & Stunz, G. W. (2017). A Comparison of Size Structure, Age, and Growth of Red Snapper from Artificial and Natural Habitats in the Western Gulf of Mexico. Transactions of the American Fisheries Society, 146(4), 762-777. https://doi.org/10.1080/00028487.2017.1308884

Su, N.-J., Sun, C.-L., Punt, A. E., Yeh, S.-Z., & Dinardo, G. (2011). Evaluation of a spatially sex-specific assessment method incorporating a habitat preference model for blue marlin (Makaira nigricans) in the Pacific Ocean. Fisheries Oceanography, 20(5), 415-433. https://doi.org/https://doi.org/10.1111/j.1365-2419.2011.00593.x

Then, A. Y., Hoenig, J. M., Hall, N. G., Hewitt, D. A., & Jardim, H. e. E. (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. https://doi.org/10.1093/icesjms/fsu136

Thompson, B. A., Beasley, M., & Wilson, C. A. (1999). Age, distribution, and growth of greater amberjack, *Seriola dumerili*, from the north central gulf of mexico. Fishery Bulletin, 97, 362-371.

Wang, S.-P., Sun, C.-L., Punt, A. E., & Yeh, S.-Z. (2005). Evaluation of a sex-specific age-structured assessment method for the swordfish, Xiphias gladius, in the North Pacific Ocean. Fisheries Research, 73(1), 79-97. https://doi.org/https://doi.org/10.1016/j.fishres.2005.01.001

Wells, R. J. D., & Rooker, J. R. (2004). Distribution, age, and growth of young-of-the-year greater amberjack (*Seriola dumerili*) associated with pelagic Sargassum. Fisheries Bulletin, 102, 545-554.

Williams, A. J., Farley, J. H., Hoyle, S. D., Davies, C. R., & Nicol, S. J. (2012). Spatial and Sex-Specific Variation in Growth of Albacore Tuna (Thunnus alalunga) across the South Pacific Ocean. PLoS ONE, 7(6), e39318. https://doi.org/10.1371/journal.pone.0039318

Young, K. A. (2005). Life–history variation and allometry for sexual size dimorphism in Pacific salmon and trout. Proceedings of the Royal Society B: Biological Sciences, 272(1559), 167-172.

Smart, J. J., & Grammer, G. L. (2021). Modernising fish and shark growth curves with Bayesian length-at-age models. PLoS ONE, 16(2), e0246734. https://doi.org/10.1371/journal.pone.0246734

8. Acoustics Overview

Active Acoustics Approach to Estimate Greater Amberjack Abundance and Advancements Towards Echo Classification

Kevin M. Boswell and Allison L. White

8.1 Introduction

Active acoustics is commonly used by researchers to conduct fish abundance surveys in marine ecosystems (Simmonds and MacLennan, 2005). By transmitting sound pulses into the water column and measuring the intensity of returning echoes, echosounders allow researchers to quantify the density and spatial distribution of organisms over broad temporal and spatial scales (Medwin and Clay, 1998; Demer et al., 2015) and play a central role in many fishery-independent survey programs. In the context of fisheries acoustics, active acoustic surveying represents a non-invasive methodology that enables expansive spatial coverage and synoptic sampling of the full water column making it particularly valuable for modern ecosystem-based fisheries management frameworks (Horne, 2000).

Despite these advantages, the application of active acoustic approaches to enumerate reef-associated species presents unique challenges. The foremost challenge is that species of interest are often closely associated with structurally complex environments such as natural banks, artificial reefs, and oil and gas platforms, where bottom interference, multipath reflections, and dense biological scattering from aggregations complicate signal interpretation (Ona and Mitson, 1996; White et al., 2022). Furthermore, unlike pelagic systems dominated by relatively homogeneous schools, reef-associated assemblages are often taxonomically diverse and spatially patchy (Barange, 1994; Horne, 2000). These conditions require more refined processing and classification approaches than those traditionally applied in open-water surveys.

A central limitation of active acoustics in reef systems is attributing measured backscatter to specific taxa because the observed backscatter is a direct representation of the morphological (and acoustic) properties of the organism, and not necessarily its taxonomic classification (Boswell et al., 2020). Consequently, echo-classification approaches are needed to attribute measured backscatter into biologically meaningful groups. Traditional echo-classification approaches rely on a combination of backscatter thresholds, aggregation structure, spatial morphology, and increasingly, spectral frequency responses to isolate echoes consistent with focal taxa while excluding non-target scatterers (Demer et al., 2009; Korneliussen et al., 2009; Korneliussen, 2018; Urmy et al., 2024). However, in systems that are characterized by taxonomically diverse assemblages, overlap in acoustic characteristics among co-occurring species can limit discrimination and result in classification uncertainty (Lekanda et al., 2024).

Recent progress in the use of broadband methods offers potential to discriminate among targets based on their morphological and backscattering properties across a continuous range of frequencies (Jech et al., 2017; Bassett et al., 2018; Boswell et al., 2020). Unlike traditional narrowband systems that operate at a single frequency, broadband systems transmit across a range of frequencies allowing the frequency-dependent backscatter response of individual targets and aggregations to be resolved (Jech et al., 2017; Bassett et al., 2018; Boswell et al., 2020). The spectral response provides frequency-specific amplitude structure, spectral slopes, and null patterns that are all influenced by swimbladder morphology, body composition, and size. As a result, broadband approaches enable discrimination among taxa based not solely on absolute target strength, but on the shape and variability of their spectral response (Verma et al., 2017; Setiawan et al., 2020; Khodabandeloo et al., 2021; Roa et al., 2022). Consequently, in taxonomically diverse systems where overlap in single-frequency target strength is

common, these spectral features may provide an additional basis for classification and may reduce uncertainty relative to threshold-based methods alone. Although comprehensive spectral response libraries remain under development, broadband data is helping to establish the framework for more robust echoclassification and future refinement of species discrimination models.

In structure-associated systems where multiple large-bodied taxa co-occur, an integrative acoustic-optical approach can substantially reduce bias in estimating fish quantities (Scoulding et al., 2023). Acoustic data provide quantitative estimates of volumetric fish density, while video observations provide direct species identification and relative composition. Species proportions derived from video can be applied to acoustically derived abundance estimates to scale density to species-specific abundance (Stunz et al., 2021). This hybrid framework takes advantage of the strengths of both methods: the volumetric coverage of acoustic sampling coupled with the taxonomic certainty of optical surveys. Both approaches possess distinct strengths and weaknesses. Active acoustics provides high-resolution spatial coverage, depth-independent sampling, and quantitative density estimates, but requires careful calibration and processing and is limited in species discrimination. Video-based methods provide direct species identification and behavioral context but are constrained by field of view, visibility conditions, attraction bias, and non-volumetric sampling. By integrating these complementary methods, it is possible to mitigate the constraints and leverage the strengths of each method to provide a more refined estimate of abundance for reef-associated species.

The objectives of this chapter are to describe the active acoustic framework used to estimate the abundance of Greater Amberjack across structured and unstructured habitats broadly across the Gulf of America. In this section we outline the echo-classification approach developed to isolate echoes consistent with large-bodied, loosely aggregated fishes that could include *Seriola.* We further describe the methodology to integrate acoustic abundance estimates using video-derived species proportions and quantify uncertainty in abundance estimation using complementary echo integration and echo counting approaches. Throughout the scope of this project, broadband frequency-modulated signals were adopted to aid in improving the echo-classification of reef fishes, including Greater Amberjack, in the northern Gulf.

8.2 Data Collection

Mobile active acoustic surveys were conducted with split-beam Simrad EK80 echosounders at 947 sites, either concurrently with or shortly before/after camera deployments. All project regions deployed a 70 kHz echosounder, and additional echosounders (38 kHz and/or 120 kHz) were incorporated as available. Echosounders were deployed at the surface with transducers oriented towards the seafloor using towed, pole-mounted, or hull-mounted configurations according to each vessel's capabilities. On all vessels, echosounders were operated in frequency-modulated upsweep pulses with fast ramping, a 1.024 ms pulse duration, and maximum ping rate. Frequency-modulated upsweep pulses of the 38, 70, and 120 kHz echosounders were operated at 35-45, 45-90, and 90-170 kHz ranges, respectively. Echosounders were calibrated using the standard sphere method with 38.1 mm tungsten carbide (6% cobalt binder) spheres and 22 or 25 mm tungsten carbide (6% cobalt binder) spheres to fill nulls in the theoretical target strength spectra of the 38.1 mm sphere in the 45-170 kHz range (Demer et al., 2015).

During data collection, vessels maintained a speed of 4-6 knots. Three different transect designs were selected to accommodate different habitat scales and attributes, informed by the findings of White et al. (2023). The primary design used at most sites was a systematic parallel design (n=748; Figure 8.1a). This design consisted of seven 200 m long transects with 10 m spacing between the center transect and the first concentric pair of neighboring transects, 20 m between the first concentric pair and the second concentric pair of transects, and 40 m between

the second pair of concentric transects and the outermost pair of transects. The center transect was positioned over the coordinates of the camera deployment or structure in the case of artificial reefs, and the inner three transects were intended to provide at least three passes where the area sampled by the cameras and/or artificial structure were included in the echosounder beam(s). The two outer pairs of transects were spaced further apart to provide a more robust representation of variation in the distribution of fish at each site. Directional orientation of the transects was determined based on sea state or orientation of the structure at artificial reefs (though most were arranged north to south). An extended version of parallel transects was employed by the Western Gulf regions (LA and TX) to provide a larger spatial scale sample of their extensive natural bank reef systems and unconsolidated bottom habitats. The extended parallel design consisted of eight 1 km long transects evenly spaced 100 m apart (n=117; Figure 8.1b). These transects were oriented north to south or east to west depending on the sea state and/or direction of the banks. Lastly, the 82 standing oil and gas platforms surveyed in this study provided a unique challenge, as echosounders could not pass directly over structures extending above the water. At these sites, an open star transect pattern was adopted which consisted of three to eight transects driven alongside the platforms as close to the structure as the vessel was safely capable of allowing (Figure 8.1c). Transects were driven along each side of the platform complex (or between individual platforms in the case of a cluster of structures not connected above the water), and each transect extended 100 m beyond the structure on either side.

8.3 Acoustic Data Processing

Raw active acoustic data were visualized and processed using Echoview 15.0 (Echoview Software Pty. Ltd.). Acoustic data was "cleaned" to remove backscatter from unwanted sources, including noise interference, surface cavitation, and seafloor echoes. Acoustic and electrical interference signals were removed from pulse compressed volume backscatter (S_V; dB re 1 m^{-1}) and target strength (TS; dB re 1 m^2) data from each available transducer using background and impulse noise removal algorithms (DeRobertis and Higginbottom, 2007; Ryan et al., 2015). A threshold offset algorithm was applied to the 45-90 kHz S_V data to exclude any surface cavitation below the transducers (-70 dB re 1 m^{-1} boundary threshold), and data within 5 m (twice the near-field range of the lowest frequency used) of the transducer faces were removed to exclude near-field interference and transducer ringdown (Medwin and Clay, 1998). A best bottom candidate detection algorithm with a 0.5 m offset was used to exclude the seafloor and acoustic deadzone from analysis (Ona and Mitson, 1996). The maximum TS value below the seafloor exclusion boundary was calculated for each ping in the 45-90 kHz range, and any pings with a maximum TS of less than -50 dB re 1 m^2 were removed to eliminate ping dropout. As only the 70 kHz echosounder was consistently deployed across all regions, further acoustic data processing was performed using only 45-90 kHz data.

8.3.1 Echoclassification

Two levels of echoclassification were performed on 45-90 kHz data. First, backscatter was filtered to retain only echoes from fish, including all gas-bearing fishes and larger-bodied non-swimbladdered fishes such as sharks. All data points where TS was less than -46 dB re 1 m^2 were removed from the s_v echogram. This was sufficient to exclude echoes from all scatterers other than fish and some particularly dense clusters of particulates in nepheloid layers, which are common in the northern Gulf (Figure 8.2a-b). A school detection SHAPES algorithm was used to exclude the denser patches of nepheloid for a fish-only s_v echogram (Table 8.1; Figure 8.2c; Barange, 1994).

Echoes classified as fish were further filtered to exclude small-bodied fishes < 20 cm TL that were densely aggregated (Figure 8.2d). High-density aggregations of organisms generate imprecise TS measurements, and as *Seriola dumerili* do not exhibit this behavior it was beneficial to exclude such fish from further analysis

(MacLennan, 1990; Toresen, 1991). The resulting filter included only echoes from fish that could potentially be *Seriola*, defined as large-bodied fishes > 20 cm TL that were either observed as isolated individuals or loosely aggregated. The results of the two echoclassification levels were manually scrutinized and the "potential *Seriola*" filter was applied to all further acoustic data processing variables to generate abundance estimates of *Seriola* and similar fishes.

8.3.2 Abundance Estimation

Abundance of fish retained in the potential *Seriola* filter was estimated using echo integration (Eq. 8.1).

$$\rho_v = \frac{\underline{S_V}}{\underline{\sigma_{bs}}} \qquad \text{(Equation 8.1)}$$

Echo integration calculates the volumetric density of fish (ρ_v; fish m^{-3}) in a given cell of water as the mean S_V of that cell ($\underline{s_v}$) divided by the average backscattering cross section of targets in the cell ($\underline{\sigma_{bs}}$). Echointegrated abundance (N_{ei}) is then obtained by multiplying ρ_v by the beam volume sum (m^3), or the sum of the beam volumes across all pings in the cell. To calculate N_{ei} of fishes that could potentially be *Seriola*, 1 m distance by 5 m depth cells were echointegrated. The potential *Seriola* filter was applied to cleaned 45-90 kHz s_v echograms, and $\underline{S_V}$ was calculated for each cell.

In order to further exclude any imprecise TS measurements from high-density fish aggregations in $\underline{\sigma_{bs}}$ single targets were first detected from cleaned, unclassified TS echograms using Echoview's wideband single target detection algorithm (Table 8.2). Detected single targets were then integrated into 50 ping by 5 m depth cells, and any cells where the number of targets per acoustic sampling volume (N_v) was > 0.01 were excluded from further analysis (Parker-Stetter et al., 2009; Eq. 8.2).

$$N_v = \frac{c\tau\psi r^2 \rho_v}{2} \qquad \text{(Equation 8.2)}$$

N_v was calculated following Sawada et al. (1993), where ρ_v was calculated from cleaned, unclassified S_V and cell-averaged TS of single targets, then multiplied by the speed of sound in water (c; ms^{-1}), the effective pulse duration (τ; s), the linear equivalent beam angle (ψ; steradians), and the squared range (r^2; m) of the cell from the transducer face. The potential *Seriola* filter was then applied to the single target echogram where $N_v \leq 0.01$, and $\underline{\sigma_{bs}}$ in Equation 1 was calculated by taking the average TS of the filtered single targets in each 1 m distance by 5 m depth cell.

In addition to echo integration, echo counting was performed to provide a measurement of uncertainty around the N_{ei} estimate for each site. Echo counting (N_{ec}) relies on the identification of echo traces, or multiple echoes of the same target from successive pings. An alpha-beta tracking algorithm was used to detect echo traces of potential *Seriola* (Table 8.3; Blackman, 1986). N_{ec} is then calculated as the sum of detected potential *Seriola* echo traces. Uncertainty in the abundance of potential *Seriola* (u(N)) was estimated as the mean absolute difference between total N_{ei} and N_{ec} at each site (Eq. 8.3).

$$u(N) = \frac{|N_{ei} - N_{ec}|}{2} \qquad \text{(Equation 8.3)}$$

To eliminate vessel-based bias in the area sampled by echosounders at each site, N_{ei} and N_{ec} of potential *Seriola* per 1 m distance by 5 m depth cell were exported from Echoview and the spatial extent of surveyed cells was reduced in R (R Core Team, 2024). For data collected along systematic parallel transects, only cells within a 100 m

radius of the site focal point (the location of either the camera deployment or artificial structure) were retained for the summation of N_{ei} and N_{ec}. At sites where the extended parallel transect design was used, only cells within a 500 m radius of the site focal point were retained. At standing oil and gas platforms where the open star transect design was employed, any cells further than 100 m from the transects between the locations where they intersected were excluded from N_{ei} and N_{ec}. The area sampled by the echosounders was calculated as the sum of the area swept in each retained 1m cell interval. Cell area swept was the footprint of the 45-90 kHz beam, taken as the max seafloor depth of the within the cell multiplied by the tangent of the beam half angle and the distance of the cell multiplied by two. The acoustically-derived abundance of greater amberjack was estimated by multiplying the potential *Seriola* abundance ($N_{ei} \pm u(N)$) by the proportion of greater amberjack MaxN from the MaxN of similar fishes (loosely aggregated and > 20 cm TL) derived from video data.

The results of the regional abundance estimates are provided within each regional chapter.

8.4 Advancements Towards Echo Classification of Active Acoustic Data

The wealth of broadband data collected throughout this project are being utilized to further the development of spectral response libraries for reef fishes in the southeastern USA. Two primary challenges of establishing a framework for acoustic-independent echoclassification from broadband spectra are grouping similar spectra from targets observed *in situ* and linking identified spectral signatures back to their source taxa/size classes. Ancillary studies are currently being conducted in parallel to address both of these challenges, as described below.

8.4.1 Clustering Broadband Spectra from in situ Fish Targets

An unsupervised machine learning framework was developed to cluster similar broadband spectra from targets detected in the mobile active acoustic surveys. We began with data collected in the South Atlantic region, which consistently included 58-84 and 98-167 kHz echosounders. As a wider range of frequencies were available from this dataset, echoes were classified to gas-bearing fishes from cleaned s_v in Echoview using dB summing as described in Fernandes (2009). The resulting gas-bearing fish filter was applied over the single target detection, and the TS broadband spectra were exported from each frequency range (Figure 8.3).

A two-step machine learning framework was implemented in Python 3.12 to remove incoherent target spectra from the dataset and identify clusters among the remaining coherent spectra. The first step in the machine learning framework developed here to cluster broadband spectra from *in situ* fish targets is to identify and remove incoherent signals via hierarchical density-based spatial clustering of applications with noise (HDBSCAN; Campello et al., 2015; Chawarski et al., 2025). A single fish often generates multiple targets in the echosounder beam(s), and there is variability in the TS spectra of these targets detected from the same fish due to a combination of coherent and incoherent energy in their backscatter. Coherent energy is the clean reflectance of a fish's swimbladder that creates the deterministic signal distinct to the size, orientation, and material properties of that fish. Incoherent energy from changes in the tilt of the fish, fish movement, reflectance of other parts of the fish's anatomy (such as the head or tail), and high-density aggregations of fish introduces variability around coherent target signals from the same fish (Demer et al., 2009; MacLennan, 1990; Toresen, 1991). These incoherent signals introduce noise which widely applied clustering algorithms, such as hierarchical clustering and k-means, are highly influenced by.

After HDBSCAN is applied to identify and remove incoherent spectra, the remaining coherent spectra are clustered via hierarchical agglomerative clustering (HAC). HAC builds a taxonomy of spectra that are grouped together by

their similarities until all spectra are grouped into one cluster (Figure 8.4). Spectral similarities can be calculated via several different distance algorithms, some of which quantify similarities between spectra based on differences in the amplitude of their TS across frequencies and others which quantify similarities strictly on spectral shape. To select the optimal distance metric and value at which to separate clusters of broadband spectra, the Optuna package was used to develop a custom optimization function (Akiba et al., 2019). The silhouette score was paired with custom constraints that penalized biologically unrealistic scant mega-clusters and hyper-fragmented micro-clusters while rewarding ecological diversity (Rousseuw, 1987). This approach honors the high diversity of fish communities observed across multiple reef habitats while avoiding the bias toward a small number of broad clusters inherent in an unconstrained silhouette framework.

In an attempt to biologically ground clusters of coherent broadband spectra, the machine learning framework was first applied to targets within 10 m of camera deployment locations (a rough estimate of the average S-BRUV field of view in the South Atlantic). This included 1,973 targets in the 58-84 kHz range and 1,641 targets in the 98-167 kHz range. Of the targets estimated to be within the field of view of the camera, 35.07% and 25.68% were identified as incoherent spectra in the HDBSCAN step from the 58-84 and 98-167 kHz ranges, respectively. For both frequency ranges, a shape-based distance metric (cosine distance) was selected from the Optuna-optimization for HAC of coherent spectra over amplitude-based distance metrics. This indicates that the frequencies at which resonance peaks occur is more significant than the difference in TS when clustering broadband spectra. Optuna-optimization of HAC for coherent spectra selected distance values resulting in 159 clusters for both frequency ranges. The between- and within-cluster stability of these 159 clusters were evaluated by performing 100 iterations of HAC with non-parametric bootstrapping of 90% of the target spectra. For both frequency ranges, all 100 trials resulted in 140-160 clusters (Figure 8.5). This possibly suggests that broadband spectra are distinguishable to the species-level or lower, as 108 fish species were observed in the camera data at the same South Atlantic sites. Within-cluster stability was examined by calculating the Jaccard coefficient (J) for each cluster across HAC bootstrapping trials. Clusters with $J \geq 0.6$ were considered “stable”, indicating that the spectra within that cluster remain together more often than not across resampling events. 14 stable clusters were identified in the 58-84 kHz range, and 30 stable clusters were found in the 98-167 kHz range (Figure 8.6). These stable clusters of coherent spectra likely represent commonly recurring broadband spectra unique to a certain size class and/or tilt angle of a distinct species. We continue to attempt to trace these stable clusters of coherent spectra back to their source taxa by expanding the machine learning framework described here to *in situ* broadband spectra observed in other regions and through comparison with theoretical scattering models of broadband spectra predicted from *ex situ* mapping of swimbladder morphologies of common species at these habitats.

8.4.2 Scattering Models of Broadband Spectra Predicted from ex situ CT Scans

Morphological characteristics of reef fish were investigated through the quantitative use of computed tomography (CT) scans on 19 individuals of Greater Amberjack. Conducting CT scans is an efficient means of acquiring high-resolution three-dimensional models of fish, including the internal musculature, skeletal components and swim bladder. This method is superior to traditional bi-planar x-ray imagery as the curvature and fine-scale structure of the bladder can be discerned and modeled. Fish were scanned along the longitudinal axis and resulted in a series of cross-sectional slices through the body yielding detailed images of the internal structure of each individual that could be quantified through image processing. Three dimensional models of bladders were derived for each individual, permitting examination of species-specific morphological variation (Figure 8.7). TS was computed from 30-200 kHz at 2 kHz intervals and over a range of tilt angles from -45 to +45 degrees in 2-degree increments, where 0 degrees represents the dorsal orientation. This range in orientation

was included to examine the TS variation relative to how a fish might be distributed within the water column (Gastauer et al. 2016; see Figures 8.8 and 8.9). These models can be added to other pre-existing models (e.g., Boswell et al. 2020) to refine our understanding of the backscatter patterns observed in these dynamic and biologically heterogeneous habitats in the Southeastern United States.

8.5 Conclusions

Leveraging the advancements of wideband echosounding, machine-learing based echoclassification and modeling has the potential to greatly improve our ability to quantify patterns of change in communities associated with complex habitats in the northern Gulf. While modeling approaches are powerful means to understand the broader patterns expected for backscatter, the importance of *in situ* ground truth data cannot be understated. When combined with additional independent observations these acoustic based approaches will offer robust non-invasive approaches to examine patterns in change, even in high-diversity fish communities.

Tables

Table 8.1. School detection parameters used to classify fishes from all other scatterers and potential *Seriola* fish echoes from small-bodied (<20 cm TL) fishes that were densely aggregated.

SHAPES Parameter	Fishes	Potential *Seriola*
Minimum s_v	-70 dB re $1m^{-1}$	-60 dB re $1m^{-1}$
Minimum Candidate Height	0.01 m	0.5 m
Minimum Candidate Length	3 pings	10 pings
Minimum Total Height	0.01 m	1 m
Minimum Total Length	3 pings	10 pings
Maximum Vertical Linking Distance	0.1 m	0.3 m
Maximum Horizontal Linking Distance	1 ping	3 pings

Table 8.2. Parameters applied in the wideband single target detection algorithm.

Single Target Detection	Value
Minimum TS	-46 dB re 1 m^2
Pulse Length Determination Level	12 dB re 1W
Minimum Normalized Pulse Length	0.2
Maximum Normalized Pulse Length	1.5
Minimum Target Separation	0 m
Maximum SD of Minor-Axis Angles	2°
Maximum SD of Major-Axis Angles	2°

Table 8.3. Alpha-beta tracking algorithm parameters used to detect echo traces of potential *Seriola* targets.

Echo Trace Detection	Value
α	0.7
β	0.5
Exclusion Distance	4 m
Missed Ping Expansion	0 %
Minor-Axis Weights	30 %
Major-Axis Weights	30 %
Range Weights	40 %
TS Weights	20 %
Ping Gap Weights	20 %
Minimum # Single Targets in an Echo Trace	5 single targets
Minimum # Pings in an Echo Trace	4 pings
Maximum Gap Between Single Targets	3 pings

Figures

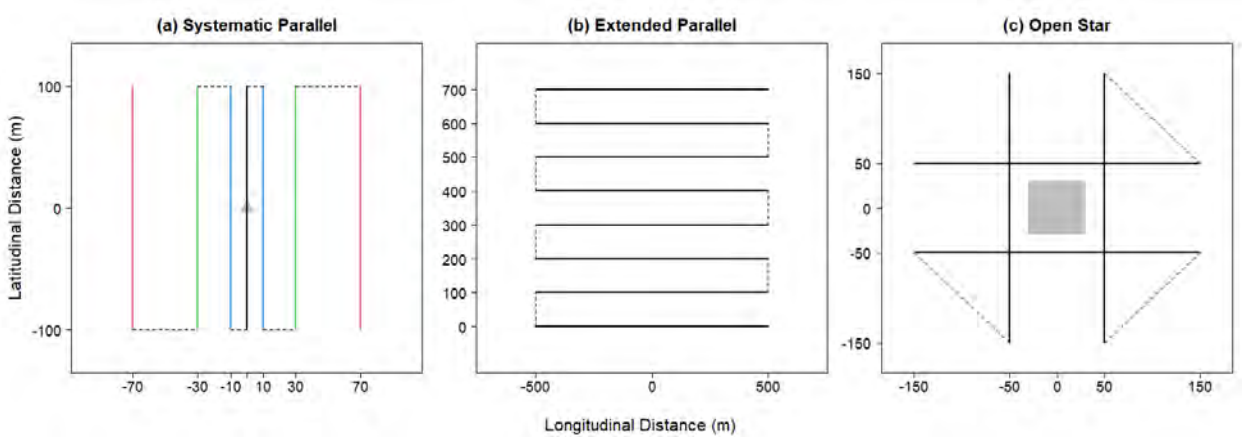


Figure 8.1. Three transect designs used during mobile echosounder surveys. Solid lines indicate portions of the transects where data was retained for analysis, and dotted lines indicate turns between transects which were excluded from analysis. In the systematic parallel design (left panel), the center transect (black) passes over the camera deployment/artificial structure (grey triangle), which is also captured by the first concentric pair of transects (blue). The second pair of concentric transects (green) and outer pair of transects (red) provide more coverage of the habitats in order to obtain a more accurate estimation of the variance in fish abundance. The extended parallel transect design (center panel) covers a larger spatial area and was used at natural bank reefs and unconsolidated habitats in the Western Gulf. The open star design (right panel) was implemented around standing oil/gas platforms (grey square) that extended above the water surface.

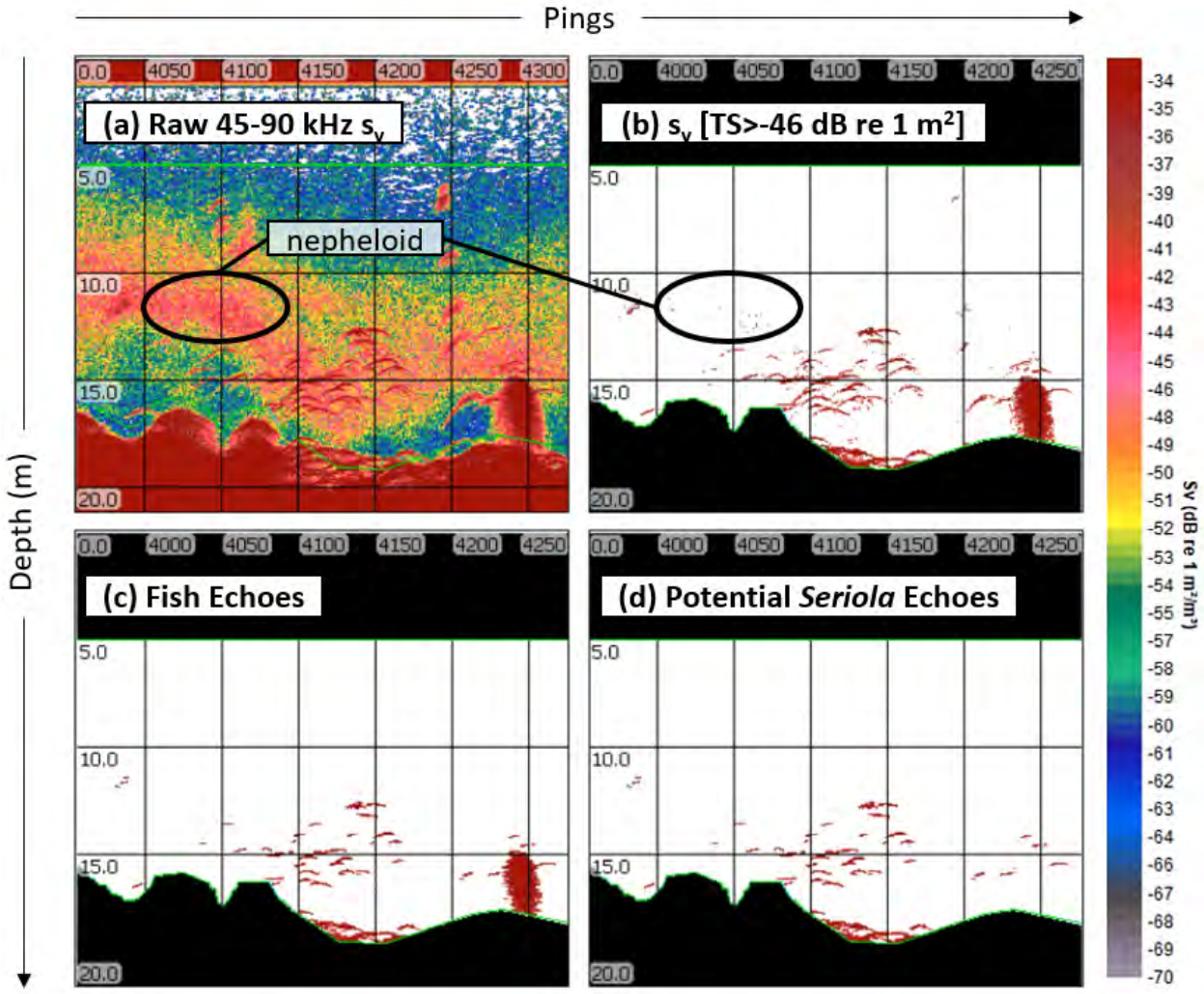


Figure 8.2. Echo classification scheme applied to (a) raw 45-90 kHz s_v data. (b) shows the cleaned S_V echogram filtered to retain only data points between 5 m from the transducer face and 0.5 m above the seafloor (green lines) where TS > -46 dB re 1 m^2. At sites with particularly dense nepheloid layers, such as the one displayed above, particulates may be retained by this filter. (c) shows the results of the SHAPES algorithm to exclude these particulates, retaining only echoes from fish. A second SHAPES algorithm is applied in (d) to retain only echoes from fish that could potentially be *Seriola*, or large-bodied (> 20 cm TL) loosely aggregated fish. Data points in the (d) echogram where s_v > -999 dB re 1 m^{-1} make up the potential *Seriola* filter applied to further analysis.

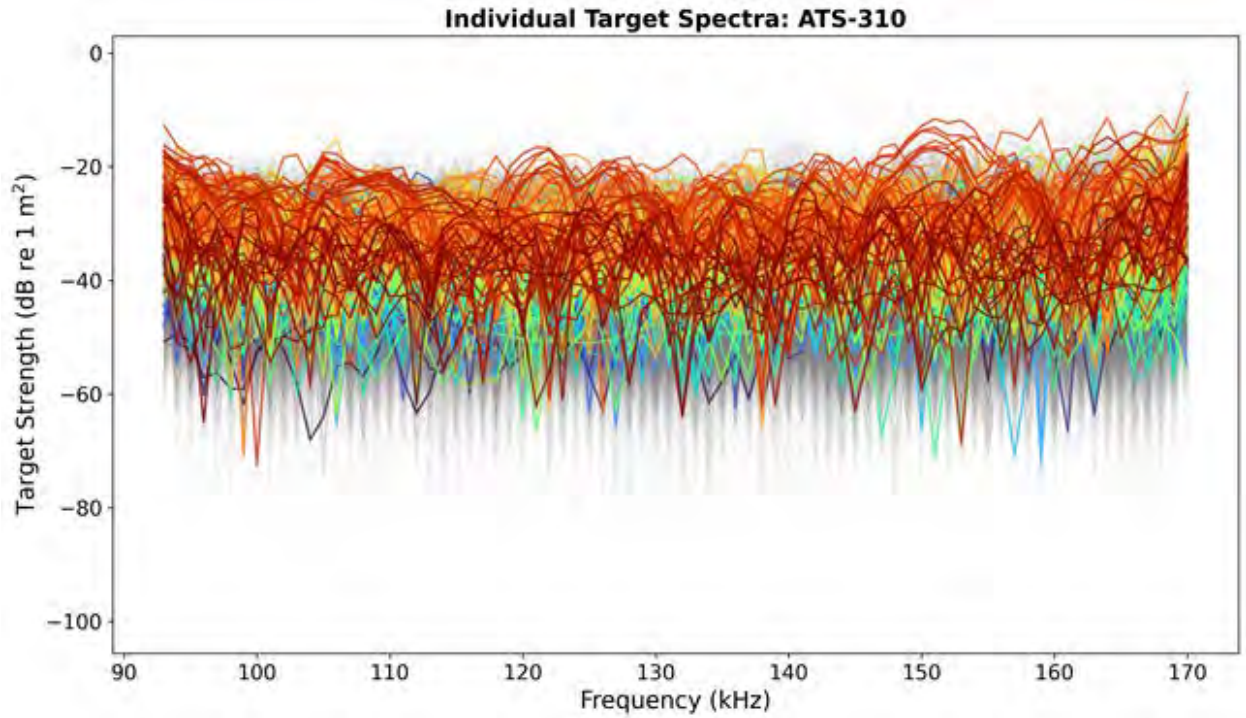


Figure 8.3. Example TS broadband spectra of targets from gas-bearing fishes within a 10 m radius of the S-BRUV at a natural reef in the South Atlantic. Each line represents the broadband spectra of a single target from 98-167 kHz.

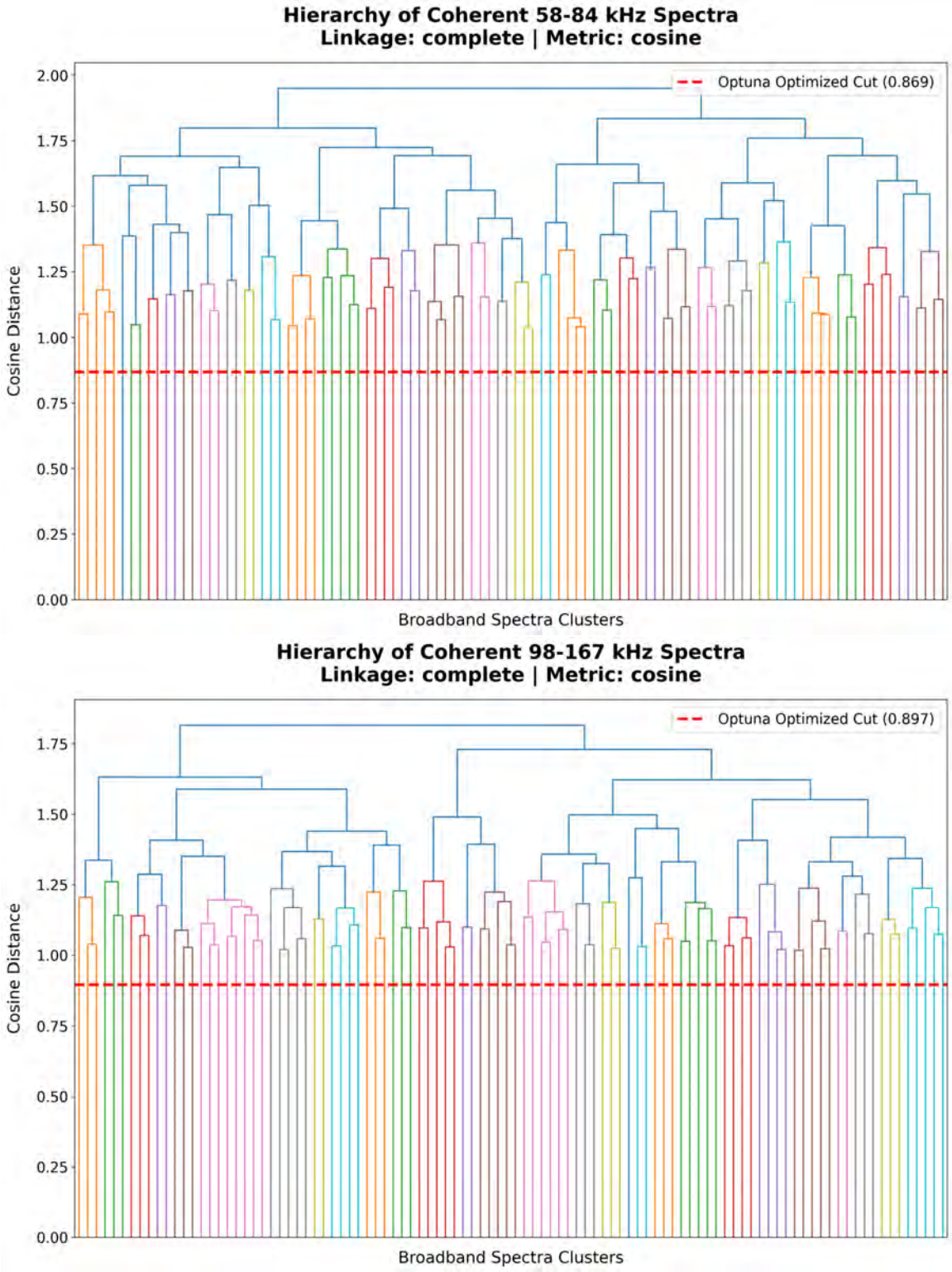


Figure 8.4. Dendrograms showing taxonomies of coherent broadband spectra from hierarchical agglomerative clustering of targets in 58-84 kHz (top) and 98-167 kHz (bottom). The distance values at which to separate clusters optimized via the Optuna package are shown as dashed red lines.

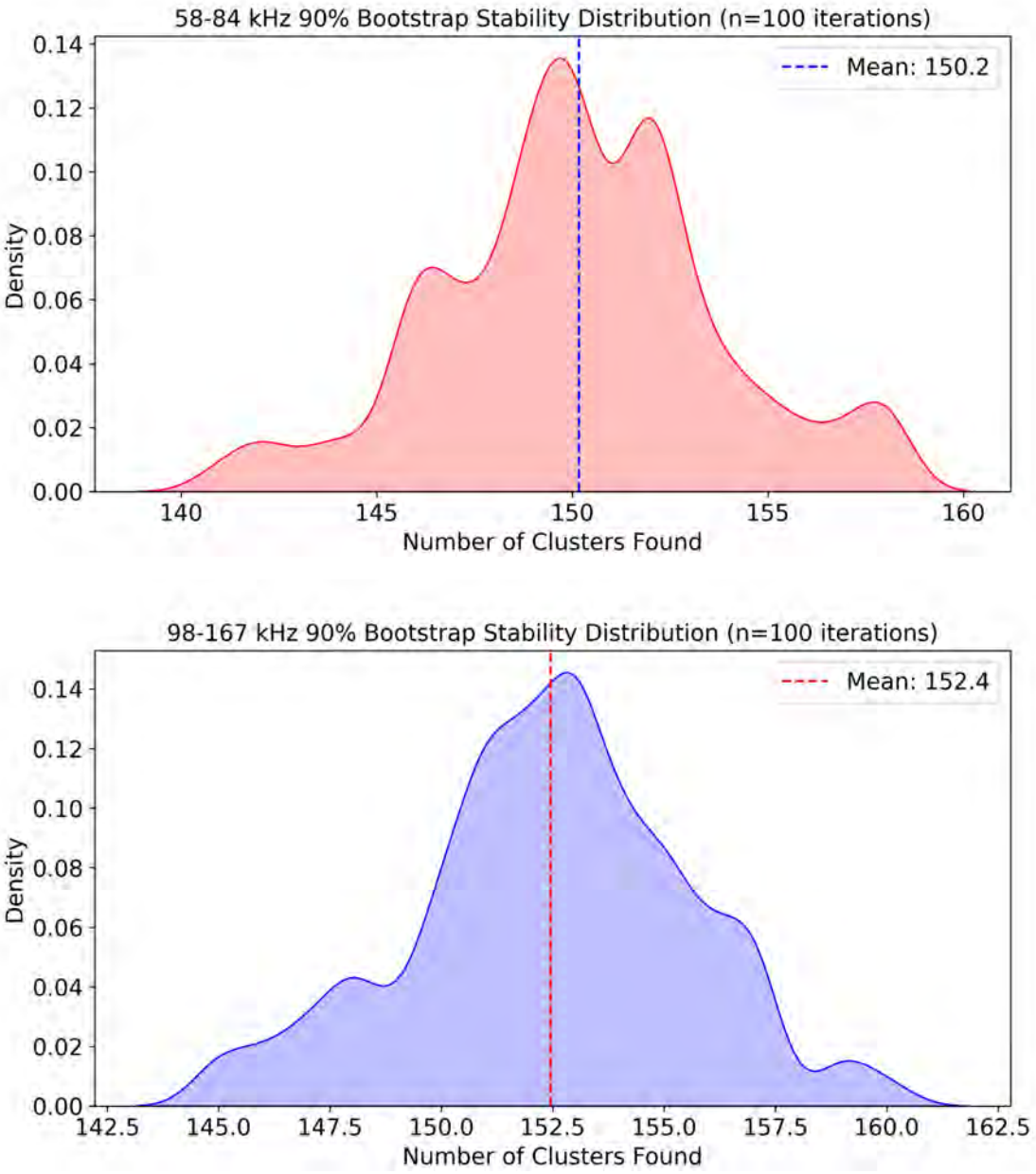


Figure 8.5. Density histograms of the number of clusters found in the 100 HAC iterations of non-parametric bootstrapping from 90% of broadband targets in the 58-84 kHz range (top) and 98-167 kHz range (bottom). Dashed vertical lines represent the mean number of clusters found across iterations of HAC on 90% of targets.

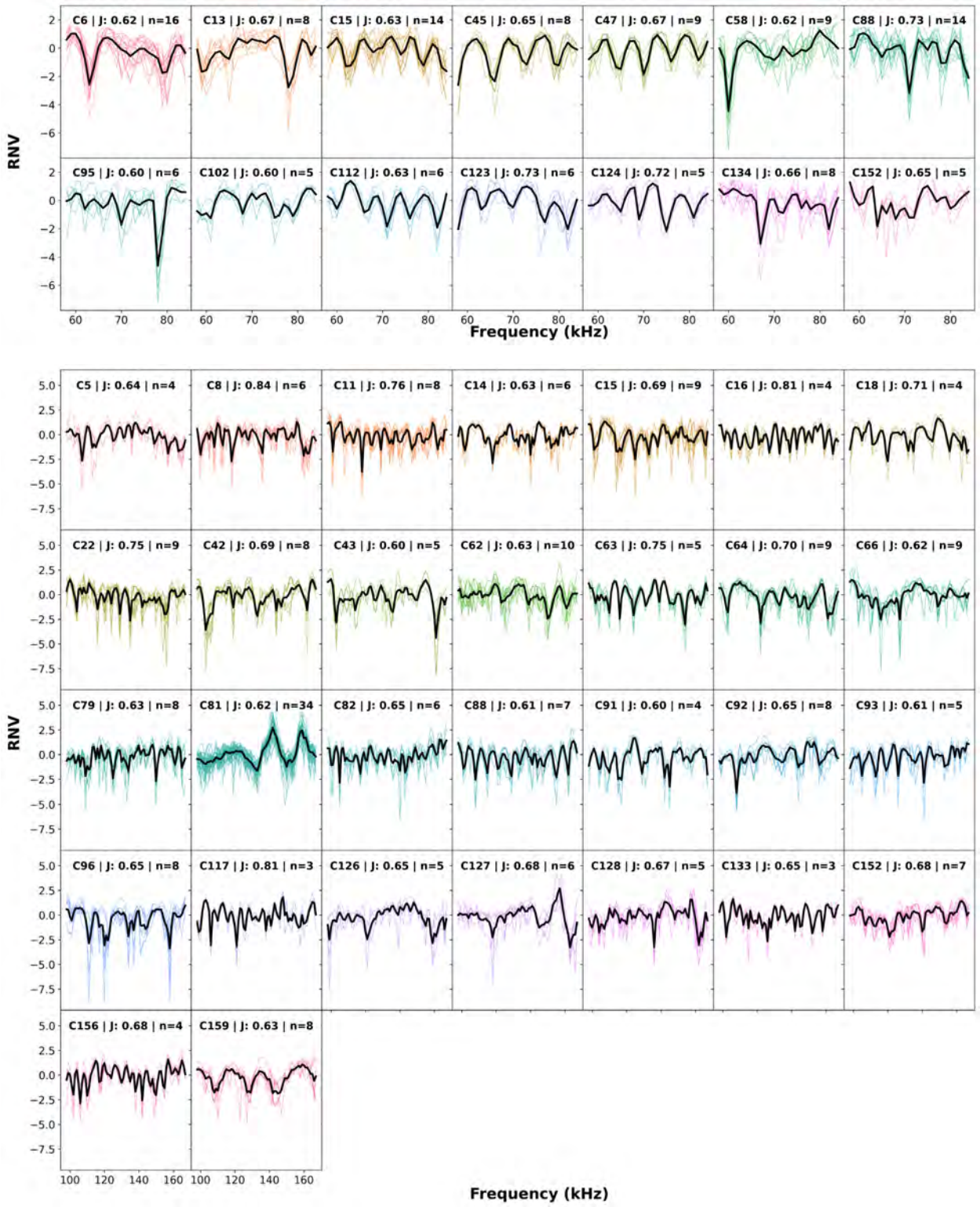


Figure 8.6. Stable clusters of coherent spectra with Jaccard coefficients ≥ 0.6 in the 58-84 kHz range (top 14 tiles) and 98-167 kHz range (bottom 30 tiles). Black solid lines indicate the cluster medians, and robust normal variate (RNV) target strength frequency responses are shown.

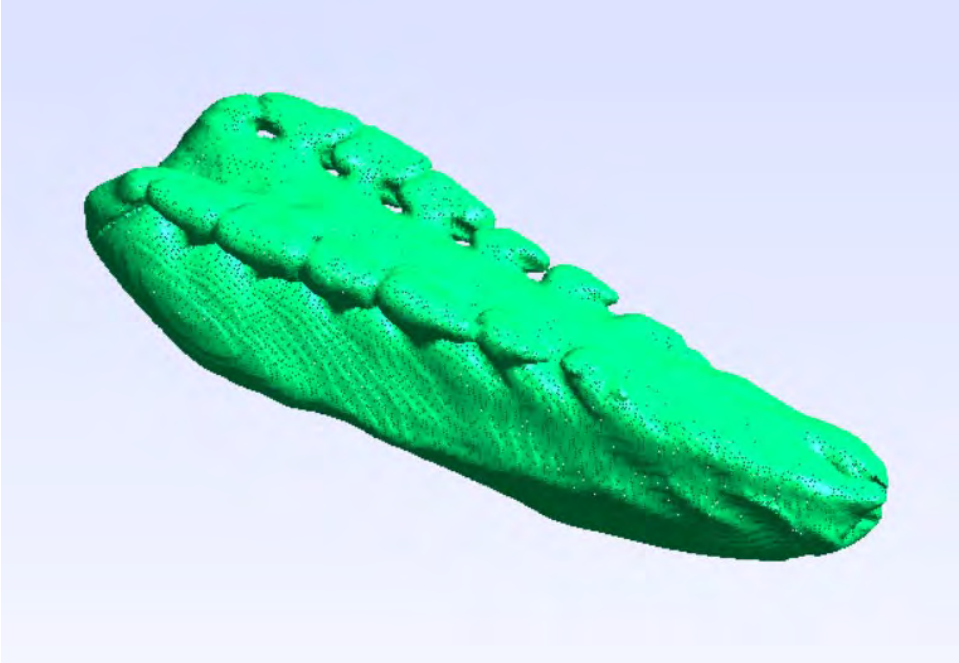


Figure 8.7. CT derived 3D bladder model of a Greater Amberjack (74 cm fork length) segmented from the body model.

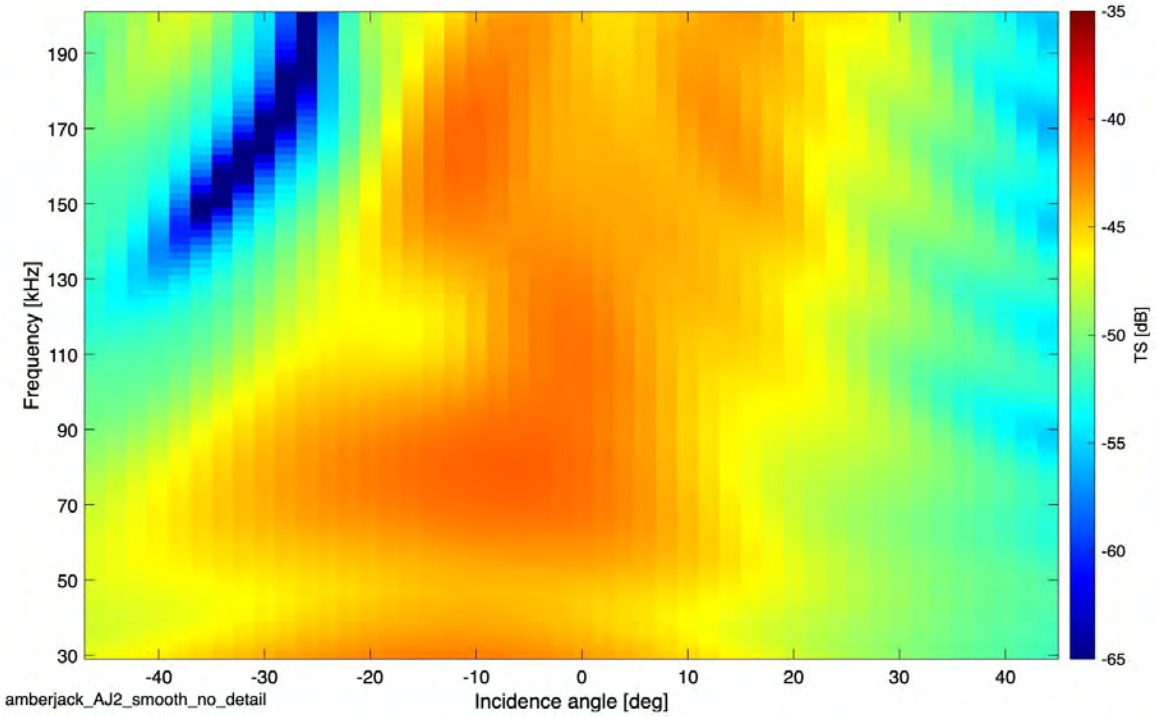


Figure 8.8. Example heatmap of modeled TS across the frequency spectrum 30-200 kHz and orientation (-44 : + 44 degrees) in the dorsal plane. Modeled Greater Amberjack is the same as in Figure 8.7.

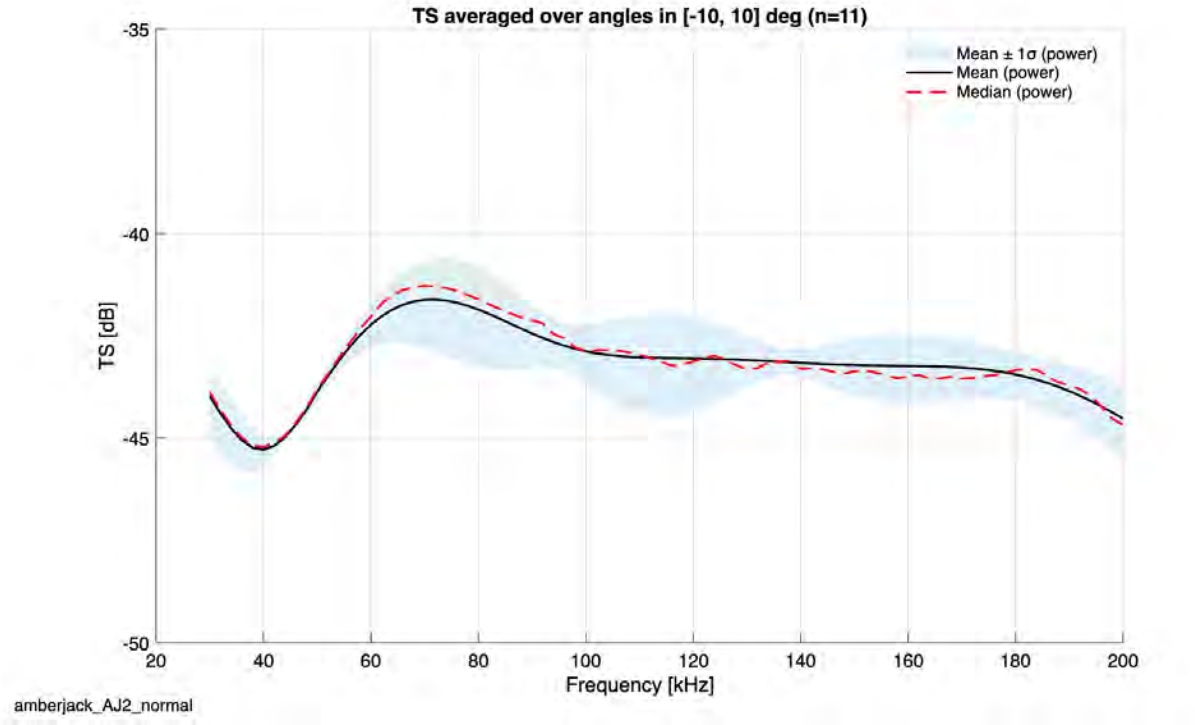


Figure 8.9. Average target strength response across frequency spectrum (30-200kHz) and averaged from -10:+10 degrees in the dorsal plane. Greater Amberjack modeled is the same as in Figure 8.7.

Literature Cited

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2623–2631.

Barange, M. (1994). Acoustic identification, classification and structure of biological patchiness on the edge of the Agulhas Bank and its relation to frontal features. South African Journal of Marine Science, 14: 333-347.

Bassett, C., De Robertis, A. and Wilson, C.D., 2018. Broadband echosounder measurements of the frequency response of fishes and euphausiids in the Gulf of Alaska. *ICES Journal of Marine Science*, *75*(3), pp.1131-1142.

Blackman, S. S. (1986). Multiple Target Tracking with Radar Applications. Artech House, Massachusetts.

Boswell, K.M., Pedersen, G., Taylor, J.C., LaBua, S. and Patterson III, W.F., 2020. Examining the relationship between morphological variation and modeled broadband scattering responses of reef-associated fishes from the Southeast United States. Fisheries Research, 228, p.105590.

Campello, R.J., Moulavi, D., Zimek, A. and Sander, J., 2015. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *10*(1), pp.1-51.

Chawarski, J., Coté, D., & Geoffroy, M. (2025). Unsupervised classification increases precision in acoustic broadband target strength measurements of mesopelagic fish. The Journal of the Acoustical Society of America, 158(5), 3994-4006.

Demer, D.A., Cutter, G.R., Renfree, J.S. and Butler, J.L., 2009. A statistical-spectral method for echo classification. *ICES Journal of Marine Science*, *66*(6), pp.1081-1090.

Demer, D.A., Berger, L., Bernasconi, M., Bethke, E., Boswell, K., Chu, D., Domokos, R., et al. (2015). Calibration of acoustic instruments. ICES Cooperative Research Report No. 326.

De Robertis, A., and Higginbottom, I. (2007). A post-processing technique for estimation of signal-to-noise ratio and removal of echosounder background noise. ICES Journal of Marine Science, 64, 1282-1291.

Fernandes, P.G., 2009. Classification trees for species identification of fish-school echotraces. *ICES Journal of Marine Science*, *66*(6), pp.1073-1080.

Gastauer, S., Scoulding, B., Fässler, S.M., Benden, D.P. and Parsons, M., 2016. Target strength estimates of red emperor (Lutjanus sebae) with Bayesian parameter calibration. Aquatic Living Resources, 29:301.

Horne, J.K., 2000. Acoustic approaches to remote species identification: a review. *Fisheries oceanography*, *9*(4), pp.356-371.

Jech, J.M., Lawson, G.L. and Lavery, A.C., 2017. Wideband (15–260 kHz) acoustic volume backscattering spectra of Northern krill (Meganyctiphanes norvegica) and butterfish (Peprilus triacanthus). *ICES Journal of Marine Science*, *74*(8), pp.2249-2261.

Khodabandeloo, B., M. D. Agersted, T. Klevjer, G. J. Macaulay, and W. Melle. 2021. Estimating target strength and physical characteristics of gas-bearing mesopelagic fish from wideband in situ echoes using a viscous-elastic scattering model. The Journal of the Acoustical Society of America 149(1):673-691.

Korneliussen, R.J., Heggelund, Y., Eliassen, I.K. and Johansen, G.O., 2009. Acoustic species identification of schooling fish. *ICES Journal of Marine Science*, *66*(6), pp.1111-1118.

Korneliussen, Rolf J. (Ed.). 2018. Acoustic target classification. ICES Cooperative Re-

search Report No. 344. 104 pp. http://doi.org/10.17895/ices.pub.4567.

Lekanda, A., Boyra, G. and Louzao, M., 2024. Weakly supervised classification of acoustic echo-traces in a multispecific pelagic environment. *ICES Journal of Marine Science*, *81*(7), pp.1247-1262.

MacLennan, D.N. (1990). Acoustical measurement of fish abundance. J. Acoust. Soc. Am. 87: 1-15.

Medwin, H., and Clay, C.S. (1998). Fundamentals of acoustical oceanography. Academic Press, Boston, MA.

Ona E. and Mitson R.B. (1996). Acoustic sampling and signal processing near the seabed: the deadzone revisited. ICES Journal of Marine Science 53: 677-690.

Parker-Stetter, S.L., Rudstam, L.G., Sullivan, P.J., and Warner, D.M. (2009). Standard operating procedures for fisheries acoustic surveys in the Great Lakes. Great Lakes Fisheries Commission Special Publication. 09-01.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Roa, C., Pedersen, G., Bollinger, M., Taylor, C. and Boswell, K.M., 2022. Taxonomical classification of reef fish with broadband backscattering models and machine learning approaches. *The Journal of the Acoustical Society of America*, *152*(2), pp.1020-1034.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65.

Ryan T. E., Downie R. A., Kloser R. J., and Keith G. (2015). Reducing bias due to noise and attenuation in open-ocean echo integration. ICES Journal of Marine Science 72, 2482-2493

Sawada, K., Furusawa, M., and Williamson, N.J. (1993). Conditions for the precise measurement of fish TS in situ. J. Mar. Acoust. Soc. Japan. 20: 73-79.

Scoulding, B., Gastauer, S., Taylor, J.C., Boswell, K.M., Fairclough, D.V., Jackson, G., Sullivan, P., Shertzer, K., Campanella, F., Bacheler, N. and Campbell, M., 2023. Estimating abundance of fish associated with structured habitats by combining acoustics and optics. *Journal of Applied Ecology*, *60*(7), pp.1274-1285.

Setiawan, W., I. Jaya, T. Hestirianoto, S. Pujiyati, A. Priatna, and A. Ma'mun. 2020. Comparison of wideband vs narrowband acoustic approach to target strength value on bullet tuna (Auxis rochei) and bonito (Euthynnus affinis). International Journal of Acoustics and Vibration 25(4):489-497.

Simmonds, J. and MacLennan, D.N., 2005. *Fisheries acoustics: theory and practice*. John Wiley & Sons.

Toresen, R. (1991). Absorption of acoustic energy in dense herring schools studied by the attenuation in the bottom echo signal. Fish. Res. 10: 317-327.

Urmy, S.S., De Robertis, A. and Bassett, C., 2024. A Bayesian inverse approach to identify and quantify organisms from fisheries acoustic data. *ICES Journal of Marine Science*, *81*(7), pp.1461-1477.

Verma, A., R. J. Kloser, and A. J. Duncan. 2017. Potential use of broadband acoustic methods for micronekton classification. Acoustics Australia 45(2):353-361.

White, A.L., Patterson III, W.F. and Boswell, K.M., 2022. Distribution of acoustic fish backscatter associated with natural and artificial reefs in the Northeastern Gulf of Mexico. *Fisheries Research*, *248*, p.106199.

White, A. L., Sullivan, P. J., Binder, B. M., Boswell, K. M. (2023). An evaluation of survey designs and model-based inferences of fish aggregations using active acoustics. Frontiers in Marine Science, 10, 1176696.

9. Gear Calibrations

9.1. Model-based calibration of survey gears

Calibration of gear-specific MaxN's from calibration experiment conducted in AL-MS, with supplemental data from AL-MS survey

Grace S. Chiu[1]

SUMMARY

Chapter 9.1 here focuses on two scenarios of the regional surveys: when a single camera type is deployed alone, or when a single camera type and the echosounder are deployed in tandem. The main part of Chapter 9.1 is adapted from the preprint by Chiu et al. (2025, revisions submitted), most of which describes the modeling framework and operational product developed based on the calibration experiment conducted in the AL-MS region between March 30 and April 25, 2024. The small calibration experiment involved 21 "boat trips," each of which simultaneously deployed 2, 3, or 4 of the drop camera, stereo baited remote underwater video (SBRUV) camera, trap camera, and remotely operated vehicle (ROV) camera. Each trip always deployed the acoustic echosounder. None of C-BASS (gear), platforms (habitat), and unconsolidated / uncharacterized bottoms (habitat) was part of the calibration experiment.

Our Bayesian hierarchical modeling framework addresses a conundrum: standard pairwise approaches for gear calibration requires one gear from each pair to serve as the "ground truth," yet when calibrating a set of 3+ gears against each other, there cannot be more than a single "ground truth." Statistically, this is a type of *change-of-support problem* (COSP) that is akin to reconciling multiple spatial datasets at various resolutions; though, the context here is to reconcile abundance datasets observed at various gear-specific scales.

In the case of a MaxN only (unpaired with acoustic count), our Bayesian framework is used to derive calibration formulae for translating camera-specific MaxN's to the actual abundance scale, accompanied by uncertainty estimates. Each of the 4 camera types has its own formula. Thus, in an operational setting, a formula takes as input a MaxN observed by the corresponding camera type that is unpaired with the echosounder. To examine the conversion formulae's performance, cross-validation is conducted using mark-recapture abundance estimates (only available for 10 trips, all observed at a single habitat type) and through a separate simulation study.

In the case of a MaxN paired with an acoustic count, a much simpler modeling approach, namely, multiple linear regression, is used to derive a deflation factor to scale down the acoustic count; the deflation factor is an estimated fraction (with standard error) of our target species, *Seriola dumerili* (henceforth, GAJ), among similarly large-bodied fish (henceforth, GAJ-like).

The conversion formula in both cases are implemented into a single web-browser-based app.

Chapter 9.1 ends with an addendum that describes the additional data that augment the 21 "boat trips" to a total of 45, using relevant data from AL-MS that were collected as part of this GAC project's regional surveys. These additional data were only finalized and made available for modeling in January 2026, many months after Chiu et al. (2025, revisions submitted) was made available. The aforementioned app is based on this augmented dataset. A major caveat of the calibration formulae in this app is the the likely lack of robustness across the

entire Gulf region, because they are constructed using a small dataset with limited habitat types across minimal spatial coverage within AL.

9.1.1. INTRODUCTION

With each camera type's unique field of view and associated video processing protocol, it produces a gear-specific MaxN that (1) does not reflect actual abundance, and (2) is only meaningful when compared within gear type but not between. Campbell et al. (2015) investigated the relationship between MaxN and mean actual abundance for a given camera gear, and Campbell et al. (2018), that between MaxN's from various camera gears. However, the purpose of these works was to explore ways in operational settings that can make better use of camera gears to reflect actual abundance. Instead, our goal is to translate camera-specific MaxN counts to the scale of actual abundance, which is the scale of acoustic counts. However, while the echosounder records acoustic signals of fish on the actual count scale, it is highly subject to confounding from acoustically similar species, thus requiring an externally derived correction factor. Conversely, a camera allows visual distinction between species but records data on a relative scale that is specific to the camera type and video processing protocol.

Chapter 9.1 here focuses on two scenarios of the GAC surveys: when a single camera type is deployed alone, or when a single camera type and the echosounder are deployed in tandem. Which camera type was deployed in a survey depends on the region and habitat being surveyed. To consolidate all regional surveys for the estimation of a study-wide absolute abundance (population size), all regional datasets must be reconciled to represent regional absolute abundances. Thus, we are faced with the following operational challenges:

How can we reconcile the inherent difference in representation between a camera-based MaxN (camera-specific relative abundance of GAJ) and an acoustic count (actual count of GAJ and GAJ-like individuals), given that each of the two counts is inherently related to the same true abundance in its unique way?

How do we ensure that the conversion from MaxN to actual abundance for any specific camera type is done in a unified manner that allows us to derive a meaningful study-wide total abundance based on regional conversions that involve different camera types?

In Chapter 9.1, we simultaneously address the challenges (1) and (2) whenever a camera is deployed, either alone or paired with an acoustic echosounder. This requires an understanding of the joint relationship among all gear-specific MaxN counts and the acoustic count. We develop our models on the data from the small-scale calibration experiment conducted in the AL-MS region, in which 21 separate "boat trips" surveyed 13 unique reef structures, each trip deploying the acoustic echosounder alongside two to all of four types of cameras: drop camera, SBRUV, trap camera, and ROV camera. By "boat trips," we refer to visits to reef structures. Thus, even if a single boat trip stopped at two different reef structures, we refer to them as "two boat trips." Note that "no structure" and "platform" were not involved in the calibration experiment. As the augmented calibration data (see Section 7) only include a single platform (and also two non-structure observations that are subsequently omitted from consideration), **any resulting conversion formula should be used to convert MaxN's at non-platform reef structures only**.

Also, very few boat trips in the experiment resulted in a 0 value for MaxN and/or acoustic counts (see Section 2). This suggests that, had a calibration tool been developed from the experimental data alone, it might not be

appropriate for converting 0 or small counts. Section 7 describes the augmented dataset that includes survey results with a non-trivial number of 0's and near-0's.

In the case of an unpaired camera, we develop a cross-calibration modeling framework through a single Bayesian hierarchical model that jointly relates all available MaxN data to acoustic data. The unifying calibration methodology produces a joint regression fit with confidence bands for each camera-specific MaxN. The calibrated value and associated uncertainty estimate produced as such simultaneously accounts for all camera types and the acoustic gear under a single overarching model.

Here, our methodological approach to make inference for an underlying "truth" is similar to those in Chiu et al. (2011), Chiu, Lehmann, et al. (2013), Chiu, Wu, et al. (2013), and Kuh et al. (2023, under revision), all of which exploit the hierarchical model structure to address a challenge that is akin to the change-of-support problem (COSP). COSP is traditionally viewed as a model-based approach to reconcile spatial data observed at multiple spatial scales (Gelfand 2010; Gotway Crawford and Young 2005). In this context, several sources of data — each reflecting the same underlying phenomenon of interest — are jointly utilized to make inference for this phenomenon. COSP then arises from spatial misalignment, whereby each data source has its unique spatial coordinate system, e.g., one dataset is recorded at the state level but another is recorded at the satellite pixel level, so that a single to multiple countries may partially appear in any given pixel; the COSP model stipulates a joint relationship among the various data sources and the unobservable "truth." Even if all data sources share the same spatial coordinate system, each data source may be a different proxy of the same variable of interest, so that none is a direct measurement of the underlying phenomenon. We consider the latter scenario as a wide-sense COSP, in the context of reconciling quantitative data that are qualitatively different. Both senses of COSP are present in Chiu, Lehmann, et al. (2013), but wide-sense COSP alone is addressed by Chiu et al. (2011), Chiu, Wu, et al. (2013), and Kuh et al. (2023, under revision). In Chapter 9.1 here, the wide-sense COSP refers to reconciling abundance data observed at various gear-specific scales, some being relative, and others, actual.

In the case of a camera paired with the echosounder, a much simpler modeling approach, namely, multiple linear regression, is used to derive a deflation factor to scale down the acoustic count; the deflation factor is a regression-based estimate of the fraction (with standard error) of GAJ among GAJ-like fish. Thus, the regression-based deflation factor is more robust than a factor derived based solely on the single camera type deployed in a regional survey.

The rest of Chapter 9.1 is structured as follows. Section 2 describes some key features of the data from the calibration experiment. Section 3 delves into our calibration modeling framework, associated rationale, assumptions, caveats, and results. Supplementary details for both sections appear in the Supplement. In Section 4, we develop a procedure that operationalizes the framework from Section 3 in the case of a camera that is deployed alone (unpaired with the echosounder) in a field survey; the procedure converts the camera-based MaxN into an actual GAJ abundance estimate, alongside uncertainty estimates. As the dataset from our calibration experiment is very small, the model used to develop the conversion from MaxN to actual abundance is substantially simpler than the general framework. To illustrate the potential utility of the general framework in data rich calibration studies, we present a simulation study in Section 5. In Section 6, we describe our multiple linear regression approach to handle the case in which both a camera and the echosounder are deployed simultaneously during a survey — instead of modeling GAJ abundance, the approach in this case derives a rather robust adjustment factor for the acoustic count. The procedures from Section 4 and Section 6 are implemented as a single app, which is described in these sections. Section 7 describes the supplemental data made available

long after the calibration modeling framework had been developed, and how they have been used to augment the calibration experimental data to update the conversion formulae discussed in Section 2 to Section 6.

9.1.2. DATA FROM THE CALIBRATION EXPERIMENT

Figure 9.1.1 shows the locations of all 13 unique reef structures (east of New Orleans, Louisiana and south of Mobile, Alabama) visited in the calibration experiment. Some columns of the data table appear in Table 9.1.1.

A total of 13 unique reef structures were visited by 21 separate boat trips; different structures visited by the same boat on the same day constitute separate trips; similarly, for the same structure visited on different dates, e.g., the artificial super pyramid structure `Spmid 18-030` generated five trips, and similarly for the artificial super pyramid `Spmid 18-061`.

Each sth trip is associated with key abundance-related variables that include

- $y_{s\ell}$ = video-derived MaxN for GAJ based on the ℓth camera type — thus, $y_{s\ell}$ is missing (denoted by `NA`) if type ℓ camera was missing; all 21 trips involved both SBRUV and trap cameras, 2 trips lacked the drop camera only, 5 lacked the ROV camera only, and 2 lacked both the drop and ROV cameras;
    - for $s = 12$: $y_{sD} = 0$ and $y_{sS}, y_{sT} > 0$ and $y_{sR} =$ `NA`,
    - for $s = 15$: $y_{sD} = y_{sS} = 0$ and $y_{sT} > 0$ and $y_{sR} =$ `NA`,
    - for $s = 16$: $y_{sD} = y_{sT} = y_{sR} = 0$ and $y_{sS} > 0$,
    - for $s = 17$: $y_{sD} = y_{sS} = y_{sT} = y_{sR} = 0$;
- $N_s = \sum_t N_{st}$ (never `NA`) is the total count of GAJ and GAJ-like fish (collectively denoted by "GAJ+") derived from acoustic signals along all transects combined, i.e., the sum of N_{st} across all tth transects traversed by the echosounder;
    - $N_s = 0$ for $s = 12,14,17$;
- $N_s^{(f)}$ = total acoustic-derived count of GAJ+ (never `NA`) pooled from focal transects only, i.e., transects that are most relevant to GAJ activities around the sth reef structure — thus, $N_s^{(f)} \leq N_s$;
- $N_s^{(mr)}$ = mark-recapture estimate of actual GAJ abundance, computed using the MARK software (https://sites.warnercnr.colostate.edu/gwhite/program-mark) (White and Burnham 1999) while assuming an unidentified marks immigration-emigration mark-resight logit-normal model (McClintock 2026) under a standard mark-recapture protocol that was administered at the structures `Spmid 18-030` and `Spmid 18-061` only (thus, $N_s^{(mr)}$ is `NA` for 11 trips, and actual observations are always at the same type of habitat, being a "super pyramid"); and scaling factor

$$r_s = \frac{\sum_{\{\ell:\ell \text{ present}\}} y_{s\ell}}{\sum_{\{\ell:\ell \text{ present}\}} \{(s,\ell)\text{th MaxN for all large-bodied species in the water column}\}} + 10^{-6}$$

(see Table 9.1.2 for the list of species included in the denominator of r_s) — the shift factor of 10^{-6} avoids ln(0) in the calibration model.

In addition, each sth trip is associated with variables of a potentially explanatory nature to abundance, including

- reef origin — of the 13 structures, 2 are natural and the rest are artificial;
- reef type — 6 types (see Table 9.1.1), of which 3 (rock outcrop, natural; large rocks, natural; and chicken coop, artificial) visited over 4 trips are considered small structures, and the other 3 types visited over 17 trips are considered large structures;
- boat ID — 2 boats, one of which never carried an ROV camera; and
- date-boat combination — 14 combinations across 21 trips.

The above variable descriptions reveal that the calibration experiment exhibited an extremely unbalanced experimental design with very few data points. This lopsidedness leads to substantial modeling challenges: missing data due to `NA` under many treatment combinations would cause many explanatory effects to be inestimable. To preemptively mitigate inestimability challenges, we assume that

- reef size (large or small under reef type) encapsulates any effects on abundance due to reef origin and reef type, and
- trip date is irrelevant to abundance.

Thus, our unifying model can only include a maximum of two explanatory variables (boat and reef size) and possibly their interaction with camera type. Even with a reduced set of candidate explanatory variables, severe imbalance remains in the experimental design, e.g., $\ell = R$ appeared on only one of the boats, and only 4 of 21 trips involved small reef structures. The imbalance is exacerbated when $N_s^{(mr)}$ is included in the model for exploration/validation purposes, as mark-recapture data are associated with only two large structures over 10 trips; this results in painstaking adjustments being required each time a potential explanatory variable is omitted from or added to the model (see Supplement for details).

9.1.2.1 VISUALIZING THE EXPERIMENTAL DATA

Species count data are best visualized on the log scale due to substantially skewed distributions.

Recall that raw acoustic counts inherently reflect more than one species, whereas mark-recapture estimates and camera-based MaxN counts refer to a single target species. As such, poor alignment between unscaled acoustic counts and the other two count types is to be expected. Indeed, as can be seen from the top panels in Figure 9.1.2, unscaled acoustic counts (N_s or $N_s^{(f)}$) are somewhat negatively correlated with mark-recapture estimates, but the bottom panels show that the correlation is noticeably positive when the acoustic counts have been scaled by the pooled camera ratios (r_s). Similarly, from the top panels of both Figure 9.1.3 and Figure 9.1.4, we see that MaxN counts ($y_{s\ell}$) and unscaled acoustic counts (N_s or $N_s^{(f)}$) generally do not correlate if the smallest counts (all of which originate from small reef structures) are omitted; the bottom panels show that the correlation improves noticeably when the acoustic counts have been scaled. All of Figure 9.1.2, Figure 9.1.3, and Figure 9.1.4 indicate that focal acoustic counts ($N_s^{(f)}$) have weaker correlations with MaxN counts and mark-recapture estimates when compared to acoustic counts pooled across all transects (N_s).

With respect to camera ratios, Figure 9.1.5 indicates that camera-specific ratios are weakly correlated among the four camera types. This poor alignment is also unsurprising, given each camera type's unique field of view,

image capturing protocol, and image processing protocol. In contrast, the pooled ratio, r_s, correlates reasonably well with each of the camera-specific ratios, providing justification for our pooling method.

9.1.3. CALIBRATION MODELING FRAMEWORK

Here in Section 3, the key objective of our modeling framework is to facilitate the conversion of the relative abundance metric, MaxN for GAJ (whose value differs across camera types), to an actual count of GAJ individuals — the conversion is to be informed by the acoustic count of GAJ+ individuals.

If the entire field survey involves MaxN's from a single camera type (paired with acoustic data or otherwise), then a traditional calibration approach may be to run a calibration experiment that pairs the camera with the echosounder, fit a simple regression of acoustic count on MaxN to these data, then use the regression curve as the calibration formula to convert any future MaxN from the same camera type to an actual abundance estimate; the confidence bands around the regression curve would provide an estimate of the calibration uncertainty. However, in our study, (i) a MaxN reflects the species of interest but an acoustic count reflects multiple species combined, and (ii) each region in the continental-scale field survey deploys its own camera type.

Thus, our key objective is to construct a unifying statistical modeling framework that reflects the single, joint relationship among all camera-specific MaxN counts of GAJ (y_{Ds}, y_{Ss}, y_{Ts}, and y_{Rs}) and the acoustic count of GAJ+ (either N_s or $N_s^{(f)}$). Such a cross-calibration modeling framework replaces a series of simple regression fits; we construct a single Bayesian hierarchical model that jointly relates the vector $\mathbf{y}_s$ to N_s or $N_s^{(f)}$ observed in the calibration experiment. The framework allows boat and reef size as predictor variables, while the pooled camera ratio r_s enters the model as the offset term in the Poisson log-link function for the true mean GAJ abundance.

The calibration model under this framework is trained by data from the 21 boat trips in the calibration experiment discussed in Section 2, in which all trips involved the acoustic gear and many trips involved all four camera types.

9.1.3.1 KEY ASSUMPTIONS

Our methodology assumes that

- the scaling factor r_s is a reasonable approximation of the true ratio between the number of GAJ individuals and the number of all acoustically indistinguishable individuals,
- our calibration experiment — involving 13 reef structures visited by a total of 21 boat trips combined — provides a reasonable basis for training the joint relationship between $\mathbf{y}_s = [y_{sD}, y_{sS}, y_{sR}, y_{sT}]'$ and N_s or $N_s^{(f)}$ across the study-wide region of interest, so that
- regional absolute abundance estimates, computed based on the calibration modeling methodology described here in Chapter 9.1, are justifiably additive, i.e., can be totaled into a study-wide estimate of GAJ absolute abundance.

9.1.3.1.1 Caveats

We acknowledge that MaxN counts are not necessarily additive, so that r_s as defined above may be an unreliable approximation of the ratio between GAJ count and GAJ+ count, although an improved definition of r_s would require video processing that is deemed impractical given the scope of this project. Notwithstanding,

cross-validation against mark-recapture abundance estimates suggests that r_s as defined is a reasonable correction factor for acoustic data (see Section 9.1.3.4).

Secondly, although we use a single list of species deemed to be "large-bodied" across the calibration experiment (see Table 9.1.2), the MaxN counts for any given "large-bodied" species may include individual fish that is smaller in size than those recognized by the algorithm that converts echosounder signals to a fish count.

Thirdly, the calibration experiment was confined to a tiny, isolated swath inside the study-wide spatial domain, with possibly limited representation of the regional conditions and habitat features across the study; however, challenges of this nature were unavoidable due to the lack of existing pilot calibration studies prior to the current continental-scale project with very ambitious objectives.

Finally, neither N_s nor $N_s^{(f)}$ is the ideal acoustic data for calibration purposes. Instead, a value (e.g., from interpolation) that reflects the type of structure (i.e., point-level or areal) would be necessary for all resulting acoustic counts to share a common areal/voluminal denominator (see Section 8.6), although such an undertaking for all data from the full-scale study will require ongoing data processing beyond the scope of Chapter 9.1.

9.1.3.2 MOST COMPREHENSIVE FRAMEWORK

Our most comprehensive — hence, ambitious — Bayesian hierarchical mixed-effects framework jointly models camera-specific MaxN counts ($y_{s\ell}$ where $\ell = D, S, T, R$), acoustic counts (N_s), and mark-recapture estimates ($N_s^{(mr)}$), while accounting for boat (two types), reef size (small or large), and their possible interactions with camera type (four types). With data from only 21 boat trips at the time of model development — each of which has at least one missing value for $y_{s\ell}$ and $N_s^{(mr)}$ due to missing {boat, reef, camera type} combinations — the model is expected to be inestimable unless upfront constraints are stipulated. Thus, our framework upfront disregards the interaction between boat and reef size, but it allows the interaction between camera type with each of boat and reef size. It also allows reef-size-specific regression error variances within the model hierarchy. Slope parameters are constrained to be non-negative, and fixed effects must sum to zero. The count variables, namely, $y_{s\ell}, N_s$, and $N_s^{(mr)}$, are all modeled as conditional Poisson random variables, whose conditional means are log-normally distributed random quantities. The four MaxN counts are explicitly expressed as multivariate in the model hierarchy.

Instead of indexing the boat trips by $s = 1, 2, \ldots, 21$, we reindex the boat trips as boat-reef-specific: trip k_{ij} is the kth trip for the (i, j)th boat-reef combination, where $i, j = 1,2$, and $k_{11} = 1,2,\ldots,13$; $k_{12}, k_{22} = 1,2$; and $k_{21} = 1,2,3,4$.

Next, let μ denote expected MaxN counts, τ denote expected acoustic and mark-recapture counts, ϕ denote the underlying true expected GAJ abundance, β_0 and β_1 denote regression intercept and slope parameters, respectively; ν, γ, and ξ denote fixed effects for boat, reef size, and actual abundance count type, respectively; ε and ζ denote linear regression noise terms, σ^2 denote regression error variances, and ρ (constrained to be non-negative) denote the inter-MaxN correlation in an exchangeable correlation structure; interactions are denoted by index combinations. Note that all μ's and τ's are modeled as functions of ϕ; thus, the underlying true expected GAJ abundance, ϕ, is the single model parameter that unifies the relationship among all MaxN counts, acoustic counts, and mark-recapture estimates. Then, our comprehensive hierarchical model is

Level 1.1: $y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell})$

Level 1.2: for $h = 1$ (acoustic): $N_{ijk} \sim \text{Pois}(\tau_{hijk})$ for all i, j, k

for $h = 2$ (mark-recapture): $N_{i1k}^{(mr)} \sim \text{Pois}(\tau_{hi1k})$ for all i, k ($j = 1$ only)

Level 2.1: for $\ell = D, S, T$:

$$\ln\mu_{ijk\ell} = \beta_{y,0,\ell} + \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\widetilde{\ln\phi}_{ijk} + \varepsilon_{ijk\ell}$$

$$\begin{bmatrix}\varepsilon_{ijkD}\\ \varepsilon_{ijkS}\\ \varepsilon_{ijkT}\end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2\mathbb{P})$$

$$\mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \begin{cases}\rho & \text{if } \ell_1 \neq \ell_2\\ 1 & \text{if } \ell_1 = \ell_2\end{cases}$$

for $\ell = R$:

$$\ln\mu_{ijkR} = \beta_{y,0,R} + \gamma_{y,jR} + \beta_{1,jR}\widetilde{\ln\phi}_{ijk} + \beta_{1,D}\widetilde{\ln\mu}_{ijkD} + \beta_{1,S}\widetilde{\ln\mu}_{ijkS} + \beta_{1,T}\widetilde{\ln\mu}_{ijkT} + \varepsilon_{ijkR}$$

$$\varepsilon_{ijkR} \sim \mathcal{N}(0, \sigma_{yR}^2)$$

for all ℓ: $0 = \sum_i \nu_{y,i\ell} = \sum_j \gamma_{y,j\ell}$ for identifiability

Level 2.2: for $h = 1$: $\ln(r_{ijk}\tau_{hijk}) = \ln\phi_{ijk} + \xi_1 + \zeta_{hijk}$

for $h = 2$: $\ln\tau_{hi1k} = \ln\phi_{i1k} + \xi_2 + \zeta_{hi1k}$ ($j = 1$ only)

for $h = 1,2$: $\zeta_{hi1k} \sim \mathcal{N}(0, \sigma_{x,1}^2)$

for $h = 1,2$: $\zeta_{hi2k} = 0$ for identifiability due to $k_{i2} \leq 2$

$$0 = \sum_h \xi_h \quad \text{for identifiability}$$

Level 3:

$$\ln\phi_{ijk} = \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk}$$

$$\delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2)$$

$$0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}$$

where, for computational stability, $\widetilde{\ln\mu}_{ijk\ell}$ is taken to be either $\ln\mu_{ijk\ell}$ (uncentered) or $\ln\mu_{ijk\ell} - (1/21)\sum_{i,j,k}\ln\mu_{ijk\ell}$ (centered), and similarly for $\widetilde{\ln\phi}_{ijk}$. We employ relatively diffuse prior distributions

$$\left.\begin{array}{cc}\text{logit}\,\rho & \sim \mathcal{N}(0, 2^2)\\ \ln\sigma & \sim \mathcal{N}(0, 2^2) \text{ for all } \sigma\text{'s}\\ \ln\beta_1 & \sim \mathcal{N}(0, 2^2) \text{ or } \mathcal{N}(0, 3^2) \text{ for all } \beta_1\text{'s}\\ \beta_0, \gamma, \nu, \xi & \sim \mathcal{N}(0, 3^2) \text{ for all } \gamma\text{'s}, \nu\text{'s}\end{array}\right\} \quad (1)$$

Note that in Level 2.1 of the hierarchy, the quadvariate relationship is broken down into a trivariate normal for drop-, SBRUV-, and trap-based MaxN regression noise terms, and a univariate log-linear regression of the mean ROV-based MaxN on the other three log-transformed MaxN means and reef size effect. This breakdown reflects the fact that ROV cameras were never deployed on one of the two boats.

The comprehensive model of this subsection has the following features:

- The true expected abundance (ϕ) can depend on boat (ν) and reef size (γ), and the regression error variance (σ_ϕ^2) can be reef-specific (i.e., indexed by j).
- On the log-mean scale, both the expected acoustic and expected mark-recapture counts (τ)
    - are equal to the true expected abundance (ϕ), but possibly subject to a count-type-specific offset (ξ) and noise (ζ); and
    - have a regression error variance (σ_x^2) that can be reef-size-specific (i.e., indexed by j).[1]
- On the log-mean scale, the expected value of MaxN counts (μ) that are observed at both reef sizes ($\ell = D, S, T$) can depend on the true expected abundance (ϕ), boat (ν), and reef size (γ). Moreover,
    - any dependence on boat and reef size at the MaxN-level is in addition to the dependence on boat and reef size at the ϕ-level;
    - the effect of ϕ on μ (i.e., slope parameter β_1) can be reef-size-specific (i.e., indexed by j); and
    - the regression noise terms (ε) follow a multivariate normal distribution with mean 0 and exchangeable correlation structure (i.e., all pairwise correlations are ρ) with a common error variance (σ_y^2) that is not specific to boat or reef size.
- On the log-mean scale, the expected value of ROV MaxN counts that are observed on one boat only ($\mu_{\ell=R}$)
    - can depend on the other expected MaxN counts ($\mu_{\ell\neq R}$) and reef size (γ), but not on boat (ν);
    - the effect of ϕ on $\mu_{\ell=R}$ (i.e., slope parameter β_{1R}) can be reef-size-specific (i.e., indexed by j);
    - the effects of other expected MaxN counts ($\mu_{\ell\neq R}$) on $\mu_{\ell=R}$ (i.e., slope parameters $\beta_{1D}, \beta_{1S}, \beta_{1T}$) are constant across reef sizes; and
    - the regression noise term ($\varepsilon_{\ell=R}$) does not share the same error variance (σ_y^2) as those for the other expected MaxN counts ($\varepsilon_{\ell\neq R}$).

While the complexity of the comprehensive framework presented above may be impractical for the small dataset from our calibration experiment, we present a simulation study in Section 5 that illustrates the potential utility of our calibration model in scenarios that are more data rich.

9.1.3.3 SIMPLER MODELS

For the 21 observations from our calibration experiment, we consider simpler versions of the general model in Section 3.2. They include a combination of the following model reductions:

[1] Note the constraint $\sigma_{x,j=2}^2 = 0$ for expected counts at small reefs (model Level 2.2), which is equivalent to $\zeta_{j=2} = 0$.

- Omit $h = 2$ from Level 1.2, thus collapsing Levels 2.2 and 3 into a single level and omitting any inestimable quantities that pertain to the presence of two sets of actual abundance data.
- Omit r from consideration, thus, ϕ is modeled as the expectation of the raw acoustic count N in the absence of $h = 2$, and of N with an offset ξ in the presence of $h = 2$.
- Omit v's in Level 2.1, thus collapsing the $\ell = R$ and $\ell \neq R$ cases into a single log-linear regression of μ on ϕ and reef size effect, with a quadvariate normal distribution for the regression error terms. Omit any of $\{\beta_{y,0}\}, \{v_x\}, \{v_y\}, \{\gamma_x\}$, and $\{\gamma_y\}$.
- Taking $\beta_{1,j\ell_1} = \beta_{1,j\ell_2}$ for some ℓ_1, ℓ_2.

Each model version considered (see Supplement) is implemented using the NIMBLE Bayesian programming software for R (Ponisio et al. 2020; de Valpine et al. 2024), which produces Markov chain Monte Carlo (MCMC) simulated draws from the posterior distribution.

We consider a model to be adequate if all of the key criteria below are met:

- Apparent convergence of the model, based on visual inspection of MCMC trace plots.
- The posterior distribution suggests that all fixed-effects and slope parameters are credibly non-zero, and that ρ is credibly bounded away from 0 and 1.
- The posterior distribution suggests a credibly close alignment between ϕ (true expected GAJ abundance) and
    - the camera-ratio-adjusted acoustic counts (i.e., $\phi_{ijk} \approx r_{ijk}N_{ijk}$ with high credibility) if r_{ijk} is part of the model; or
    - the raw acoustic counts (i.e., $\phi_{ijk} \approx N_{ijk}$ with high credibility) if r_{ijk} is omitted from the model.
- Cross-validation against mark-recapture estimates: the posterior distribution suggests a credibly positive correlation between ϕ and $N^{(mr)}$ — as $N^{(mr)}$ is an abundance estimate that is only available at one type of habitat ("super pyramid") and thus may be inherently biased, we do not require $\phi \approx N^{(mr)}$ with high credibility so as to declare model adequacy.
- The posterior distribution is noticeably different from the prior distribution, indicating that non-trivial Bayesian updating has taken place.

Many versions of our modeling framework exhibit poor model performance (see Supplement). For example, several exploratory versions that omit the camera ratios r from the model yield very poor alignment between ϕ and N on both the count- and log-scales. That is, $\phi \approx c_0 + c_1 N$ where c_0 and c_1 are credibly very different from 0 and 1, respectively; and the same is true when ϕ and N are replaced with $\ln\phi$ and $\ln(N + 1)$.

9.1.3.4 Final model adopted for producing calibration formulas based on experimental data only

Among the model versions we have implemented for these experimental data, the least complex yet still adequate version is

$$
\begin{aligned}
\text{Level 1:} \quad & y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell}) \\
\text{Level 2.1:} \quad & N_{ijk} \sim \text{Pois}(\phi_{ijk}/r_{ijk}) \\
\text{Level 2.2:} \quad & \ln\mu_{ijk\ell} = \beta_{y,0,\ell} + \beta_{1,j\ell}\widetilde{\ln\phi}_{ijk} + \varepsilon_{ijk\ell} \\
& \begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
& \mathbb{P}_{\ell_1,\ell_2} = \begin{cases} \rho & \text{if } \ell_1 \neq \ell_2 \\ 1 & \text{if } \ell_1 = \ell_2 \end{cases} \quad \text{for all } \ell_1, \ell_2 = D, S, T, R \\
\text{Level 3:} \quad & \ln\phi_{ijk} = \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
& \delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2) \\
& 0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

with $\widetilde{\ln\phi}_{ijk}$ being centered and prior distributions as given in Equation 9.1.1. Note that first fitting the model variant with respective intercepts $\beta_{y,0,\ell} + \nu_{y,i\ell} + \gamma_{y,j\ell}$ in Level 2.2 and β_0 in Level 3 suggests that $\nu_{y,i\ell}, \gamma_{y,j\ell}$, and β_0 are unnecessary.

Again, model adequacy is assessed according to the criteria given in Section 9.1.3.3. For example, Figure 9.1.6 and Figure 9.1.7 show non-trivial Bayesian updating and the desired alignment between ϕ and each of the mark-recapture count and the adjusted acoustic count. In particular, Figure 9.1.6 shows that the marginal posterior distributions for population-level model parameters are noticeably different from the corresponding prior distributions, with the unsurprising exceptions of $\beta_{1,j=2,\ell=D}$ (left column, 7th panel) and $\beta_{1,j=2,\ell=R}$ (right column, 8th panel) that correspond to some issues with drop and ROV camera data: at the four small reef structures, the pair $(y_{\ell=D}, y_{\ell=R})$ takes on the values (0, 0), (0, NA), (NA, 8), and (NA, NA), respectively, whereas $y_{\ell=S}$ and $y_{\ell=T}$ each takes on a variety of values across these four reefs. Figure 9.1.7 indicates that, *a posteriori*, the correlation between ϕ and $N^{(mr)}$ is high, and the regression line of ϕ on rN has intercept $c_0 \approx 0$ and slope $c_1 \approx 1$. Note that Figure 9.1.7 is produced as follows.

Upon fitting the final model, each Mth draw from the posterior distribution is a vector of all model parameter values, namely, $\{\mu_{ijk}^{(M)}, \phi_{ijk}^{(M)}, \beta_0^{(M)}, \beta_{1,j\ell}^{(M)}, \sigma_{\phi,j}^{(M)}, \sigma_y^{(M)}, \rho^{(M)}$: for all $i, j, k, \ell\}$. Let us once again index a boat trip by $s = 1, 2, \ldots, 21$. Thus, for each M, frequentist least-squares regression can be applied to regress $\{\phi_s^{(M)}: s = 1, \ldots, 21\}$ on each of $\{N_s^{(mr)}: s = 1, \ldots, 10\}$ and $\{r_s N_s: s = 1, \ldots, 21\}$ while prohibiting a negative slope; each Mth regression line has intercept $c_0^{(M)} \in (-\infty, \infty)$ and slope $c_1^{(M)} \geq 0$. The median among the $c_0^{(M)}$'s and that among the $c_1^{(M)}$'s together yield the white "median regression line" in the panels of Figure 9.1.7.

9.1.3.4.1 Caveat

The use of this final model for converting MaxN counts to expected actual abundance is possibly handicapped by the lack of data (21 boat trips only) and of spatial variability used to train the calibration approach.

9.1.4. OPERATIONALIZING OUR CALIBRATION MODEL FOR UNPAIRED CAMERA DEPLOYMENTS

Again, any given camera type was sometimes deployed in a regional survey without being paired with an echosounder. For example, given a trap camera MaxN from a survey, we wish to couple it with the final model in

Section 3.4 to produce an estimate of the expected GAJ abundance, with uncertainty. A fully rigorous approach may be to refit the final model to an augmented dataset, one that includes the original 21 (i, j, k) combinations of $\{y_{ijkD}, y_{ijkS}, y_{ijkT}, y_{ijkR}, N_{ijk}, r_{ijk}\}$, with an additional 22nd that is {NA, NA, y_T^*, NA, NA, r^*}, where a '$*$' denotes the survey observation. However, this fully rigorous approach can be impractical, for two reasons: 1) the indexing of (i, j) for the calibration experiment does not apply to the survey (with i denoting boat type that was specific to the experiment), so that the 22nd boat trip is incompatible with the final model; and 2) even if $\nu_{x,i}$ were set at 0 in the final model (i.e., boat type were assumed to not affect actual abundance) to restore compatibility, model refitting is highly computationally intensive if one were to use the calibration tool during a field survey.

A more practical alternative — especially for use in the field — is to derive a set of calibration formulae that take on a camera-specific MaxN as input and provide an estimated expected GAJ abundance as output. To this end, we conduct least-squares regression similar to that described in Section 3.4, as follows.

For each M and each $\ell(= D, S, T, R)$, the least-squares regression line of $\{\phi_s^{(M)}: s = 1, \ldots, 21\}$ on $\{y_{s\ell}: s = 1, \ldots, 21$ and $y_{s\ell}$ not missing$\}$ has intercept $b_{0,\ell}^{(M)}$ and slope $b_{1,\ell}^{(M)}$. Thus, the corresponding linear formula can take an ℓth camera-specific MaxN count as input and map it to the Mth estimated expected actual abundance of GAJ as output. Next, fixing ℓ, we gather $(b_{0,\ell}^{(M)}, b_{1,\ell}^{(M)})$ across all values of M to form the joint posterior distribution of the *calibration intercept and slope* for the ℓth camera type. The posterior median intercept and posterior median slope can be used to produce *the* calibration formula for converting an ℓth MaxN count to an actual abundance estimate.

Figure 9.1.8 shows screen shots of the interactive R Shiny app (Chang et al. 2025) that we have created to display posterior summaries and visualizations of the calibration formulae for cases in which a camera is deployed alone (without being accompanied by the echosounder). The app user chooses the camera type to see the posterior median, mean, and variance for each of b_0 and b_1 from least-squares regression of posterior draws of $\boldsymbol{\phi}$ on this camera type's MaxN data in the calibration experiment; additionally, the posterior correlation between b_0 and b_1 is shown. The app user can use these posterior statistics to manually convert a regional survey's MaxN count to an estimated expected GAJ abundance and compute an approximate standard error (SE). Alternatively, the widget displays an interactive plot of 5000 least-squares lines (blue) and the "posterior median" least-squares line (white); the app user can click on the white line at a particular MaxN value to see the corresponding estimated value of expected GAJ abundance; hovering the mouse over any part of the plot gives the coordinates of the mouse, which facilitates the visual evaluation of the approximate SE.

Note that the notion of this SE corresponds to the uncertainty in the $\boldsymbol{\phi}$ vector associated with the final model in Section 9.1.3.4. Specifically, each posterior draw of $\boldsymbol{\phi}$ from the final model yields a different regression line (in blue in Figure 9.1.8). Thus, this SE does not reflect the goodness-of-fit of the calibration formula (white line in Figure 9.1.8). For example, the posterior median R^2 value for the ϕ-on-y_T least-squares regression is 0.16, reflecting a poor goodness-of-fit, and hence, a possibly large *calibration error*. To quantify this calibration error with more rigor, the widget displays the posterior median and 80% credible interval for the quantity $\epsilon_{sT} = \phi_s - \hat{\phi}_{sT}$ for each boat trip s, where $\hat{\phi}_{sT}$ is the value according to the calibration formula (white line in Figure 9.1.8) with input value y_{sT} from the experiment. Additionally, each Mth set of posterior draws of the calibration errors, namely, $\{\epsilon_{sT}^{(M)} = \phi_s^{(M)} - \hat{\phi}_{sT}: s = 1, \ldots, n_T\}$ (where n_ℓ is the number of observations in the experiment with the ℓth camera), is associated with the Mth posterior draw of the calibration root-mean-square error

(RMSE), namely, $\text{RMSE}_T^{(M)} = \sqrt{(n_T - 1)^{-1} \sum_{s=1}^{n_T} (\epsilon_{sT}^{(M)})^2}$; the widget displays the median of these $\text{RMSE}_T^{(M)}$'s. Thus, the user enters a survey value y_T^* as the widget input, and receives an output calibrated value $\hat{\phi}_T^*$ and the displayed posterior median RMSE as an approximate error associated with $\hat{\phi}_T^*$.

9.1.5. DATA RICH SCENARIOS — A SIMULATION STUDY

In a larger-scale calibration experiment with more data / fewer missing data / a more balanced experimental design, our most comprehensive calibration model may be desirable. Here, we conduct a simulation study in which the model in Section 3.2 is used to generate data, while fixing population-level parameters at values that are inspired by the results from Section 3.4. The computer implementation of assigning parameter values and simulating data can be found in the Supplementary Materials. An overview of the procedure is as follows.

First, we directly adopt from our 21 boat trips the data on vessel, reef size, pooled ratio r, MaxN's, and acoustic and mark-recapture counts. We also extract $\widehat{\ln\phi}$ (= posterior median values for $\ln\phi$) from Section 3.4. Values assigned for $\beta_0, \nu_{x,i}, \gamma_{x,j}$, and $\sigma_{\phi,j}$ in Level 3 of the simulation are inspired by a least-squares regression of $\widehat{\ln\phi}$ on vessel and reef size. Next, $\{\xi_h + \zeta_{hijk}\}$ in Level 2.2 is approximately $\{\ln(r_{ijk}N_{ijk} + 1) - \widehat{\ln\phi}_{ijk}\}$ for $h = 1$ and $\{\ln(N_{ijk}^{(mr)} + 1) - \widehat{\ln\phi}_{ijk}\}$ for $h = 2$, which can inspire the assigned values for $\sigma_{x,h} > 0$ for all h, and for ξ_h (subject to a sum-to-zero constraint). Similarly, for $\ell \neq R$ in Level 2.1, a least-squares regression of $\ln y (\approx \ln\mu)$ on $\widetilde{\ln\phi}$ (computed based on $\widehat{\ln\phi}$), vessel, and reef size can inspire the assigned values of $\beta_{y,0,\ell}$ and $\beta_{1,j\ell}$, which are in turn used to approximate $\{\nu_{i\ell} + \gamma_{j\ell} + \varepsilon_{ijk\ell}\}$ to further inspire σ_y, as well as $\nu_{i\ell}$ and $\gamma_{j\ell}$ (both subject to sum-to-zero constraints). For $\ell = R$ in Level 2.1, we again take $\ln y_{ijk\ell} \approx \ln\mu_{ijk\ell}$, and regress $\ln y_{\ell=R}$ on $\widetilde{\ln\phi}, \widetilde{\ln y}_{\ell=D}, \widetilde{\ln y}_{\ell=S}, \widetilde{\ln y}_{\ell=T}$, and reef size to inspire the assigned values of $\beta_{1,jR}, \beta_{1,S}$, and $\beta_{1,T}$. However, complications result from camera types D and R never being paired at small reef structures, so that a regression of $\{\ln y_{\ell=R} - \beta_{1,1R}\widetilde{\ln\phi} - \beta_{1,S}\widetilde{\ln y}_{\ell=S} - \beta_{1,T}\widetilde{\ln y}_{\ell=T}\}$ on $\widetilde{\ln y}_{\ell=D}$ is separately fitted for large reefs only to inspire the assigned value of $\beta_{1,D}$. All these assigned slope values can be substituted back into Level 2.1 in the model to inspire the assigned values for $\beta_{y,0,R}, \gamma_{y,jR}$, and σ_{yR}. If at any step a least-squares slope coefficient is negative, it is first multiplied by -1 before being used to inspire other model parameters.

The final step before simulating data is to jitter the pooled camera ratio r on the natural log scale by a standard deviation of approximately 0.27. These and all assigned parameters are then used to simulate 126 boat trips on MaxN (y), acoustic count (N), and mark-recapture estimates ($N^{(mr)}$) according to Levels 1–3 in the model. The predictor variables in each set of 126 ($= 21 \times 6$) boat trips is made up of replicating 6 times the original 21 observations on vessel, reef size, and camera ratio, with the exception that the camera ratios are jittered instead of cloned.

We simulate 50 datasets of 126 boat trips in this manner, then fit the comprehensive calibration model to each dataset, again using NIMBLE and visual assessment of MCMC convergence. We extract nominal 90% credible intervals for each population-level parameter in the model, and record the capture rate out of 50. We additionally monitor the capture rate for the derived parameters $\mathcal{M}_{ij}^{\ln} = E(\ln\phi_{ijk}) = \beta_0 + \nu_{x,i} + \gamma_{x,j}$ and $\mathcal{M}_{ij} = E(\phi_{ijk}) = \exp\{\mathcal{M}_{ij}^{\ln} + \sigma_{\phi,j}^2/2\}$.

9.1.5.1 RESULTS

Table 9.1.3 presents the capture rate of nominal 90% credible intervals for all population-level parameters, as well as $\mathcal{M}_{ij}^{\ln}$ and $\mathcal{M}_{ij}$.

While the capture rate for each parameter in Levels 2.2 and 3 is near or exceeds the nominal rate of 90%, that for many parameters in Level 2.1 is as poor as 0%, although 6 of the Level 2.1 parameters exhibit capture rates between 0.70 and 0.98. The overall capture rate (for all parameters combined, including the derived parameters) is just shy of 0.60. These results may suggest that the upper levels in the model hierarchy are poorly identified. Notwithstanding, the key purpose of the calibration model is to reconcile gear-specific MaxN's, acoustic data, and mark-recapture estimates for the unified inference of ϕ, the expected count of the target species. In that sense, the model performs very well, whereby both $\mathcal{M}_{ij}^{\ln} = E(\ln\phi)$ and $\mathcal{M}_{ij} = E(\phi)$ are successfully recovered at a rate that is near or above the nominal capture rate.

9.1.6. ADJUSTING ACOUSTIC DATA PAIRED WITH CAMERA

In the survey, a single camera type was often deployed alongside acoustic gear. In these cases, inference of ϕ should account for the joint presence of N and y_ℓ in the survey. Thus, an immediate concern becomes the derivation of an appropriate adjustment factor r to correct N. However, each camera type may have its unique limitations, so that an adjustment factor based on the single camera type deployed in the survey may be biased. A more robust adjustment factor is needed in this case.

As seen in Figure 9.1.5, the pooled ratio r_s from our calibration experiment aligns all four camera types against each other, so that r_s as an adjustment factor for acoustic counts should be more robust than any individual camera-specific ratio. As such, for a camera-specific ratio and MaxN observed in a field survey that also deploys the echosounder, we use the ratio and MaxN to predict a pooled r_s; our prediction is based on a least-squares regression fit of r_s on the ratio and MaxN from the calibration experiment.

Figure 9.1.5 suggests that r_s depends rather linearly on each camera-specific ratio. Figure 9.1.9 further suggests a linear relationship between the camera-specific ratio and the square of ln(MaxN + 1). Thus, the mean of r_s appears to be a quadratic function of ln(MaxN + 1); due to collinearity between the square term and the camera ratio, the quadratic function can be re-expressed as a linear combination of ln(MaxN + 1) and the camera ratio.

Finally, homogeneity in Figure 9.1.5 and Figure 9.1.9, together with the trends described above, suggest that a standalone least-squares linear regression of r_s on ln(MaxN + 1) and camera ratio can be used to predict r_s in this case. That is, unlike the integrative modeling framework in Section 9.1.3, the methodology of this section results in four standalone least-squares fits — one for each camera type. The substantially reduced rigor of our simple approach here can be justified by the fact that acoustic data already reflect actual abundance, so that the rigor required to ensure its utility towards the estimation of a study-wide absolute abundance is of much less concern than the rigor required to convert data from a relative abundance metric to an actual abundance measure.

The resulting least-squares fits are operationalized into a prediction tool as part of our R Shiny widget (Figure 9.1.10). The app user chooses the camera type to display the least-squares regression coefficients and their covariance matrix based on the calibration experiment. The app user can use these statistics to manually compute an $\hat{r}$ (i.e., the predicted pooled-camera ratio) and prediction standard error given any MaxN count and associated camera-specific ratio of GAJ to GAJ+. Alternatively, the app user can enter an observed value for each of MaxN and the camera-specific GAJ:GAJ+ ratio. The widget then displays $\hat{r}$ and the associated prediction standard error.

9.1.7. ADDENDUM: Supplemental data from AL-MS regional survey

At the time of model development described above, information on the effective sampling area (ESA) of any of the sampling gears (camera or echosounder) were not yet available. Months after the calibration modeling framework of Chapter 9.1 had been finalized, the ESA data, as well as supplemental samples, emerged that could augment the calibration experimental data to improve statistical power of the calibration tool described in Section 4 and Section 6. These supplemental data, observed between March and November, 2023, comprise 24 boat trips to reef structures (one of which is a platform) that deployed all three gear types: ROV camera, SBRUV camera, and acoustic echosounder. Many MaxN or acoustic counts were 0: of the 24 trips, 1 had $y_{sS} = y_{sR} = N_s = 0$, 12 had $y_{sS} = y_{sR} = 0$ and $N_s > 0$, 2 had $y_{sS} = 0$ and $y_{sR}, N_s > 0$, and 1 had $y_{sR} = 0$ and $y_{sS}, N_s > 0$.

The final model of Section 3.4 and the least-squares regression model of Section 6 were fitted to the augmented dataset of a total of 45 observations (including the original 21 boat trips from the calibration experiment and the additional 24 observations from the AL-MS survey), and the calibration tool of Section 4 and Section 6 were updated accordingly. While the modeling approach from Section 6 did not change for the case of paired deployment of camera and echosounder, that for the case of camera-only deployment from Section 3.4 required the incorporation of the echosounder's ESA into the model. Specifically, the model statement in Level 2.1 now becomes

$$\text{Level 2.1:} \qquad N_{ijk} \sim \text{Pois}(w_{ijk}\phi_{ijk}/r_{ijk})$$

where w_{ijk} denotes the echosounder's ESA in hectares, equal to the total spatial coverage of the transects traversed by the echosounder. In this updated model, ϕ_{ijk} now denotes the expected value of the underlying GAJ density per hectare in the echosounder's view, which includes habitat and non-habitat.

Figure 9.1.8 and Figure 9.1.10 show the interface of the updated calibration tool for both scenarios.

9.1.8. SUPPLEMENT: TIMELINE OF CALIBRATION MODELING

9.1.8.1 Cross calibrating camera counts only

Model variants based on the following were fitted for exploratory purposes:

$$
\begin{aligned}
\text{Level 1:} \qquad & y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell}) \\
\text{Level 2:} \qquad \text{for } \ell \neq R: \quad & \ln\mu_{ijk\ell} = \beta_{y,0,\ell} + \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell} \\
& \begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
& \mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \rho_{\ell_1,\ell_2} \text{ for all } \ell_1 \neq \ell_2 \text{ and } \ell_1, \ell_2 = D, S, T \\
\text{for } \ell = R: \quad & \ln\mu_{ijkR} = \beta_{y,0,\ell} + \gamma_{y,jR} + \beta_{1,jR}\ln\phi_{ijk} + \\
& \qquad \beta_{1,D}\ln\mu_{ijkD} + \beta_{1,S}\ln\mu_{ijkS} + \\
& \qquad \beta_{1,T}\ln\mu_{ijkT} + \varepsilon_{ijkR} \\
& \varepsilon_{ijkR} \sim \mathcal{N}(0, \sigma_{yR}^2) \\
\text{for all } \ell: \qquad & 0 = \sum_i \nu_{y,i\ell} = \sum_j \gamma_{y,j\ell} \quad \text{for identifiability} \\
\text{Level 3:} \qquad & \ln\phi_{ijk} = \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
& \delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2) \\
& 0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

where $\beta_{y,0,\ell} \equiv 0$ to reduce confounding, and ρ_{ℓ_1,ℓ_2} may not be constant across $\ell_1, \ell_2 = D, S, T$.

NOTES

- The purpose of these exploratory models was not true calibration.
- We decided to force ρ_{ℓ_1,ℓ_2} to be constant for all future versions.
- Further model refinement for these exploratory models was not pursued.

9.1.8.2 Mark-recapture counts (GAJ only) and "raw" acoustic counts (all GAJ+ species, unadjusted)

1. Model variants based on the following were fitted to incorporate mark-recapture (MR) data as true absolute abundance observations, i.e., we calibrated MaxN's against expected MR count values. However, camera counts were observed on both `Reef_Size` values (`L`, `S`), yet MR counts were observed at `Reef_Size=L` only. Thus, MR counts were further calibrated against a latent "true expected abundance" that could depend on `Reef_Size`.

$$
\begin{aligned}
\text{Level 1.1:} \qquad & y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell}) \\
\text{Level 2.1:} \qquad & N_{i1k}^{(mr)} \sim \text{Pois}(\phi_{i1k}) \quad \text{for all } i,k \quad (j=1 \text{ only}) \\
\text{Level 2.2:} \quad \text{for } \ell \neq R: \quad & \ln\mu_{ijk\ell} = \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell} \\
& \begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
& \mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2 \text{ and } \ell_1, \ell_2 = D, S, T \\
\text{for } \ell = R: \quad & \ln\mu_{ijkR} = \gamma_{y,jR} + \beta_{1,jR}\ln\phi_{ijk} + \\
& \qquad \beta_{1,D}\ln\mu_{ijkD} + \beta_{1,S}\ln\mu_{ijkS} + \\
& \qquad \beta_{1,T}\ln\mu_{ijkT} + \varepsilon_{ijkR} \\
& \varepsilon_{ijkR} \sim \mathcal{N}(0, \sigma_{yR}^2) \\
\text{for all } \ell: \quad & 0 = \sum_i \nu_{y,i\ell} = \sum_j \gamma_{y,j\ell} \quad \text{for identifiability} \\
\text{Level 3:} \qquad & \ln\phi_{ijk} = \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
& \delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2) \\
& 0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

RESULTS: Posterior median of $\ln\phi$ and $\ln(N+1)$ were negatively correlated at super pyramids (Figure 9.1.11). We also decided to use quadvariate normal in Level 2 in subsequent models to reduce model complexity.

2. We repeated models under #1 but replaced Level 2.1 with acoustic counts and Level 2.2 with quadvariate normal, as follows:

$$
\begin{aligned}
\text{Level 2.1:} \quad & N_{ijk} \sim \text{Pois}(\phi_{ijk}) \quad \text{for all } i,j,k \\
\text{Level 2.2:} \quad & \begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
& \mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2
\end{aligned}
$$

i.e., we used acoustic counts as data on true absolute abundance, thus calibrating MaxN's against the acoustic expected value.

RESULTS: On the log scale, the estimated "true expectation" ($\ln\phi$) aligned almost perfectly with logged MR counts but poorly with logged acoustic counts (Figure 9.1.12).

3. We repeated models under #2 but replaced acoustic counts with updated data (acoustic signals transcribed with a refined algorithm).
RESULTS: On the log scale, the expected acoustic count $\ln\phi$ did not correlate with logged MR counts whatsoever, nor did it align well even with $\ln(N+1)$ itself (i.e., showed a non-0 intercept and non-unit slope in Figure 9.1.13). The misalignment may be worse on the count scale (Figure 9.1.14, top row), especially for D- and R-calibrations (Figure 9.1.14, bottom row).

4. We incorporated both MR and the updated acoustic counts as paired observations of the true absolute abundance. The models we fitted here were based on the following:

$$
\begin{aligned}
&\text{Level 1.1:} && y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell}) \\
&\text{Level 1.2:} \quad \text{for } h=1: && N_{ijk} \sim \text{Pois}(\tau_{hijk}) \quad \text{for all } i,j,k \\
&\quad \text{(acoustic)} \\
&\quad \text{for } h=2: && N_{i1k}^{(mr)} \sim \text{Pois}(\tau_{hi1k}) \quad \text{for all } i,k \quad (j=1 \text{ only}) \\
&\quad \text{(mark-recapture)} \\
&\text{Level 2.1:} && \ln\mu_{ijk\ell} = \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell} \\
& && \begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
& && \mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2 \\
& && 0 = \sum_i \nu_{y,i\ell} = \sum_j \gamma_{y,j\ell} \quad \text{for identifiability} \\
&\text{Level 2.2:} \text{for } h=1: && \ln(r_{ijk}\tau_{hijk}) = \ln\phi_{ijk} + \xi_1 + \zeta_{hijk} \\
&\quad \text{for } h=2: && \ln\tau_{hi1k} = \ln\phi_{i1k} + \xi_2 + \zeta_{hi1k} \quad (j=1 \text{ only}) \\
&\quad \text{for } h=1,2: && \zeta_{hi1k} \sim \mathcal{N}(0, \sigma_{x,1}^2) \\
&\quad \text{for } h=1,2: && \zeta_{hi2k} = 0 \quad \text{for identifiability due to } k_{i2} \leq 2 \\
& && 0 = \sum_s \xi_s \quad \text{for identifiability} \\
&\text{Level 3:} && \ln\phi_{ijk} = \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
& && \delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2) \\
& && 0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

where $r_{ijk} = 1$ for all i, j, k, i.e., we calibrated MR and acoustic counts against a latent "true expectation" (by letting each to be equal to the true GAJ density, ϕ, but subject to an offset on the log scale), and calibrated MaxN's against ϕ.

RESULTS: The parameter $\ln\phi$ barely correlated with logged MR counts, and had a non-unit slope with $\ln(N+1)$ (Figure 9.1.15).

9.1.8.3 Adjusted acoustic counts using preliminary scaling factor, r^*

These models employed a preliminary scaling factor, r^*, defined as the ratio between the MaxN for all *Seriola* species (instead of only *Seriola dumerili*, i.e., GAJ) and the MaxN for all GAJ+ species, pooled from drop- and ROV-camera observations only. At that stage of the modeling process, GAJ+ MaxN's were not yet available for the other two camera types. Also, we considered either N or $N^{(f)}$ in the model based on the updated acoustic counts from Section 9.1.8.2, Item 3.

The following model versions were fitted.

1. We calibrated GAJ MaxN's against the adjusted acoustic expected value, thus fitting model variants based on the following:

$$
\begin{aligned}
\text{Level 1.1:} \quad y_{ijk\ell} &\sim \text{Pois}(\mu_{ijk\ell}) \\
\text{Level 2.1:} \quad N_{ijk} &\sim \text{Pois}(\phi_{ijk}/r^*_{ijk}) \quad \text{for all } i,j,k \\
\text{Level 2.2:} \quad \ln\mu_{ijk\ell} &= \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell} \\
\begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} &\sim \text{MVN}(\mathbf{0}, \sigma_y^2\mathbb{P}) \\
\mathbb{P}_{\ell_1,\ell_2} &= \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2 \\
0 = \sum_i \nu_{y,i\ell} &= \sum_j \gamma_{y,j\ell} \quad \text{for identifiability} \\
\text{Level 3:} \quad \ln\phi_{ijk} &= \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
\delta_{ijk} &\sim \mathcal{N}(0, \sigma^2_{\phi,j}) \\
0 = \sum_i \nu_{x,i} &= \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

RESULTS: Again, $\ln\phi$ barely correlated with logged MR counts, and had non-zero intercept and non-unit slope with $\ln[(N+1)r^*]$ (Figure 9.1.16). However, results were desirable on the count scale (Figure 9.1.17) for ϕ vs. each of $N^{(mr)}$ (high correlation) and Nr^* (approximately 0 intercept and unit slope).

2. We replaced N in #1 in this section with $N^{(f)}$ by replacing Level 2.1 with

$$N^{(f)}_{ijk} \sim \text{Pois}(\phi_{ijk}/r^*_{ijk}) \quad \text{for all } i,j,k.$$

RESULTS: Performance was less desirable than #1 (plots not shown).

3. Similar to Item 4 under Section 9.1.8.2, here, we incorporated both Nr^* and $N^{(mr)}$ as paired data on the true the absolute abundance, i.e., we fitted model variants based on the following:

Level 1.1: $y_{ijk\ell} \sim \text{Pois}(\mu_{ijk\ell})$

Level 1.2: for $h = 1$ (acoustic): $N_{ijk} \sim \text{Pois}(\tau_{hijk})$ for all i, j, k

for $h = 2$ (mark-recapture): $N_{i1k}^{(mr)} \sim \text{Pois}(\tau_{hi1k})$ for all i, k $(j = 1 \text{ only})$

Level 2.1:
$$\ln\mu_{ijk\ell} = \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell}$$
$$\begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} \sim \text{MVN}(\mathbf{0}, \sigma_y^2\mathbb{P})$$
$$\mathbb{P}_{\ell_1,\ell_2} = \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2$$
$$0 = \sum_i \nu_{y,i\ell} = \sum_j \gamma_{y,j\ell} \quad \text{for identifiability}$$

Level 2.2: for $h = 1$: $\ln(r_{ijk}^* \tau_{hijk}) = \ln\phi_{ijk} + \xi_1 + \zeta_{hijk}$

for $h = 2$: $\ln\tau_{hi1k} = \ln\phi_{i1k} + \xi_2 + \zeta_{hi1k}$ $(j = 1 \text{ only})$

for $h = 1,2$: $\zeta_{hi1k} \sim \mathcal{N}(0, \sigma_{x,1}^2)$

for $h = 1,2$: $\zeta_{hi2k} = 0$ for identifiability due to $k_{i2} \leq 2$
$$0 = \sum_s \xi_s \quad \text{for identifiability}$$

Level 3:
$$\ln\phi_{ijk} = \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk}$$
$$\delta_{ijk} \sim \mathcal{N}(0, \sigma_{\phi,j}^2)$$
$$0 = \sum_i \nu_{x,i} = \sum_j \gamma_{x,j} \quad \text{for identifiability}$$

RESULTS: The parameter $\ln\phi$ barely correlated with logged MR counts, and had a non-unit slope with $\ln[(N+1)r^*]$ (Figure 9.1.18). Performance on the count scale was not evaluated.

4. We repeated the models in #3 under this section, but replaced N with $N^{(f)}$.
   RESULTS: Performance on the log scale was less desirable than #3 (plots not shown). Again, performance on the count scale was not evaluated.

9.1.8.4 Adjusted acoustic counts by pooling all camera ratios

At this point, it was decided that $N^{(mr)}$ should no longer be modeled alongside N, but should be used purely for model validation. The decision was made due to the poor performance of models that included $N^{(mr)}$, but also because $N^{(mr)}$ was only available at one type of habitat ("super pyramid") and thus may be inherently biased.

Seeing that #1 under Section 9.1.8.3 was promising, we proceeded to repeat #1–2 from that section, but replacing r^* with r (based on all four cameras).

We fitted model variants based on the following:

$$
\begin{aligned}
\text{Level 1.1:} \qquad y_{ijk\ell} &\sim \text{Pois}(\mu_{ijk\ell}) \\
\text{Level 2.1:} \qquad N_{ijk} &\sim \text{Pois}(\phi_{ijk}/r_{ijk}) \quad \text{for all } i,j,k \\
\text{Level 2.2:} \qquad \ln\mu_{ijk\ell} &= \nu_{y,i\ell} + \gamma_{y,j\ell} + \beta_{1,j\ell}\ln\phi_{ijk} + \varepsilon_{ijk\ell} \\
\begin{bmatrix} \varepsilon_{ijkD} \\ \varepsilon_{ijkS} \\ \varepsilon_{ijkT} \\ \varepsilon_{ijkR} \end{bmatrix} &\sim \text{MVN}(\mathbf{0}, \sigma_y^2 \mathbb{P}) \\
\mathbb{P}_{\ell_1,\ell_2} &= \mathbb{P}_{\ell_2,\ell_1} = \rho \quad \text{for all } \ell_1 \neq \ell_2 \\
0 = \sum_i \nu_{y,i\ell} &= \sum_j \gamma_{y,j\ell} \quad \text{for identifiability} \\
\text{Level 3:} \qquad \ln\phi_{ijk} &= \beta_0 + \nu_{x,i} + \gamma_{x,j} + \delta_{ijk} \\
\delta_{ijk} &\sim \mathcal{N}(0, \sigma^2_{\phi,j}) \\
0 = \sum_i \nu_{x,i} &= \sum_j \gamma_{x,j} \quad \text{for identifiability}
\end{aligned}
$$

RESULTS: Focusing on the count scale, we saw the best overall alignment of ϕ with $N^{(mr)}, rN$, and $rN^{(f)}$ among all fitted models throughout the Supplement timeline up until this point.

9.1.8.5 Adjusted acoustic counts by updating pooled camera ratios

At this point, it was decided that the adjustment factor r would be recomputed to align with how they would be computed from the field survey data. Additionally, as the approach under Section Section 8.4 was the most promising, we reintroduced $\beta_{y,0,\ell}$ to Level 2.2 before proceeding. The model variant with the best performance under this section is described in Section 3.4.

9.1.8.6 Model-based derivation of fish density using acoustic counts

Acoustic echosounder data provide geo-referenced counts for GAJ+ species along multiple transects around a reef structure (Figure 9.1.19, left panel). Ideally, model-based spatial interpolation is applied to infer fish counts across the entire sampled area made up of many grid cells, each being an observational unit for the study-wide GAJ survey. Model-based interpolation also allows uncertainty quantification associated with such inference.

If a habitat structure is part of a large contiguous reef structure or unconsolidated bottom with non-trivial spatial footprint, then the average height of the interpolated surface over the total footprint of all such areal structures within the gth grid cell can be an estimate of fish density at such structures in grid cell g. In contrast, if the habitat structure is a point in space with minimal footprint (e.g., most artificial reefs), one should identify a *region of association* around the point-level structure to create an approximate footprint, where the association refers to the footprint of fish activity due to its attraction to the point-level reef structure in question. Identification may be possible from the shape of the fitted surface over cell g (Figure 9.1.19, right panel). An ad-hoc alternative method is to pre-assign buffer zones (e.g., 100m radius) around each point-level structure to create an approximate region of association.

TABLES FOR CHAPTER 9.1

Table 9.1.1: Some identifying features of the 21 boat trips in the calibration experiment.

	CruiseID	Vessel	SiteNo	Date	SiteID	Reef_Lat	Reef_Long	Reef_Type	Reef_Origin
1	ES040524	ES	1	2024-04-05	Spmid 18-030	29.772299	-87.839747	super pyramid	Artificial
2	ES040724	ES	1	2024-04-07	Spmid 18-030	29.772299	-87.839747	super pyramid	Artificial
3	ES041324	ES	1	2024-04-13	Spmid 18-030	29.772299	-87.839747	super pyramid	Artificial
4	WI041424	WI	1	2024-04-14	Spmid 18-030	29.772299	-87.839747	super pyramid	Artificial
5	ES042524	ES	1	2024-04-25	Spmid 18-030	29.772299	-87.839747	super pyramid	Artificial
6	ES040624	ES	1(3)	2024-04-06	Spmid 18-061	29.701329	-87.81156	super pyramid	Artificial
7	WI041324	WI	1	2024-04-13	Spmid 18-061	29.701329	-87.81156	super pyramid	Artificial
8	ES041424	ES	1	2024-04-14	Spmid 18-061	29.701329	-87.81156	super pyramid	Artificial
9	WI041524	WI	1	2024-04-15	Spmid 18-061	29.701329	-87.81156	super pyramid	Artificial
10	ES042424	ES	1	2024-04-24	Spmid 18-061	29.701329	-87.81156	super pyramid	Artificial
11	ES033024	ES	1	2024-03-30	B-204_2_contact0002	29.81703	-87.70383	super pyramid	Artificial
12	WI041424	WI	2	2024-04-14	A-115 contact 0017	29.9032	-87.95587	rock outcrop	Natural
13	ES041524	ES	1	2024-04-15	B-258contact0004	29.66377	-87.58898	large rocks	Natural
14	WI041524	WI	2	2024-04-15	A-136 contact 0007	29.84015	-87.88298	chicken coop	Artificial
15	WI041524	WI	3	2024-04-15	A-38 contact 0008	29.9797	-88.04502	barge	Artificial
16	ES041924	ES	1	2024-04-19	A-6contact0006	29.98465	-88.07212	wreck	Artificial
17	ES041924	ES	2	2024-04-19	A-38contact0021	29.97072	-88.03952	chicken coop	Artificial
18	ES042024	ES	1	2024-04-20	A-182_2_contact0004	29.87623	-87.74035	super pyramid	Artificial
19	ES042024	ES	2	2024-04-20	A-7_2_contact0011	29.86917	-87.77467	super pyramid	Artificial
20	ES042424	ES	2	2024-04-24	B-6_2_contact0018	29.68735	-87.81175	super pyramid	Artificial
21	ES042524	ES	2	2024-04-25	B-0_2_contact0021	29.77343	-87.8823	super pyramid	Artificial

Table 9.1.2: “Large-bodied” species list. For this table, a ‘1’ under a camera type denotes a species that was observed in the experiment by that camera type, and a ‘1’ under the “Ratio” column denotes a large-bodied fish (total length above 20 cm) that generally are found above 0.5 m from the seafloor, do not form dense aggregations (i.e., are solitary or form loosely packed schools), and are thus considered to be acoustically indistinguishable from GAJ.

Scientific Name	Common Name	Drop	SBRUV	Trap	ROV	Ratio
Lutjanus campechanus	Red Snapper	1	1	1	1	1
Seriola dumerili	Greater Amberjack	1	1	1	1	1
Seriola rivoliana	Almaco Jack	1	1	1	1	1
Balistes capriscus	Grey Triggerfish	1	1	1	1	1
Thunnus atlanticus	Blackfin Tuna	1	0	0	0	1
Euthynnus alletteratus	Little Tunny	1	1	1	0	1
Rhomboplites aurorubens	Vermilion Snapper	1	1	1	1	0
Centropristis ocyurus	Bank Sea Bass	1	1	1	1	0
Carcharodon carcharias	Great White Shark	1	1	1	1	1
Chaetodipterus faber	Atlantic Spadefish	1	1	1	1	1
Lutjanus griseus	Grey Snapper	1	1	1	1	1
Pterois volitans	Lionfish	0	1	1	1	0
Lutjanus synagris	Lane Snapper	0	1	1	1	1
Holacanthus bermudensis	Blue Angelfish	0	1	1	1	0
Chaetodon spp.	Butterflyfish	0	0	1	1	0
Chaetodon ocellatus	Spotfin Butterflyfish	0	1	0	1	0
Rypticus maculatus	Whitespotted Soapfish	0	1	0	1	0
Pomacentridae	Damselfishes	0	0	0	1	0
Labridae	Wrasses	0	0	0	1	0
Mycteroperca phenax	Scamp	0	0	1	1	1
Aluterus monoceros	Unicorn Filefish	0	1	1	1	1
Pareques spp.	Pareques spp.	0	0	0	1	0
Archosargus probatocephalus	Sheepshead	0	1	0	1	1
Haemulon aurolineatum	Tomtate	0	1	1	1	0
Stenotomus caprinus	Longspine Porgy	0	1	1	1	0
Calamus spp.	Porgies	0	0	1	1	1
Rachycentron canadum	Cobia	0	0	1	1	1
Eques lanceolatus	Jack-knife Fish	0	0	0	1	0
Neopomacentrus cyanomos	Regal Demoiselle	0	0	0	1	0
Heteroconger sp.	Garden Eel	0	1	0	0	0
Carcharhinus plumbeus	Sandbar Shark	0	1	1	0	1
Echeneis spp.	Sharksucker	0	1	1	0	1
Diplectrum formosum	Sand Perch	0	1	0	0	0
Centropristis philadelphica	Rock Sea Bass	0	1	0	0	0
Orthopristis chrysoptera	Pigfish	0	1	0	0	0
Lagodon rhomboides	Pinfish	0	1	0	0	0
Calamus leucosteus	Whitebone Porgy	0	1	1	0	1
Scomberomorus spp.	Spanish Mackerel	0	1	1	0	1
Caranx crysos	Blue Runner	0	0	1	0	1
Carcharhinus leucas	Bull Shark	0	0	1	0	1
Pagrus pagrus	Red Porgy	0	0	1	0	1
Mycteroperca microlepis	Gag	0	0	1	0	1
Anisotremus virginicus	Porkfish	0	0	1	0	1
Sphyraena barracuda	Great Barracuda	0	0	1	0	1
Selene vomer	Lookdown	0	0	1	0	1
Caranx hippos	Crevalle Jack	0	0	1	0	1
Aluterus scriptus	Scrawled Filefish	0	0	1	0	1
Calamus nodosus	Knobbed Porgy	0	0	1	0	1
Pristipomoides sp.	Wenchmen	0	0	1	0	0
Bodianus pulchellus	Spotfin Hogfish	0	0	1	0	1
Epinephelus morio	Red Grouper	0	0	1	0	1
Decapterus punctatus	Round Scad	0	0	1	0	0

Table 9.1.3: Simulation results. Capture rate (out of 50 sets) of nominal 90% credible intervals for all population-level parameters and associated derived parameters, $\mathcal{M}_{ij}^{ln}$ and $\mathcal{M}_{ij}$, for our comprehensive calibration model (Section 9.1.3.2).

Level	Parameter	Rate	Level	Parameter	Rate	Level	Parameter	Rate
2.1	$\beta_{y,0,D}$	0.04	2.1 (R)	$\beta_{y,0,R}$	0.46	3	β_0	0.98
(D/S/T)	$\beta_{y,0,S}$	0.04		$\beta_{1,1R}$	0.00		$\nu_{x,1}$	0.87
	$\beta_{y,0,T}$	0.24		$\beta_{1,2R}$	0.00		$\gamma_{x,1}$	0.93
	$\beta_{1,1D}$	0.00		$\beta_{1,D}$	0.70		$\sigma_{\phi,1}$	0.83
	$\beta_{1,1S}$	0.00		$\beta_{1,S}$	0.83		$\sigma_{\phi,2}$	0.87
	$\beta_{1,1T}$	0.02		$\beta_{1,T}$	0.00		$\mathcal{M}_{11}$	0.87
	$\beta_{1,2D}$	0.37		$\gamma_{y,1R}$	0.96		$\mathcal{M}_{21}$	0.91
	$\beta_{1,2S}$	0.70		σ_{yR}	0.13		$\mathcal{M}_{12}$	0.89
	$\beta_{1,2T}$	0.98					$\mathcal{M}_{22}$	0.89
	$\nu_{y,1D}$	0.67	2.2	ξ_1	0.93		$\mathcal{M}_{11}^{\log}$	0.96
	$\nu_{y,1S}$	0.57		$\sigma_{x,1}$	0.89		$\mathcal{M}_{21}^{\log}$	0.85
	$\nu_{y,1T}$	0.57		$\sigma_{x,2}$	0.85		$\mathcal{M}_{12}^{\log}$	0.96
	$\gamma_{y,1D}$	0.89					$\mathcal{M}_{22}^{\log}$	0.96
	$\gamma_{y,1S}$	0.09						
	$\gamma_{y,1T}$	0.52						
	ρ	0.50						
	σ_y	0.46					Combined	0.59

FIGURES FOR CHAPTER 9.1

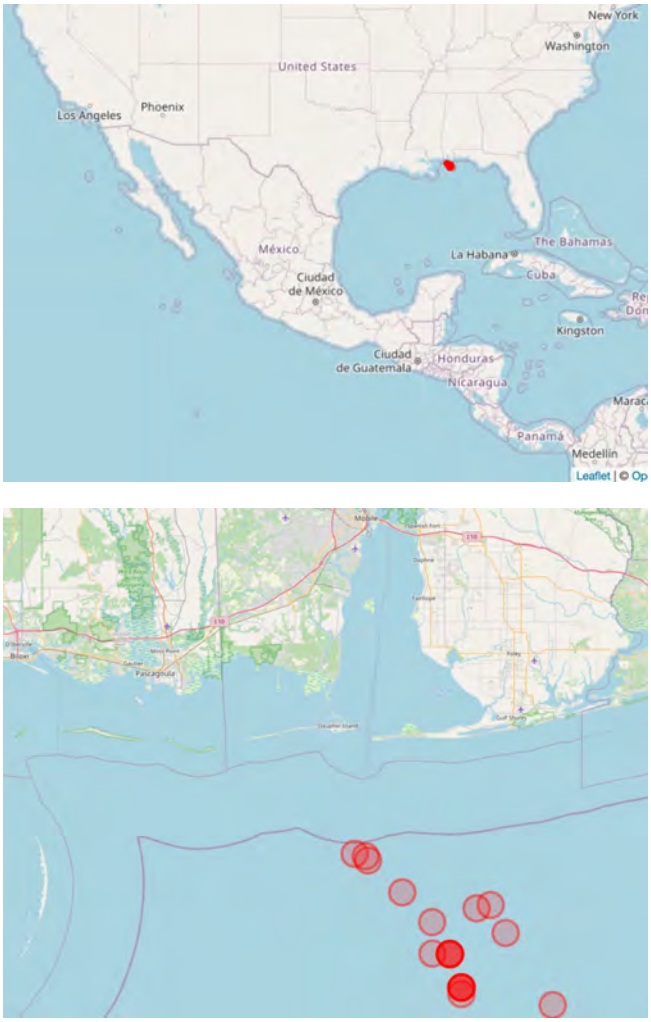


Figure 9.1.1: Maps showing the site of the calibration experiment, displayed at two spatial scales: zoomed out (top panel) and zoomed in (bottom panel). Red markers denote reef structures visited in the experiment. The two darker markers in the bottom panel denote multiple visits to the same structure (5 trips each).

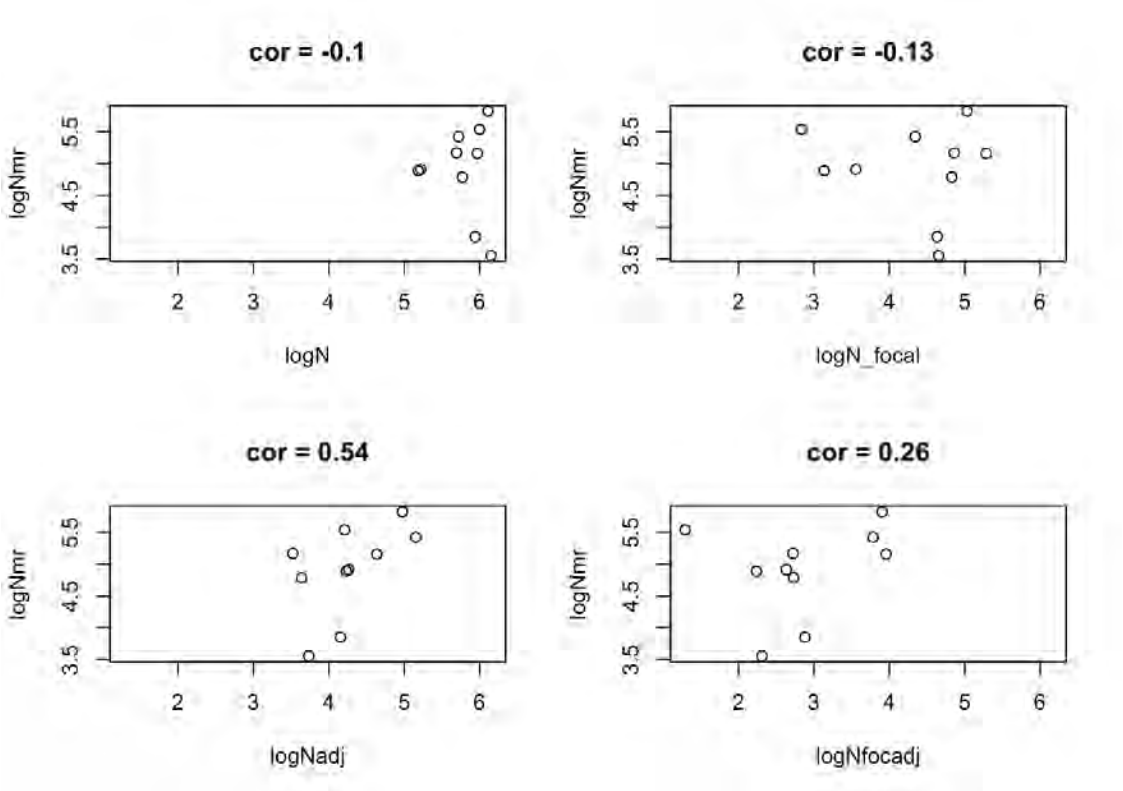


Figure 9.1.2: Mark-recapture abundance estimates ($N_s^{(mr)}$) vs. acoustic counts (N_s or $N_s^{(f)}$), the latter being unscaled (top row) or scaled by the camera ratio r_s pooled across all four camera types (bottom row). All axes are on the natural log scale after adding 1 to avoid ln(0).

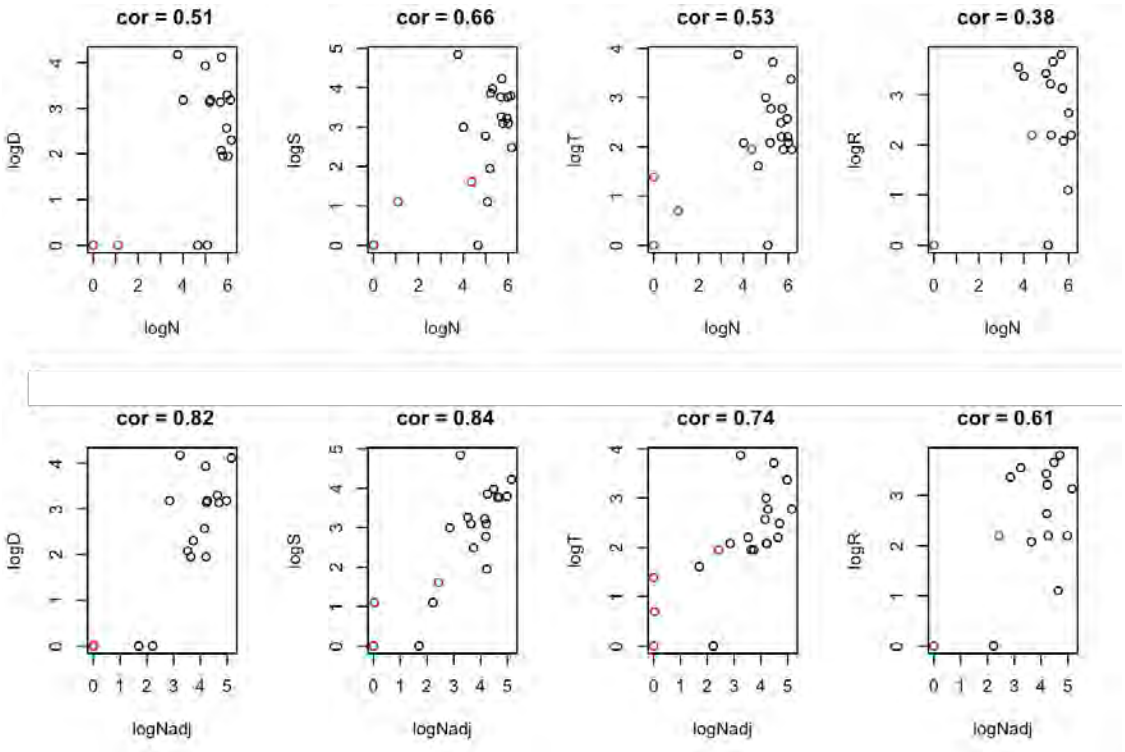


Figure 9.1.3: Camera-specific MaxN ($y_{s\ell}$) vs. acoustic count (N_s), the latter being unscaled (top row) or scaled by r_s (bottom row). Both axes are on the log scale after adding 1 to avoid ln(0). Red circles denote data points at small reef structures. Each correlation coefficient accounts for both small and large reefs.

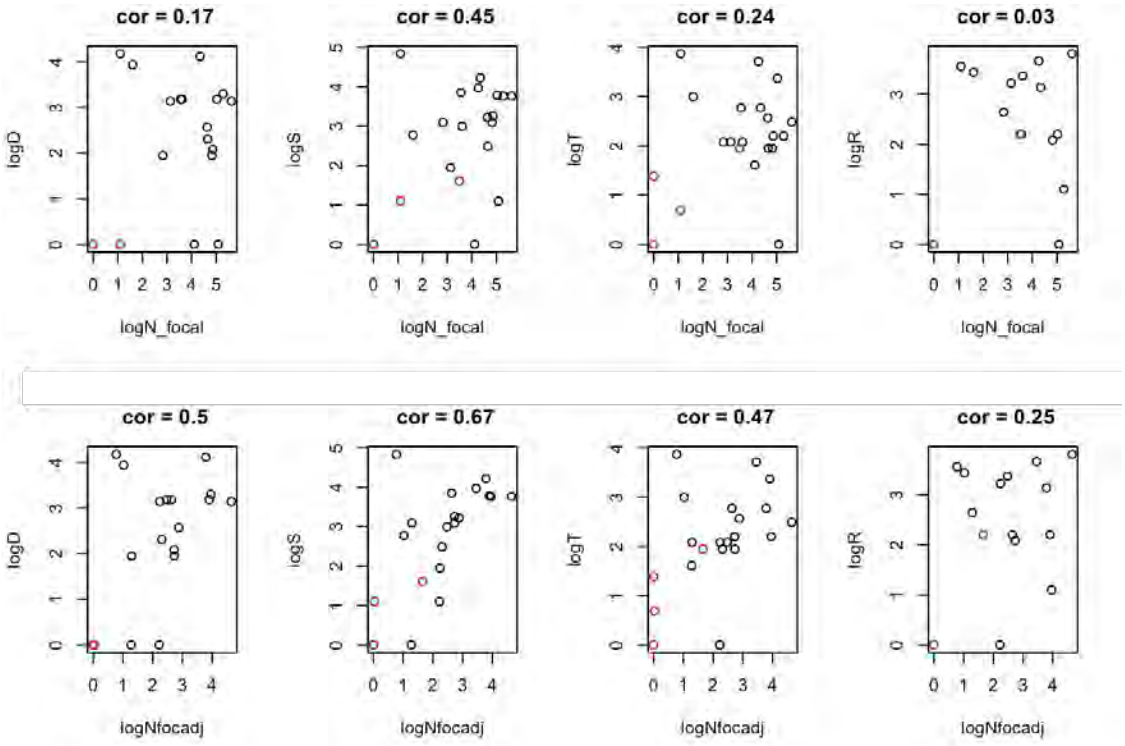


Figure 9.1.4: Camera-specific MaxN ($y_{s\ell}$) vs. focal acoustic count ($N_s^{(f)}$). See Figure 9.1.3 caption for additional details.

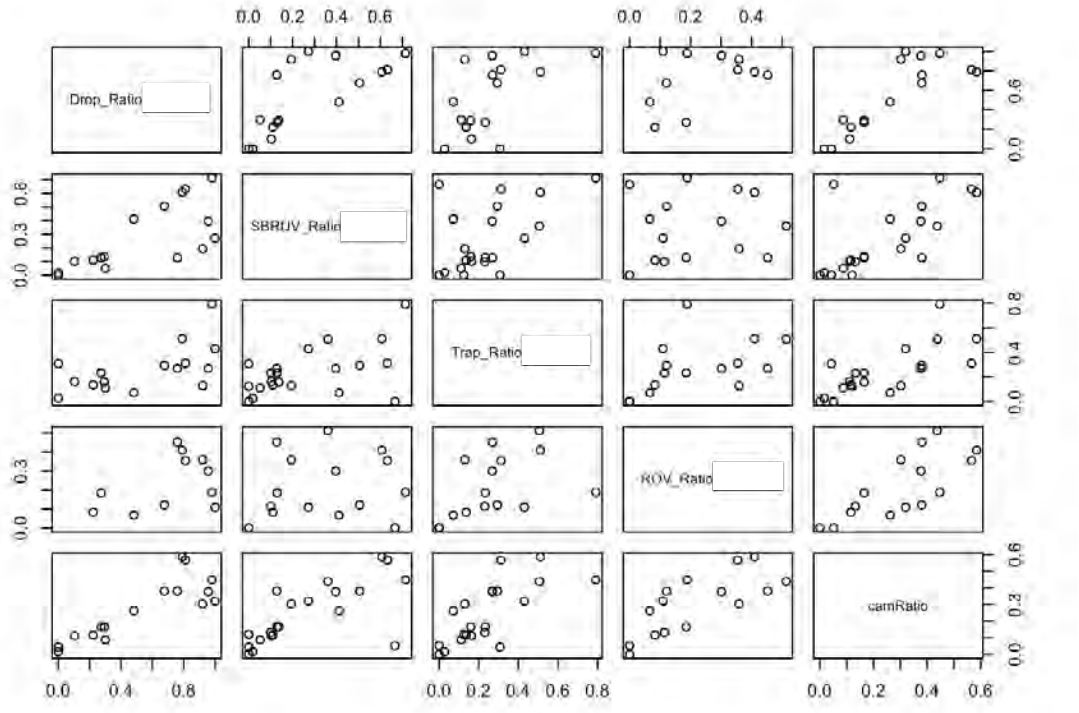
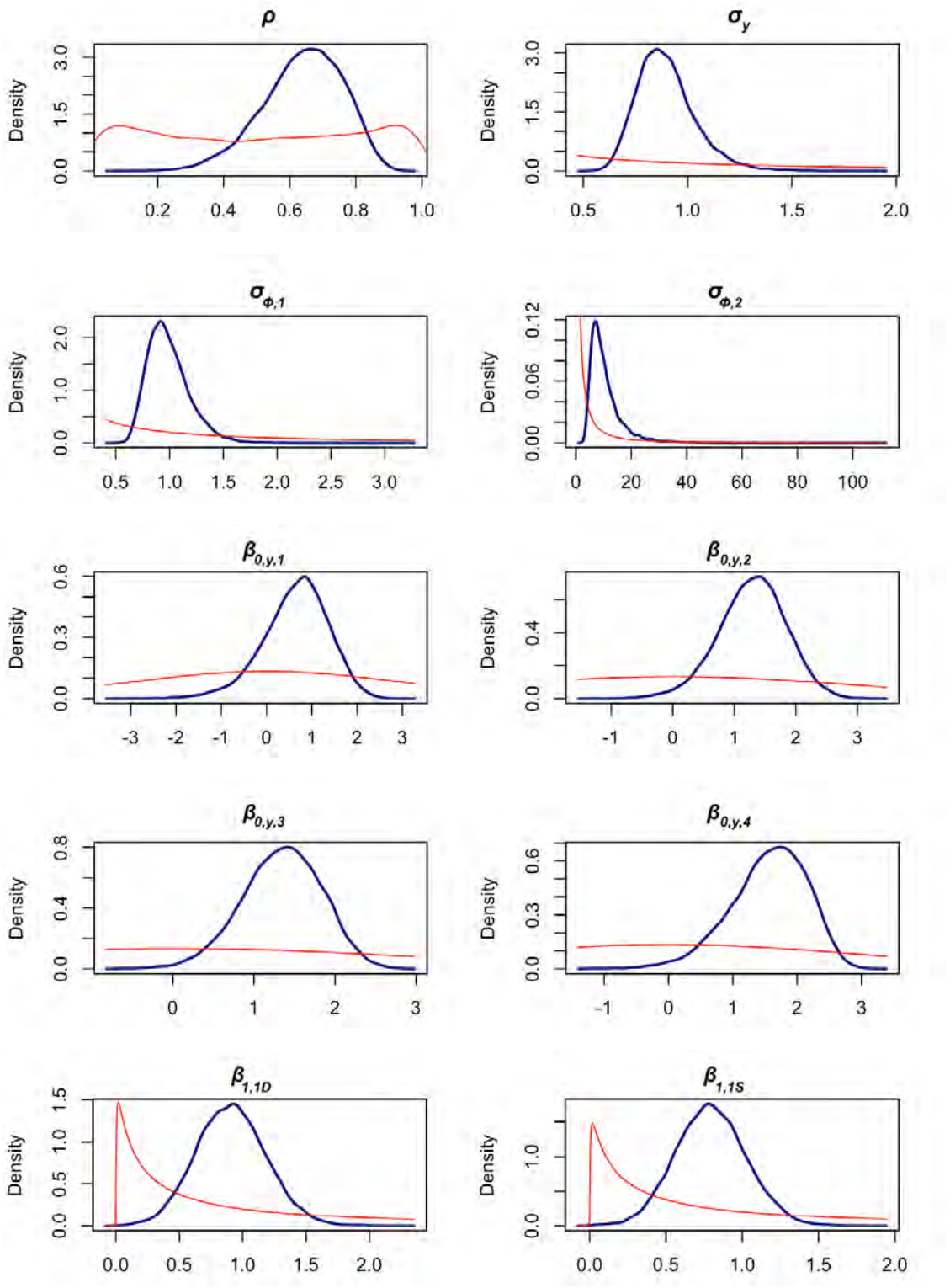


Figure 9.1.5: Scatter-plot matrix among the pooled ratio r_s (labeled as `camRatio`) and camera-specific ratios ($= y_{s\ell}/\{(s,\ell)$th MaxN for all GAJ+}, unpooled).


ρ
Density
σ_y
σ_φ,1
σ_φ,2
β_0,y,1
β_0,y,2
β_0,y,3
β_0,y,4
β_1,1D
β_1,1S

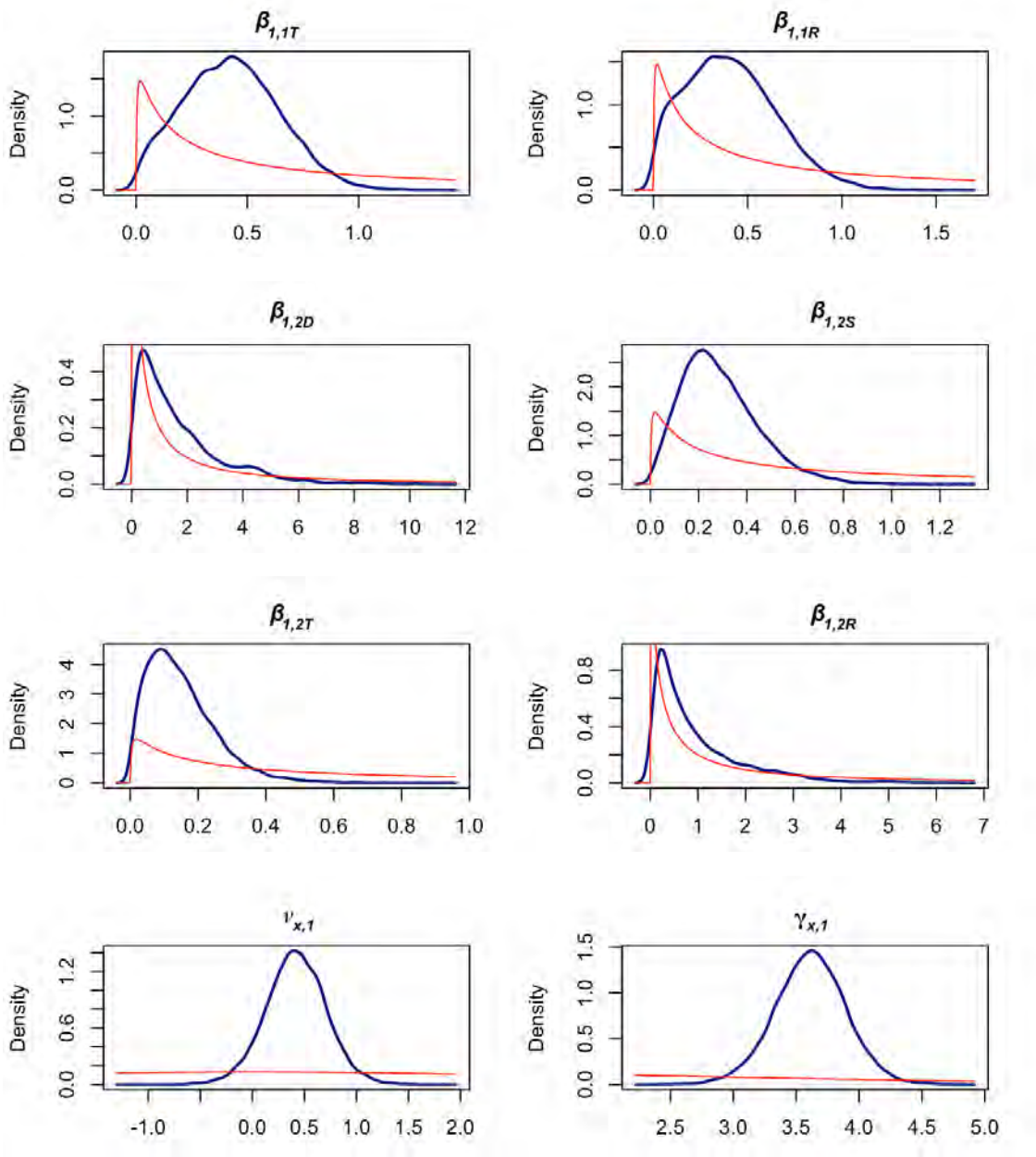


Figure 9.1.6: [This figure has 9 × 2 panels presented across 2 pages.] Posterior (blue) and prior (red) distribution for each population-level model parameter.

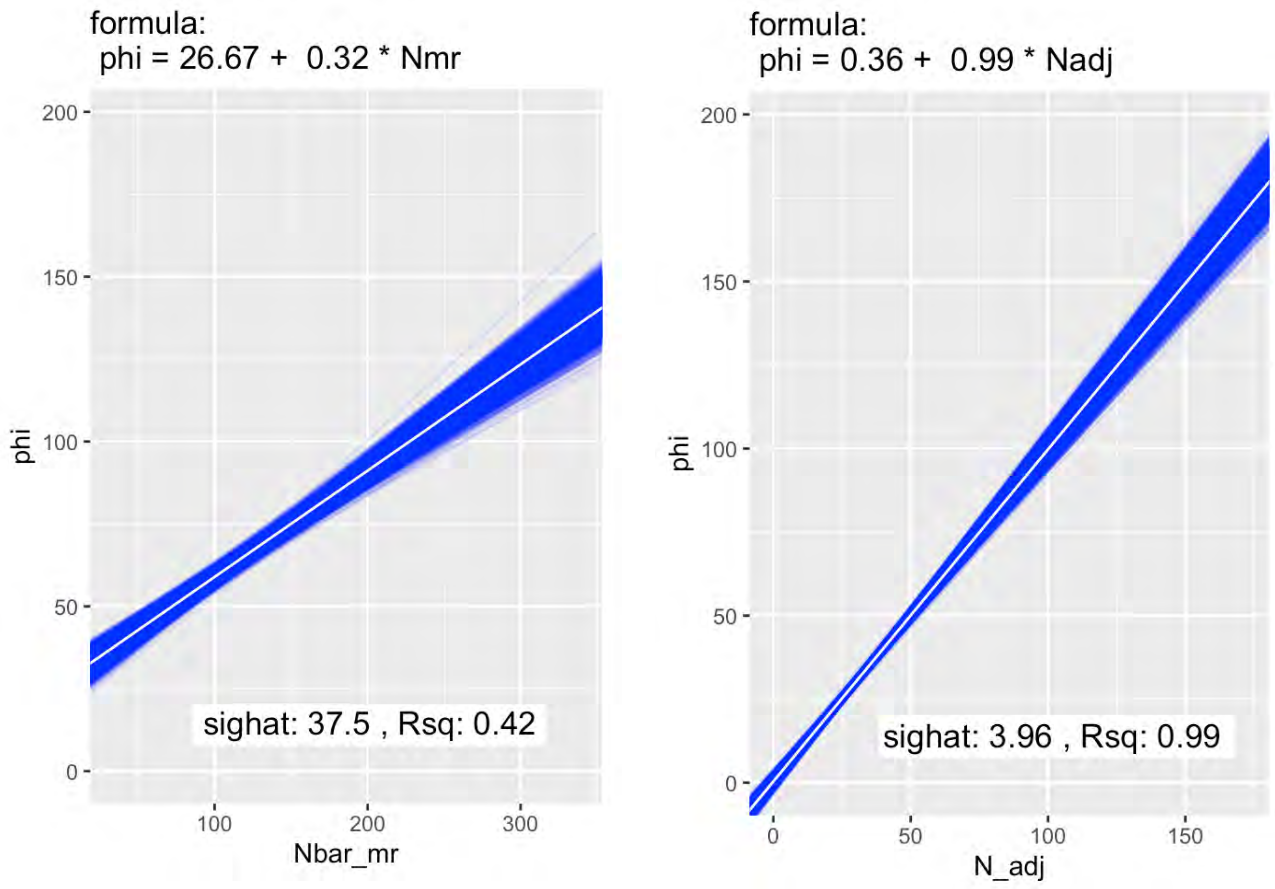


Figure 9.1.7: Simple linear regression of modeled expected GAJ abundance on mark-recapture estimate (left) and adjusted acoustic count (right): 5000 least-squares regression lines (blue) and the "median" regression line (white). Left panel: Each Mth blue line is from taking $\{\phi_s^{(M)}: s = 1,\ldots,21\}$ — the Mth posterior draw for the vector ϕ according to the final model — and regressing it on $\{N_s^{(mr)}: s = 1,\ldots,10\}$ where mark-recapture data are available (`Nmr` or `Nbar_mr` in the plot); displayed values for the intercept, slope, residual error SD (`sighat`) and R^2 are medians across the blue lines; the white line corresponds to the median intercept and slope displayed above the plot. Right panel: Same as the left panel, except regression on $\{r_s N_s: s = 1,\ldots,21\}$ (`Nadj` or `N_adj` in the plot).



Figure 9.1.8: Calibration widget for unpaired MaxN, available for download. Top-left panel: Instructions for using the widget interface in the case of either an unpaired MaxN (single camera type, unaccompanied by echosounder) or a MaxN paired with an acoustic count. Top-right panel: Interactive widget interface with calibration formula and associated output for an unpaired MaxN after changes have been made to the default input values; the white "+" denotes the location of the hovering mouse. Camera-specific posterior summaries for the corresponding Bayesian-model based calibration formula are included in the display. They include the posterior mean, median, and variance for each of the calibration intercept and slope. Also included is the posterior correlation between the calibration intercept and slope. Bottom panel: Widget output for the calibration error evaluated at each observation in the calibration experiment, with the vertical axis label displaying the posterior median of the RMSE associated with the values on the graph.

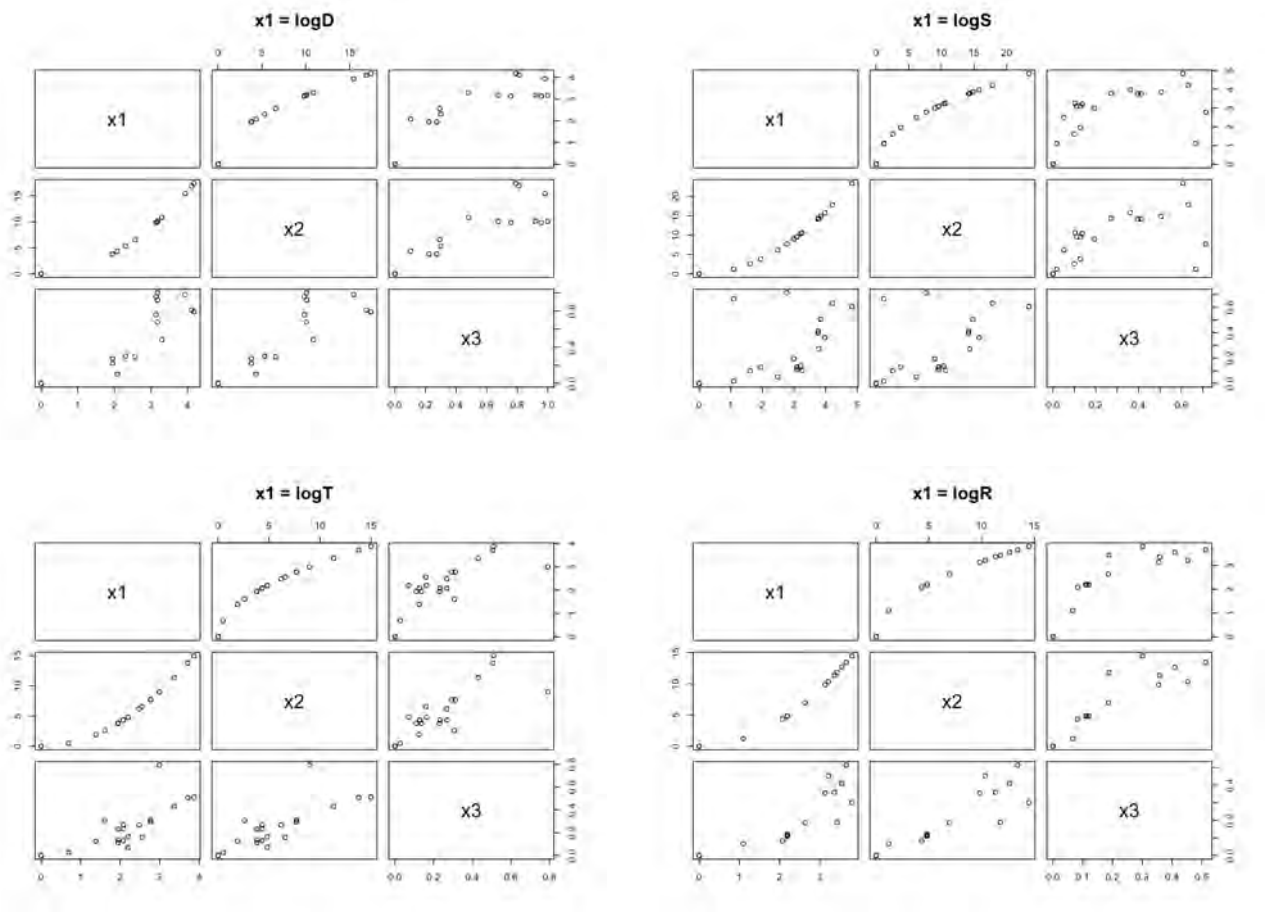


Figure 9.1.9: Scatter-plot matrices among `x1` = ln(MaxN + 1), `x2` = $(x1)^2$, and `x3` = camera ratio for drop camera (top left panel), SBRUV camera (top right panel), trap camera (bottom left panel), and ROV camera (bottom right panel).

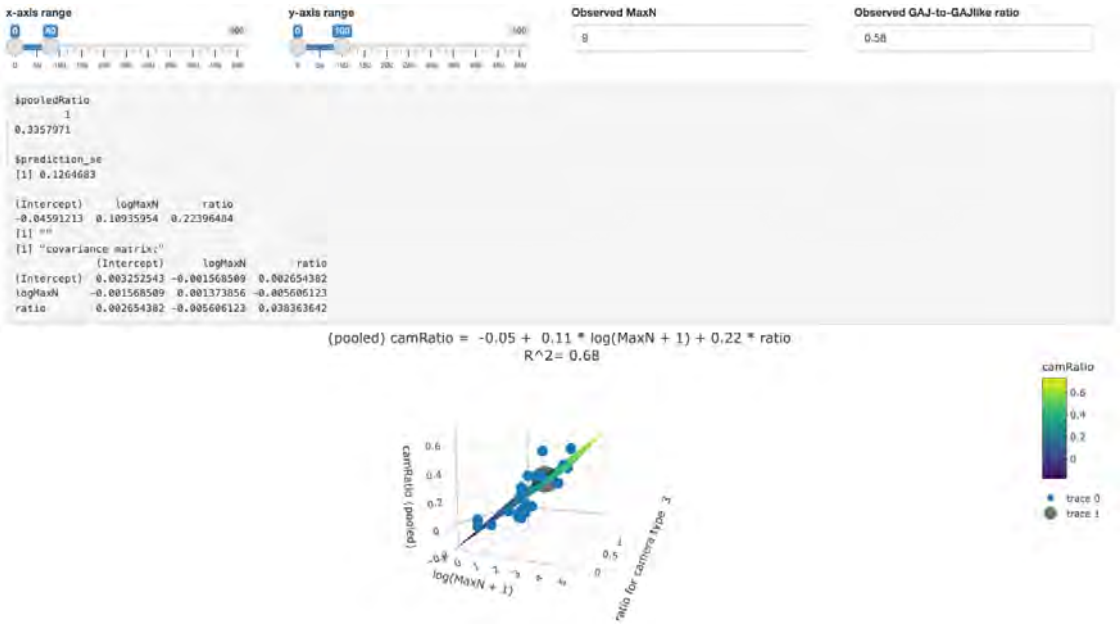


Figure 9.1.10: Calibration widget for camera paired with echosounder, available for download.
Interface with output after changes have been made to the default input values. An interactive 3D plot displays the least-squares regression of pooled camera ratio on the camera-specific MaxN (log scale) and ratio, respectively. The predicted $\hat{r}$ for the pooled camera ratio (large black dot in the 3D plot) and associated prediction standard error are provided based on the user's camera-specific MaxN and ratio. Also included are the regression coefficients, R^2, and variance-covariance matrix among the regression coefficients.

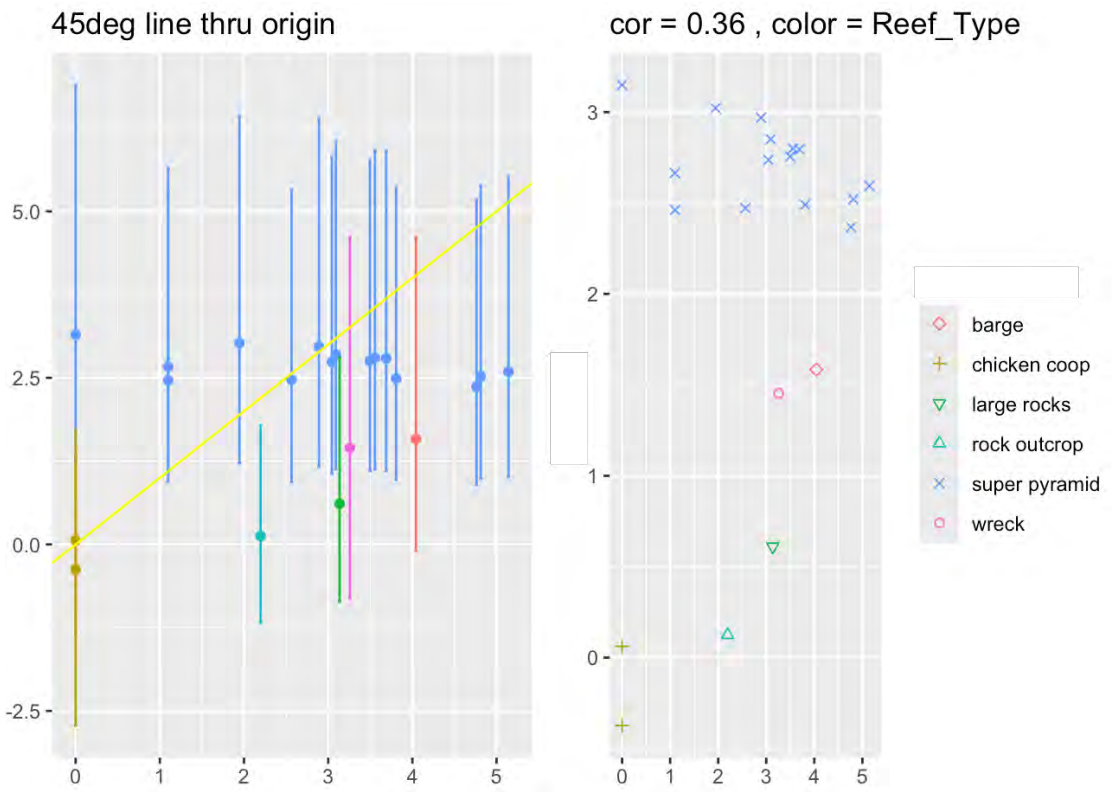


Figure 9.1.11: Posterior median for $ln\phi$ vs. $ln(N+1)$, with 80% credible intervals (left) and without (right), from one of the best fitting models under Section 9.1.8.2, Item 1.

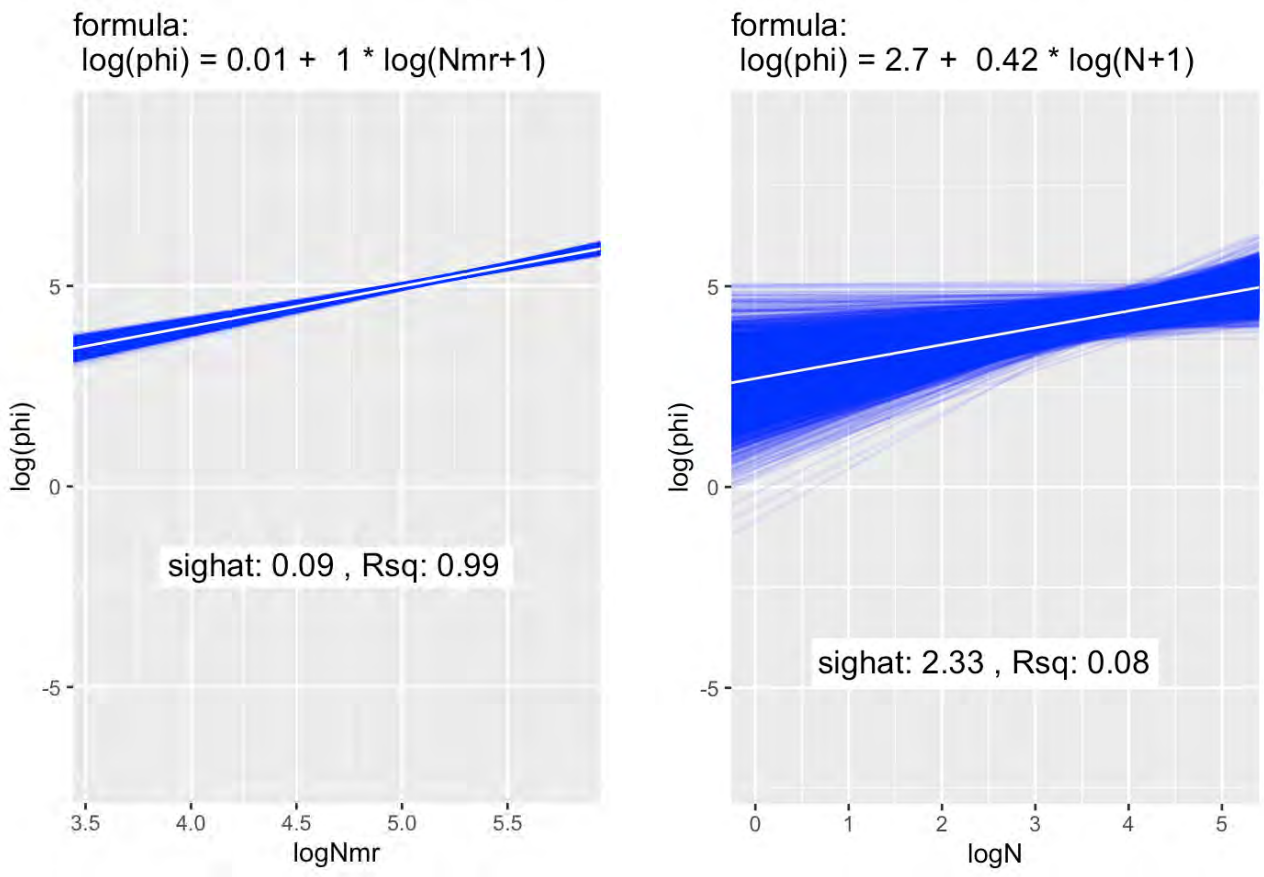


Figure 9.1.12: Posterior draws of $ln\phi$ regressed on $ln(N^{(mr)}+1)$ (left) and $ln(N+1)$ (right), from one of the best fitting models under Section 9.1.8.2, Item 2; see Figure 9.1.7 for procedural details.

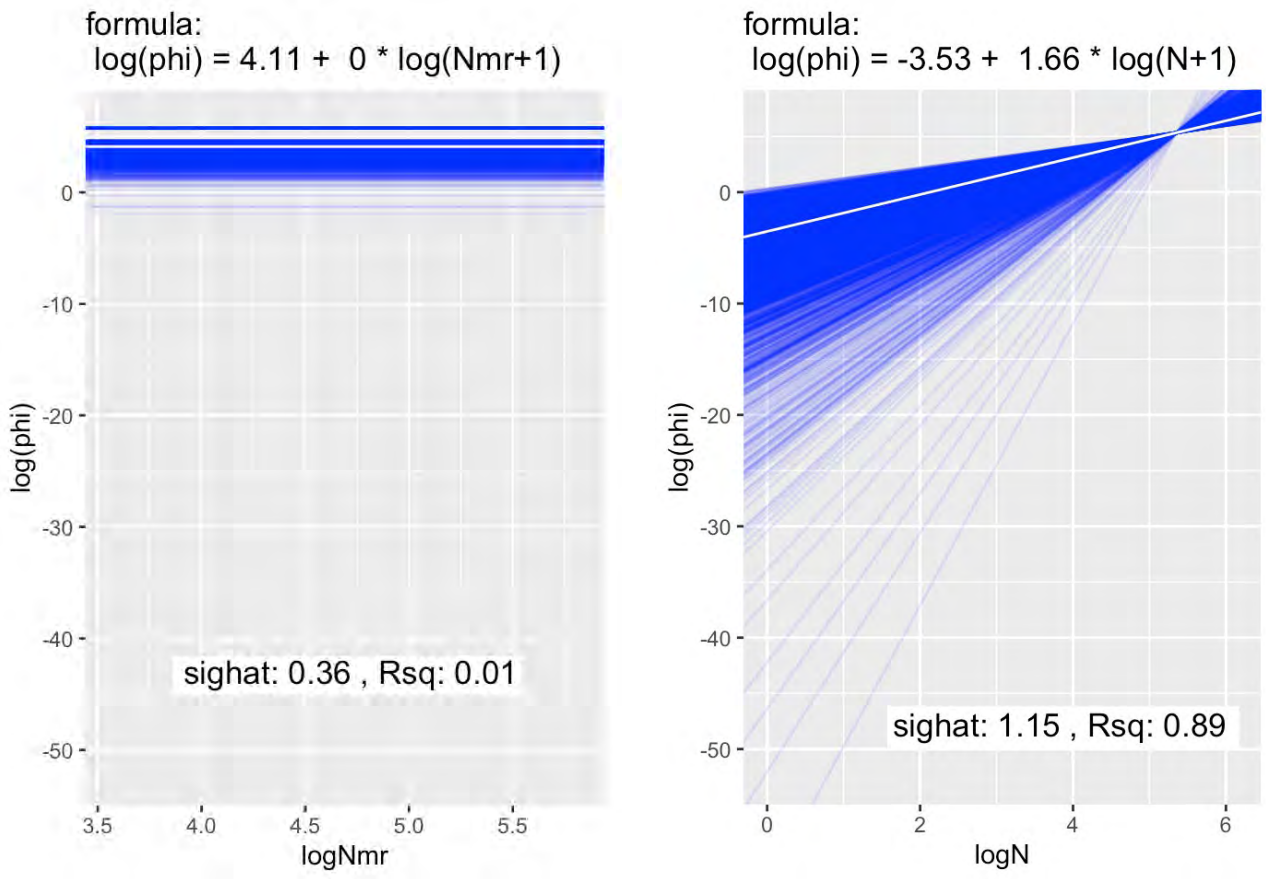


Figure 9.1.13: Same as Figure 9.1.12 but from models under Section 9.1.8.2, Item 3.

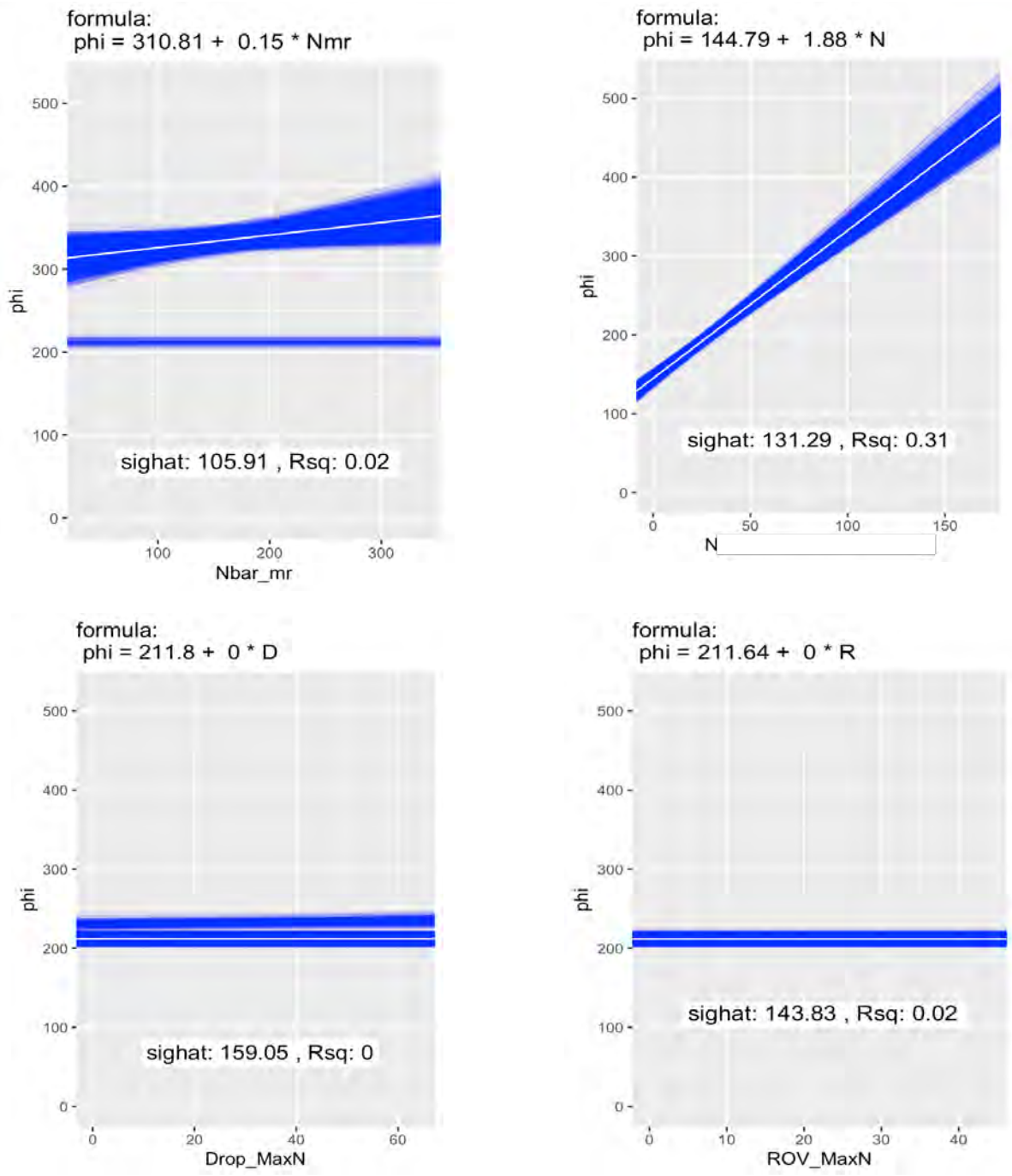


Figure 9.1.14: [This figure has 2 × 2 panels.] Same as Figure 9.1.13 but for regression on the count scale. The horizontal regression lines in three of the panels are due to posterior draws of ϕ being negatively correlated with $N^{(mr)}$, y_D, and y_R, thus producing constrained fits with a 0 slope.

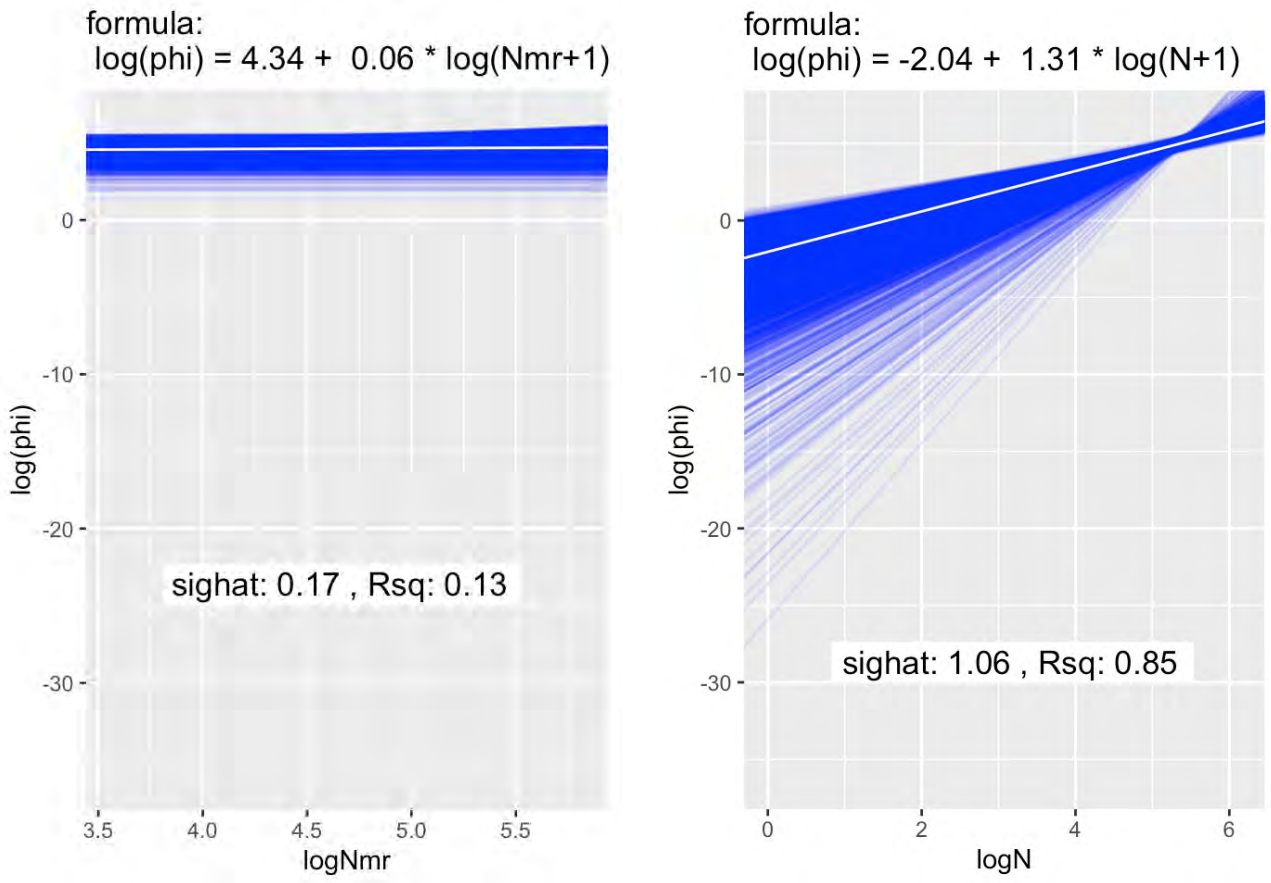


Figure 9.1.15: Same as Figure 9.1.12 but from models under Section 9.1.8.2, Item 4.

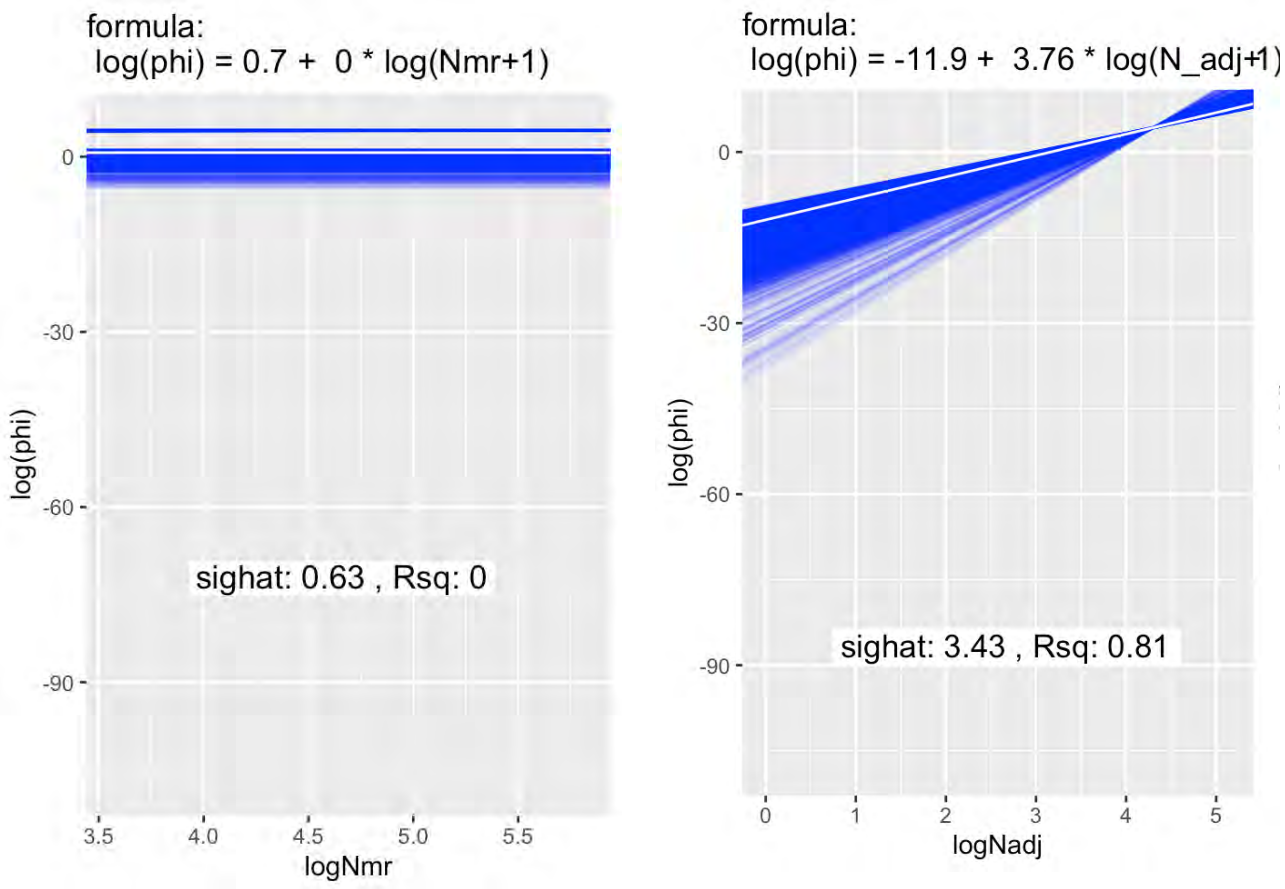


Figure 9.1.16: Same as Figure 9.1.12 but from models under Section 9.1.8.3, Item 1. `N_adj` denotes Nr^*.

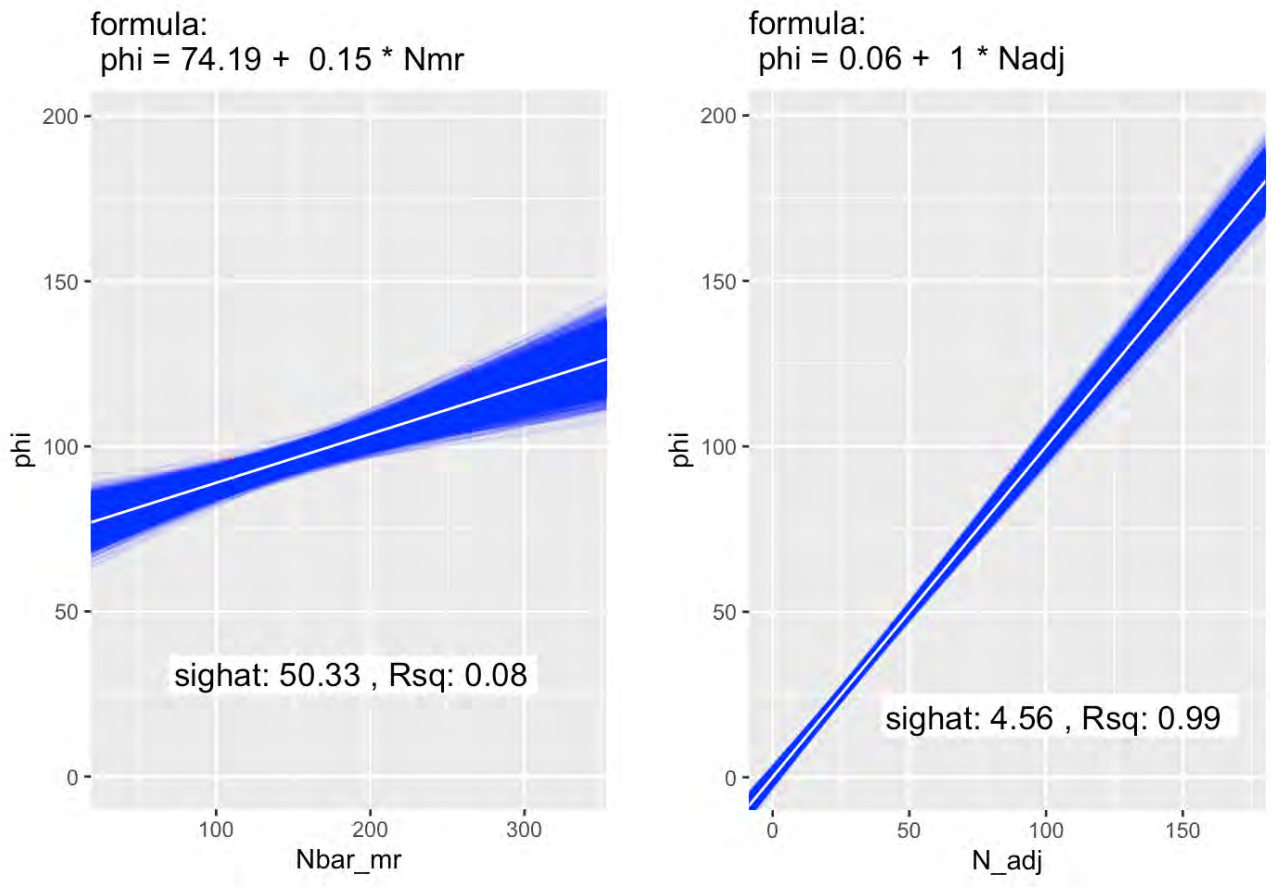


Figure 9.1.17: Same as Figure 9.1.16 but for regression on the count scale.

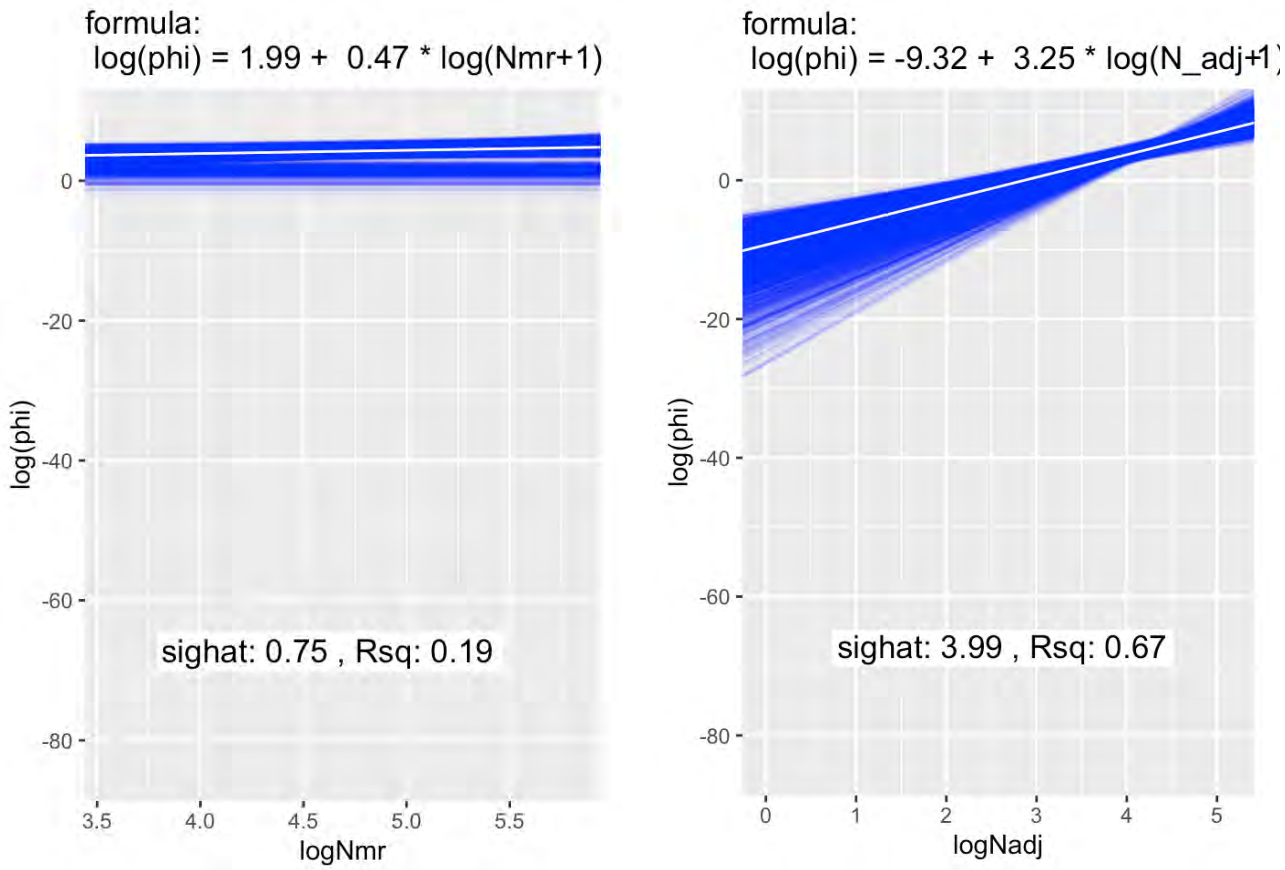


Figure 9.1.18: Same as Figure 9.1.16 but from models under Section 9.1.8.3, Item 3.

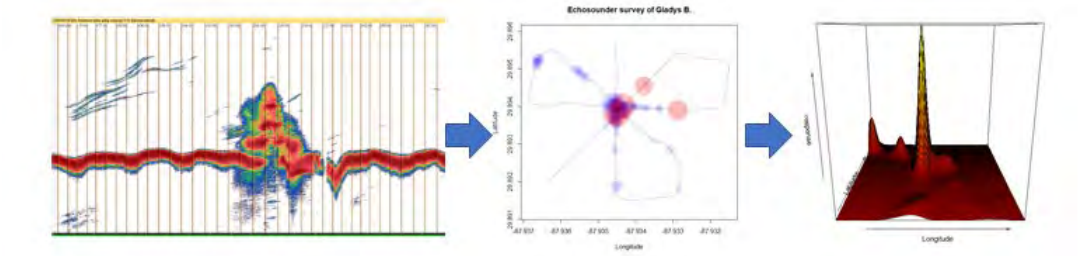


Figure 9.1.19: An example of acoustic signals recorded within a grid cell. Left panel: raw signals. Middle panel: spatial zone identified by the acoustic signal processing algorithm as the domain relevant to fish activities. Right panel: spatial interpolation of fish counts based on the middle panel.

REFERENCES FOR CHAPTER 9.1

Campbell, Matthew D., Adam G. Pollack, Christopher T. Gledhill, Theodore S. Switzer, and Douglas A. DeVries. 2015. "Comparison of relative abundance indices calculated from two methods of generating video count data." *Fisheries Research* 170 (October): 125–33. https://doi.org/10.1016/j.fishres.2015.05.011.

Campbell, Matthew D., Joseph Salisbury, Ryan Caillouet, William B. Driggers, and James Kilfoil. 2018. "Camera field-of-view and fish abundance estimation: A comparison of individual-based model output and empirical data." *Journal of Experimental Marine Biology and Ecology* 501 (April): 46–53. https://doi.org/10.1016/j.jembe.2018.01.004.

Chang, Winston, Joe Cheng, JJ Allaire, et al. 2025. *Shiny: Web Application Framework for R*. https://shiny.posit.co/.

Chiu, Grace S., Peter Guttorp, Anton H. Westveld, Shahedul A. Khan, and Jun Liang. 2011. "Latent health factor index: a statistical modeling approach for ecological health assessment." *Environmetrics* 22 (3): 243–55. https://doi.org/10.1002/env.1055.

Chiu, Grace S., Eric A. Lehmann, and Joshua C. Bowden. 2013. "A Spatial Modelling Approach for the Blending and Error Characterization of Remotely Sensed Soil Moisture Products." *Journal of Environmental Statistics* 4:9 (9). http://www.jenvstat.org/v04/i09.

Chiu, Grace S., Anton H. Westveld, Mark A. Albins, et al. 2025, revisions submitted. "Model-Based Calibration of Gear-Specific Fish Abundance Survey Data as a Change-of-Support Problem." 2025, revisions submitted. https://arxiv.org/abs/2505.05767v2.

Chiu, Grace S., Margaret A. Wu, and Lin Lu. 2013. "Model-based assessment of estuary ecosystem health using the latent health factor index, with application to the Richibucto Estuary." *PloS One* 8 (6): e65697. https://doi.org/10.1371/journal.pone.0065697.

de Valpine, P., C. Paciorek, D. Turek, et al. 2024. *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling*. R package version 1.3.0. https://cran.r-project.org/package=nimble.

Gelfand, Alan E. 2010. "Misaligned Spatial Data: The Change of Support Problem." In *Handbook of Spatial Statistics*, edited by Alan Gelfand, Peter Diggle, Peter Guttorp, and Montserrat Fuentes. CRC Press. https://doi.org/10.1201/9781420072884-37.

Gotway Crawford, C. A., and L. J. Young. 2005. "Change of support: an inter-disciplinary challenge." In *Geostatistics for Environmental Applications*, edited by Phillippe Renard, Hélène Demougeot-Renard, and Roland Froidevaux. Springer-Verlag. https://doi.org/10.1007/3-540-26535-X_1.

Kuh, F. Swen, Grace S. Chiu, and Anton H. Westveld. 2023, under revision. *Latent Causal Socioeconomic Health Index*. 2023, under revision. http://arxiv.org/abs/2009.12217v2.

McClintock, Brett. 2026. "Mark-resight models." In *Program MARK: A Gentle Introduction*, 12th ed., edited by E G Cooch and G C White. Http://www.phidot.org/software/mark/docs/book/.

Ponisio, Lauren C., Perry de Valpine, Nicholas Michaud, and Daniel Turek. 2020. "One size does not fit all: Customizing MCMC methods for hierarchical models using NIMBLE." *Ecology and Evolution* 10 (5): 2385–416. https://doi.org/10.1002/ece3.6053.

White, Gary C., and Kenneth P. Burnham. 1999. “Program MARK: Survival Estimation from Populations of Marked Animals.” *Bird Study* 46 (January): S120–39. https://doi.org/10.1080/00063659909477239.

9.2. Single to spherical camera calibration

Model-based calibration of West FL SBRUV MaxN data from single-camera scale to spherical-camera scale

Grace S. Chiu,[1] Theodore Switzer,[2] Sean Keenan[2]

[1]William & Mary's Batten School of Coastal & Marine Sciences, Virginia Institute of Marine Science; Department of Statistical Sciences and Operations Research, Virginia Commonwealth University; Department of Statistics, University of Washington; Department of Statistics and Actuarial Science, University of Waterloo

[2]Florida Fish and Wildlife Conservation Commission

SUMMARY

Chapter 9.2 here focuses on the 1129 MaxN values observed in West FL surveys, which were conducted in 2020–2023 using single-camera SBRUV reads. Of these 1129 reads, 653 were paired with corresponding spherical-camera SBRUV reads. Of the 653 pairings, 244 were part of the current GAC survey conducted in 2023. All 1129 deployments targeted either an artificial or natural structure. Note that the gear calibration modeling of Chapter 9.1 considers spherical-camera reads and not single-camera reads. Thus, to facilitate region-wide and project-wide abundance estimation, the data from single-camera reads must be converted to the scale of spherical-camera reads, whose field of view is wider.

In contrast to Chapter 9.1, for which half of the dataset was from an experiment that yielded very few 0 counts, the BRUV data in Chapter 9.2 were observed entirely from surveys with a very large number of 0s. Thus, the calibration model here must formally handle the prevalence of 0s through a model component that accounts for the probability of underlying GAJ absence. The resulting zero-inflated Poisson mixed-effects regression model is implemented as a Bayesian hierarchical model.

Of the numerous models that we considered (not all discussed here in Chapter 9.2), three of them showed similar model performance when prediction uncertainty is accounted for. However, only one of the models was adopted for converting single-camera data, on the following basis: when the posterior median of the spherical count's expected value is taken as *the converted value* across the 653 paired observations, then this converted value (on the spherical scale) based on the final adopted model tends to be higher than the original unconverted single-camera value. In contrast, the other two models yield many more converted values that are lower on the spherical scale than on the single-camera scale.

9.2.1. INTRODUCTION

Let y_1 and y_2 denote the relative abundance data (MaxN's in our case) observed by two types of camera reads conducted in tandem. Suppose y_1 corresponds to the less desirable gear (single-camera SBRUV in our case) and y_2, the more desirable one (spherical-camera SBRUV in our case) — see Figure 9.2.1. An existing approach to calibrate y_1 against y_2, e.g., by Bacheler et al. (2023), is to fit a homoscedastic simple linear regression model of y_2 on y_1 while disregarding the extreme heteroscedasticity of this type of data that is due to the high prevalence of 0s and/or the fact that the variability of abundance data increases with the underlying abundance values. One possibility to handle the prevalence of 0s might be to ignore all observations in which both camera reads yield $y_1 = y_2 = 0$. From a modeling perspective, the omission of paired 0s is not ideal, for the following reason. If GAJ were to be absent at a site, then either y_1 or y_2 should be 0; yet, a prediction tool trained on

(y_1, y_2) pairs that are never simultaneously 0 would lead to a y_2^{pred} that is never exactly 0, thus inducing a positive bias in cases where GAJ are truly absent.

A more rigorous modeling approach is to include all 0s, paired or otherwise, in the calibration model while addressing heteroscedasticity through a generalized linear modeling (GLM) structure. Miller et al. (2010) consider different types of GLMs to address heteroscedasticity of count data, although the authors regard *calibration* as equating the linear predictor for y_2 to multiplying the linear predictor for y_1. That is, the modeling objective for Miller et al. (2010) is to make inference for this multiplier in the absence of covariates. As such, instances of $y_1 = y_2 = 0$ have no effect on the inference of this multiplier, and thus they can be discarded from the training data.

While a simple multiplier to relate the linear predictors for y_1 and y_2 may be reasonable if each of y_1 and y_2 is directly proportional to the actual animal count in the sample, it may be inappropriate in our case, as MaxN is not directly proportional to the animal count in the sample. In addition, we leverage the potential explanatory power of covariates for calibration modeling, so that the relationship between the linear predictors of y_1 and y_2 must allow for an intercept and multiple slope parameters. For this reason, the prevalence of 0s (see Figure 9.2.2, top four panels) requires special handling, such as through a zero-inflated probability distribution.

We consider zero-inflated Poisson mixed-effects regression, which is well suited for small MaxN values while accounting for overdispersion, and we implement the model under a Bayesian hierarchical modeling framework.

The rest of Chapter 9.2 is structured as follows. In Section 9.2.2, we describe the data on which the calibration model is developed and trained. In Section 9.2.3, we describe the final model adopted for calibration, and describe two alternative models that performed similarly but were ultimately discarded. Section 9.2.4 describes the procedure to convert single-camera MaxN values to the scale of the spherical SBRUV, once the final model has been trained on the data from Section 9.2.2. Some final thoughts appear in Section 9.2.5. Supplementary details appear in Section 9.2.6.

9.2.2. PAIRED DATA FROM 2020–2023

Figure 9.2.3 shows the locations in West FL where the single-camera SBRUV reads are paired with spherical SBRUV reads at artificial or natural structures. Of these 653 paired observations in 2020–2023 (red and purple dots), 244 were from the current GAC survey (purple dots) conducted in 2023. The GAC survey in West FL yielded a total of 1129 observations at artificial or natural structures; the goal of Chapter 9.2 is to develop a calibration model based on the 653 paired observations, then use the model to predict SBRUV's MaxN (y_2) for the 885 (= 1129 − 244) cases of unpaired single-camera SBRUV reads in the GAC survey.

In addition to y_1 and y_2 (the two types of MaxN values), the only other variables available for modeling are $x_1 =$ depth (natural log-transformed for deskewing, see Figure 9.2.2), $x_2 =$ longitude, $x_3 =$ latitude, and habitat type (artificial or natural structure). Due to the log link function for a Poisson regression model, Figure 9.2.4 displays the relationship on the natural log scale between y_2 and other numerical variables as candidate predictors, partitioned by habitat type (also a candidate predictor). The lack of pairwise trends with y_2 in the figure aside from y_1 suggests that in the case of positive MaxN's, y_1 is the only numerical candidate predictor that has non-trivial predictive power; though, we identify that when the dataset is partitioned according to $y_1 = 0$ (585 cases) and $y_1 > 0$ (68 cases), then 1) the predictor x_1 (depth, transformed) and a newly defined predictor, $x_4 = x_2 + x_3$ (a type of average of longitude and latitude to address the high collinearity between x_2 and x_3), show a

somewhat detectable signal, and 2) habitat type is a source of heterogeneity in the linear predictor. These identified relationships are reflected by the model statements in Section 9.2.3.1.

9.2.3. MODELING FRAMEWORK

We structure each model in this section in a Bayesian hierarchical framework, and implement it using the NIMBLE Bayesian programming software for R (Ponisio et al. 2020; de Valpine et al. 2024), which produces Markov chain Monte Carlo (MCMC) simulated draws from the posterior distribution.

9.2.3.1 Model A: Final model adopted for calibration

The final adopted zero-inflated Poisson mixed-effects model for converting unpaired observations of $y_{\ell=1}$ to $y_{\ell=2}$ involves the following quantities:

- $i = 1,2$ denotes the habitat type ('1' for artificial, '2' for natural),
- j denotes the jth observation (nested within i),
- $\ell = 1,2$ denotes the type of SBRUV read ('1' for single-camera, '2' for spherical-camera),
- $x_{1ij} =$ natural-log transformed depth (m), and centered post-transformation to have mean 0,
- $x_{4ij} = x_{2ij} + x_{3ij} = 2 \times$ {average of longitude and latitude}, centered to have mean 0,
- $y_{ij\ell} =$ the (i, j, ℓ)th observed MaxN value,
- $\mu_{ij\ell} =$ the unknown expected value of $y_{ij\ell}$,
- $\varepsilon_{ij\ell} =$ the random effect associated with $\mu_{ij\ell}$,
- $\sigma_{11}^2, \sigma_{20}^2, \sigma_{21i}^2$ are variances associated with the $\varepsilon_{ij\ell}$'s,
- $\pi_{ij} =$ the probability that $\mu_{ij\ell}$ is non-zero for both BRUV types,
- $\omega_{ij} =$ the expected value of the linear predictor logit π_{ij},
- $\delta_{ij} =$ the random effect associated with π_{ij},
- $\sigma_{\pi 0}^2, \sigma_{\pi 1}^2$ are variances associated with the δ_{ij}'s,
- $\alpha_{10}, \alpha_{10}^*, \alpha_{11i}, \alpha_{21}, \beta_0, \beta_0^*, \beta_1, \beta_2, \gamma$ are regression coefficients associated with the linear predictors logit π_{ij} and $\ln\mu_{ij\ell}$.

The statement of our Bayesian hierarchical model is made up of 4 levels:

$$\text{Level 1:} \quad y_{ij\ell} \mid \mu_{ij\ell} \sim \text{Pois}(\mu_{ij\ell})$$

$$\text{Level 2:} \quad \begin{cases} P(\mu_{ij\ell} > 0) = \pi_{ij} \\ \text{logit}\, \pi_{ij} = \omega_{ij} + \delta_{ij} \end{cases} \quad \text{for } \ell = 1,2$$

$$\begin{cases} \omega_{ij} = \beta_0^* \\ \delta_{ij} \mid \sigma_{\pi 0}^2 \sim N(0, \sigma_{\pi 0}^2) \end{cases} \quad \text{for } y_{ij1} = 0$$

$$\begin{cases} \omega_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{4ij} \\ \delta_{ij} \mid \sigma_{\pi 1}^2 \sim N(0, \sigma_{\pi 1}^2) \end{cases} \quad \text{for } y_{ij1} > 0$$

$$\text{Level 3.1:} \quad \mu_{ij2} > 0: \begin{cases} \ln\mu_{ij2} = \alpha_{21}\omega_{ij} + \gamma \ln(y_{ij1} + 1) + \varepsilon_{ij2} & \text{for} \quad y_{ij1} \geq 0 \\ \varepsilon_{ij2} \mid \sigma_{21i}^2 \sim N(0, \sigma_{21i}^2) & \text{for} \quad y_{ij1} > 0 \\ \varepsilon_{ij2} \mid \sigma_{20}^2 \sim N(0, \sigma_{20}^2) & \text{for} \quad y_{ij1} = 0 \end{cases}$$

$$\text{Level 3.2:} \quad \mu_{ij1} > 0: \begin{cases} \ln\mu_{ij1} = \alpha_{10} + \alpha_{11i} x_{1ij} + \varepsilon_{ij1} \\ \varepsilon_{ij1} \mid \sigma_{11}^2 \sim N(0, \sigma_{11}^2) \end{cases} \quad \text{for } y_{ij1} > 0$$

$$\ln\mu_{ij1} = \alpha_{10}^* \quad \text{for } y_{ij1} = 0$$

$$\text{Level 4:} \quad \begin{cases} \alpha_{10}, \alpha_{10}^*, \alpha_{11i}, \alpha_{21}, \beta_0, \beta_1, \beta_2, \gamma \sim N(0, 10^2) \\ -\ln\beta_0^*, \ln\sigma_{\pi 0}, \ln\sigma_{\pi 1}, \ln\sigma_{11}, \ln\sigma_{20}, \ln\sigma_{21i} \sim N(0, 3^2) \quad \text{for } i = 1,2 \end{cases}$$

where Levels 1 and 3.2 can be considered a type of empirical Bayes framework by partitioning the behavior of μ_{ij1} (the expected value of y_{ij1}) based on the value of y_{ij1}, and Level 4 is a set of rather diffuse prior distributions (see Supplement Section 9.2.6.1 for further details on the diffuseness of these and other priors employed throughout Chapter 9.2). Note that β_0^* is constrained to be negative to make it unlikely *a priori* for the expected value of either gear's MaxN to be positive when $y_1 = 0$. Also note that heterogeneity (in this case, the dependence of a random-effects variance on i) is not universal across the model hierarchy. We arrived at the specific heterogeneity structure of Model A based on the diagnostics of many model variants (not shown) with a variety of heterogeneity structures, all of which were deemed less adequate and/or too complex compared to Model A. The same logic applies to the specific mean structures for the linear predictors of Model A.

9.2.3.2 Alternative models considered for calibration

9.2.3.2.1 Model B: Variant of Model A

This model is identical to Model A with one exception: β_0^* is unconstrained, i.e., it is allowed to take on a positive or negative value. The prior distribution employed is $\beta_0^* \sim N(0, 10^2)$.

Figure 9.2.5 shows that the inference for μ_{ij2} is very similar between Models A and B. Supplement Section 9.2.6.2 discusses how Model A of Section 9.2.3.1 was deemed to be more suitable than this model variant for calibration purposes.

9.2.3.2.2 Model 0: Standalone naive Poisson mixed-effects models

For this simplest model we considered for calibration, it partitions the dataset into two, according to $y_1 > 0$ ($n = 68$) or $y_1 = 0$ ($n = 585$). A different model is fitted to each dataset, regressing the expected value of y_2 on x_1, x_2, x_3 (all centered) and/or $\ln(y_1 + 1)$:

$$y_{ij2} \mid \mu_{ij2} \sim \text{Pois}(\mu_{ij2}) \quad \text{for Datasets 1 and 2}$$

Dataset 1 ($y_{ij1} > 0$):
$$\ln\mu_{ij2} = \alpha_0 + \alpha_1 x_{1ij} + \alpha_2 x_{4ij} + \gamma \ln(y_{ij1} + 1) + \varepsilon_{ij2}$$
$$\varepsilon_{ij2} \mid \sigma_i^2 \sim N(0, \sigma_i^2)$$

Dataset 2 ($y_{ij1} = 0$):
$$\ln\mu_{ij2} = \alpha_{0i}^* + \alpha_{2i}^* x_{2ij} + \varepsilon_{ij2}$$
$$\alpha_{2i}^* = 0 \quad \text{for } i = 1$$
$$\varepsilon_{ij2} \mid \sigma_*^2 \sim N(0, \sigma_*^2)$$

where prior distributions for all regression coefficients and random-effects variance parameters are like those in Level 4 of Section 9.2.3.1 (i.e., $N(0, 10^2)$ for regression coefficients and $N(0, 3^2)$ for $\ln\sigma$). Note that a) adding 1 to y_{ij1} here for Dataset 1 is to prevent the regression coefficient γ from ever disappearing from the model even when $y_{ij1} = 1$, and that b) the numerical predictor for Dataset 2 is x_2 (longitude).

Figure 9.2.5 shows that the inference for μ_{ij2} is very similar among Models A, B, and 0. Supplement Section 9.2.6.2 discusses how Model A of Section 9.2.3.1 was deemed to be more suitable than this simple model for calibration purposes.

9.2.4. CALIBRATING SINGLE-CAMERA DATA TO THE SPHERICAL SBRUV'S SCALE

When a calibration model from Section 9.2.3 is fitted to the data described in Section 9.2.2, we refer to it as "a model trained on the training data." Once a model has been trained, the results can be used to make predictive inference on exogenous data that are not part of the training data — exogenous quantities are labeled with a tilde overscore ("~") below.

In a calibration setting, given an unpaired observed value of $\tilde{y}_1$ and associated covariate values $\tilde{\imath}, \tilde{x}_1, \tilde{x}_2$, and $\tilde{x}_3$, the key unknown quantity of interest is $\tilde{\mu}_2$, which is the expected value of the counterfactual SBRUV's $\tilde{y}_2$. To this end, we take the trained Model A in Section 9.2.3.1, extract the posterior distribution of parameter values that are "downstream" from μ_2 in the model hierarchy, and make posterior predictive inference for $\tilde{\mu}_2 | \{\tilde{y}_1, \tilde{\imath}, \tilde{x}_1, \tilde{x}_2, \tilde{x}_3\}$, as follows:

1. Consider the mth MCMC draw from the posterior distribution.
2. Extract "downstream" parameter values $\alpha_{21}^{(m)}, \beta_0^{(m)}, \beta_0^{*(m)}, \beta_1^{(m)}, \beta_2^{(m)}, \gamma^{(m)}, \sigma_{\pi 0}^{(m)}, \sigma_{\pi 1}^{(m)}, \sigma_{20}^{(m)}, \sigma_{21\tilde{\imath}}^{(m)}$.
3. Define $\tilde{\omega}^{(m)} = \begin{cases} \beta_0^{*(m)} & \text{if} \quad \tilde{y}_1 = 0 \\ \beta_0^{(m)} + \beta_1^{(m)} \tilde{x}_1 + \beta_2^{(m)} (\tilde{x}_2 + \tilde{x}_3) & \text{if} \quad \tilde{y}_1 > 0 \end{cases}$.
4. Simulate a single random draw for each of the following:
   a. Simulate $\text{logit}\, \tilde{\pi}^{(m)} \sim \begin{cases} N\left(\tilde{\omega}^{(m)}, (\sigma_{\pi 0}^{(m)})^2\right) & \text{if} \quad \tilde{y}_1 = 0 \\ N\left(\tilde{\omega}^{(m)}, (\sigma_{\pi 1}^{(m)})^2\right) & \text{if} \quad \tilde{y}_1 > 0 \end{cases}$.
   b. Simulate $\tilde{r}^{(m)} \sim \text{Bernoulli}(\tilde{\pi}^{(m)})$.

5. If $\tilde{r}^{(m)} = 0$, then assign $\tilde{\mu}_2^{(m)} = 0$.

   If $\tilde{r}^{(m)} = 1$, then simulate

$$\ln\tilde{\mu}_2^{(m)} \sim \begin{cases} N\left(\alpha_{21}^{(m)}\tilde{\omega}^{(m)} + \gamma^{(m)}\ln(\tilde{y}_1 + 1), (\sigma_{20}^{(m)})^2\right) & \text{if} \quad \tilde{y}_1 = 0 \\ N\left(\alpha_{21}^{(m)}\tilde{\omega}^{(m)} + \gamma^{(m)}\ln(\tilde{y}_1 + 1), (\sigma_{21\tilde{\imath}}^{(m)})^2\right) & \text{if} \quad \tilde{y}_1 > 0 \end{cases}.$$

6. Repeat Steps 1–5 across all $m = 1,2,\ldots,M$ MCMC draws, then gather all predictive draws $\{\tilde{\mu}_2^{(1)}, \tilde{\mu}_2^{(2)}, \ldots, \tilde{\mu}_2^{(M)}\}$, forming the posterior predictive distribution of $\tilde{\mu}_2$.

7. Denote the posterior median of $\tilde{\mu}_2$ by $\hat{\tilde{\mu}}_2$. Take $\hat{\tilde{\mu}}_2$ as the *raw* calibrated value of $\tilde{y}_1$, with uncertainty reported as a 68% posterior predictive credible interval for $\tilde{\mu}_2$ (i.e., the 16th and 84th percentiles of the posterior predictive distribution), which is akin to a frequentist estimate bracketed by ± 1 standard error.

8. The final calibrated value, $\hat{\tilde{y}}_2$, is obtained as follows: round the raw $\hat{\tilde{\mu}}_2$ value to the nearest integer, then replace it by the original value of $\tilde{y}_1$ only if the rounded value of $\hat{\tilde{\mu}}_2$ is less than $\tilde{y}_1$.

As an informal model diagnostic, the top panel of Figure 9.2.5 shows the results of Steps 1–7 above when applied to the 653 sets of $\{y_{ij1}, x_{1ij}, x_{2ij}, x_{3ij}\}$ in the training data; the figure is a plot of the resulting $\hat{\mu}_{ij2}$ against the observed values of y_{ij2}, with 68% credible intervals (CIs). Most CIs are near the 45° line in the figure, suggesting decent alignment between $\hat{\mu}_{ij2}$ and y_2. It turns out that for these training data, all cases of $y_{ij1} = 0$ yield 68% CIs for μ_{ij2} that run from 0 to 0, although the posterior standard deviation of μ_{ij2} in these cases are all positive (but less than 1).

Note that collectively, the CIs are akin to a frequentist 68% confidence band around a regression line, which refers to the uncertainty in $\hat{\mu}_{ij2}$ as opposed to $\hat{y}_2$, and thus it is not meant to capture 68% of the data. In fact, the focus of calibration is to make inference for μ_2 given an observed y_1 and covariates. Thus, we do not formally assess the ability for the calibration model to capture y_2. Diagnostics that are more formal than Figure 9.2.5 are discussed in Supplement Section 9.2.6.2.

9.2.5. CONCLUSION

In Chapter 9.2, we have developed a zero-inflated Poisson mixed-effects model to facilitate calibration of MaxN data from the scale of a single-camera SBRUV read to that of a spherical SBRUV read. Throughout the Bayesian model hierarchy, overdispersion is addressed using random effects, and covariates are incorporated in the model to enhance statistical inference for the mean of the underlying SBRUV's MaxN. The model is trained by fitting it to paired single-camera and spherical SBRUV data. The trained model that we adopt in practice is applied to single-camera data that are unpaired with spherical-camera data. For each data point, the resulting converted value is taken as the posterior median for the mean of the counterfactual spherical-camera MaxN, subject to minor adjustments such as rounding. Each converted value is accompanied by 68% posterior predictive credible intervals.

While the principles of the approach we employ in Chapter 9.2 can be generalized to other calibration settings, the specific model statements presented here are highly tailored to the West FL SBRUV data and covariates, and are not expected to perform well if directly transcribed to other calibration settings or datasets.

9.2.6. SUPPLEMENT

9.2.6.1 DIFFUSENESS OF BAYESIAN PRIORS

In the absence of concrete prior understanding about model parameters, we avoid the use of overly restrictive prior distributions that would heavily inform parameter inference by the observed data and minimally by the specified prior distributions. On the other hand, overly diffuse prior distributions for complex Bayesian models can cause computational instability in the form of MCMC mixing / convergence issues. In our case, we chose prior distributions that we believed *a priori* to be rather nonrestrictive, fitted the models, then diagnosed the actual restrictiveness of these priors by comparing the posterior distributions to their prior distributions.

Figure 9.2.6 shows that the posterior distributions for all but two parameters differ noticeably from their prior distributions, suggesting that we did not propose restrictive priors for these parameters. The exceptions are $\sigma_{\pi 0}$ and $\sigma_{\pi 1}$, whose posterior distributions only differ slightly from their prior distributions. This suggests that those two parameters are weakly identifiable, that is, the dataset provides limited statistical power towards the inference of those two parameters. One possible reason is that Level 2 of the model might not require random effects that we have employed to account for possible overdispersion of logit π. Though, for the context of fisheries surveys, we feel that allowing overdispersion throughout the model hierarchy constitutes good modeling practice. Future iterations of this calibration work can explore alternative ways to handle overdispersion without employing random effects for logit π.

9.2.6.2 MODEL COMPARISONS

We compare various perspectives of the model fit among the candidate models of Model A (Section 9.2.3.1), Model B (Section 9.2.3.2.1), and Model 0 (Section 9.2.3.2.2).

For our calibration models, we feel that standard goodness-of-fit metrics for Bayesian models, such as the deviance information criterion (DIC) or Watanabe-Akaike information criterion (WAIC), are unsuitable for model comparison. The reason is that these metrics focus on the model's predictive power with respect to the response data (y_1, y_2 in our case), when our goal here is to translate y_1 from the single-camera scale to the spherical scale of y_2. Thus, although both y_1 and y_2 are response variables in the calibration model, the ultimate purpose of the model is, as explained in Section 9.2.4, to make inference for μ_2 alone.

Instead of goodness-of-fit metrics, we consider the uncertainty in the inference for μ_2 across the entire calibration (training) dataset to convey the notion of "calibration reliability," as follows.

For any observation among the 653 in the training dataset, there is an unknown underlying value of μ_{ij2}. While Figure 9.2.5 may suggest that all Models A, B, and 0 perform similarly at the observation level, we inspect performance at the study level. To do so, we take the posterior standard deviation (SD) of μ_{ij2} for each (i, j)th observation, which can be regarded as a type of Bayesian standard error for $\hat{\mu}_{ij2}$. We gather the 653 values of posterior SDs, and consider their distribution collectively under each model; we log-transform the posterior SDs to deskew and aid visualization. The top panel of Figure 9.2.7 shows the corresponding distributions for Models A, B, and 0: it is clear that Model A yields the smallest variability for the posterior SD, which in turn reflects the reliability of the inference for μ_{ij2}. That is, at the study level, Model A performs best with respect to the reliability of the inference for μ_{ij2}. The bottom panel of Figure 9.2.7 suggests that this superiority of Model A is due to the very small posterior SD for $\{\mu_{ij2}|y_{ij1} = 0\}$ (recall the discussion from Section 9.2.4).

This, and the fact that $\hat{\mu}_{ij2}$ (posterior median of μ_{ij2}) is almost always greater than y_{ij1} for Model A, lead us to conclude that Model A is more appropriate than either of Models B and 0 for converting MaxN's from the single-camera read to the scale of the spherical-camera read.

FIGURES FOR CHAPTER 9.2



Figure 9.2.1: A video frame from an SBRUV deployment. The entire image corresponds to the spherical-camera read, and the portion bounded by the red box corresponds to the single-camera read.

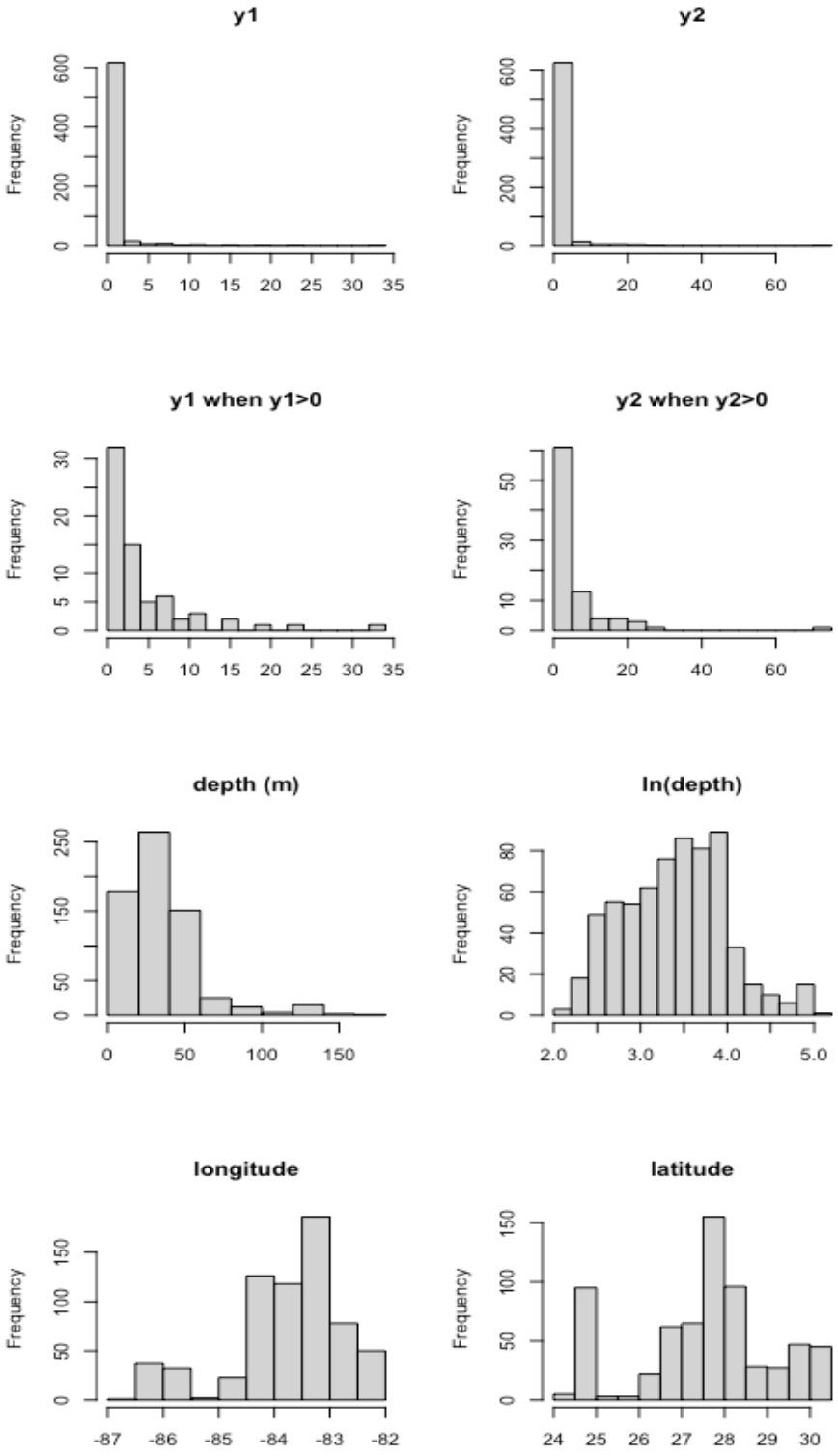


Figure 9.2.2: Histograms of numerical variables.

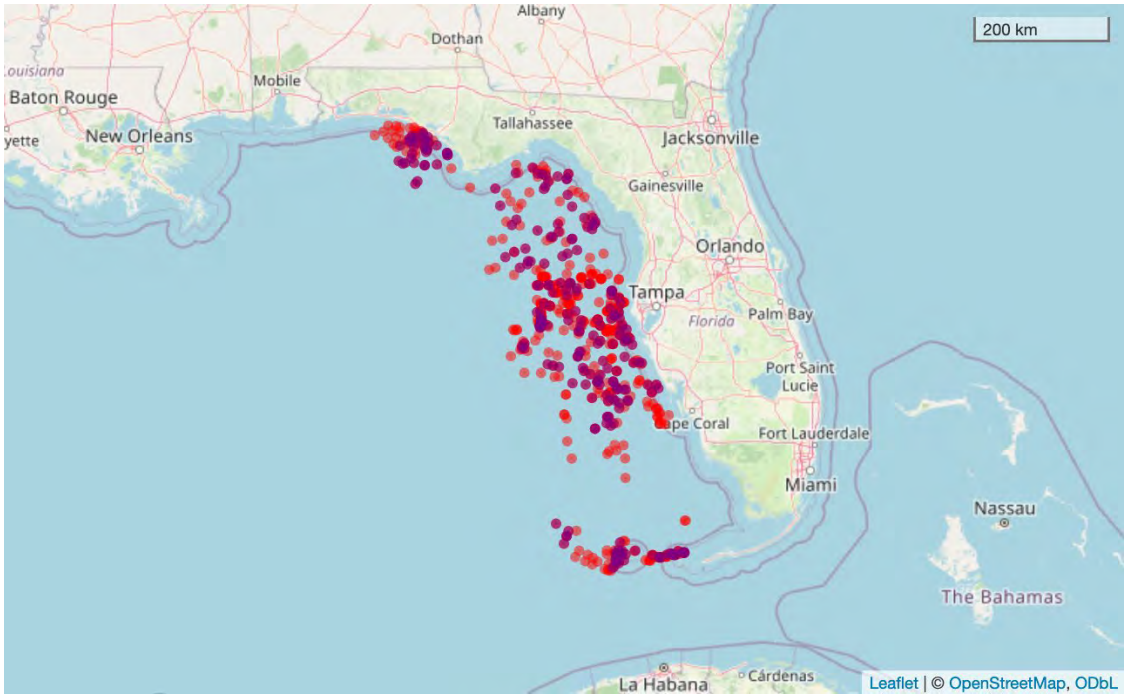


Figure 9.2.3: Locations of West FL calibration and survey data involving both single-camera and spherical SBRUV reads. Paired observations in 2020–2023 (red and purple dots) include those from the current GAC survey (purple dots) conducted in 2023.



Figure 9.2.4: Scatterplots and cross-correlations between numerical variables for cases where $y_1, y_2 > 0$. Colors denote habitat type (Artifical / Natural). Longitude and latitude exhibit high collinearity. Replacing them with their average leads to collinearity between this average and log-transformed depth, although it less severe than that between longitude and latitude (prior to averaging).

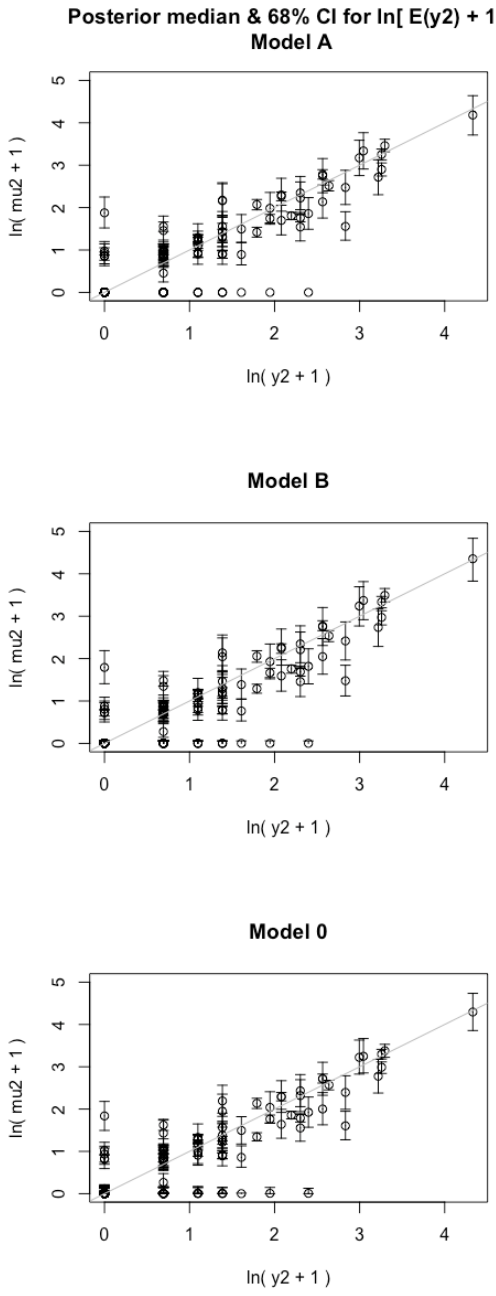


Figure 9.2.5: Calibration results for the training data in the form of a scatter plot of $\hat{\mu}_2$ (with 68% CIs) vs. y_2, both on the natural log scale to aid visualization. Values of $\hat{\mu}_2$ are obtained by applying Steps 1–7 in Section 9.2.4 to the 653 sets of $\{y_{ij1}, x_{1ij}, x_{4ij}\}$ from the training data; a 45° line is shown for reference. Top panel: Steps 1–7 directly refer to the final adopted Model A from Section 9.2.3.1. Middle panel: results for Model B from Section 9.2.3.2.1, after modifying Steps 1–7 to match the model statements for Model B. Bottom panel: results for Model 0 from Section 9.2.3.2.2, after modifying Steps 1–7 to match the model statements for Model 0.

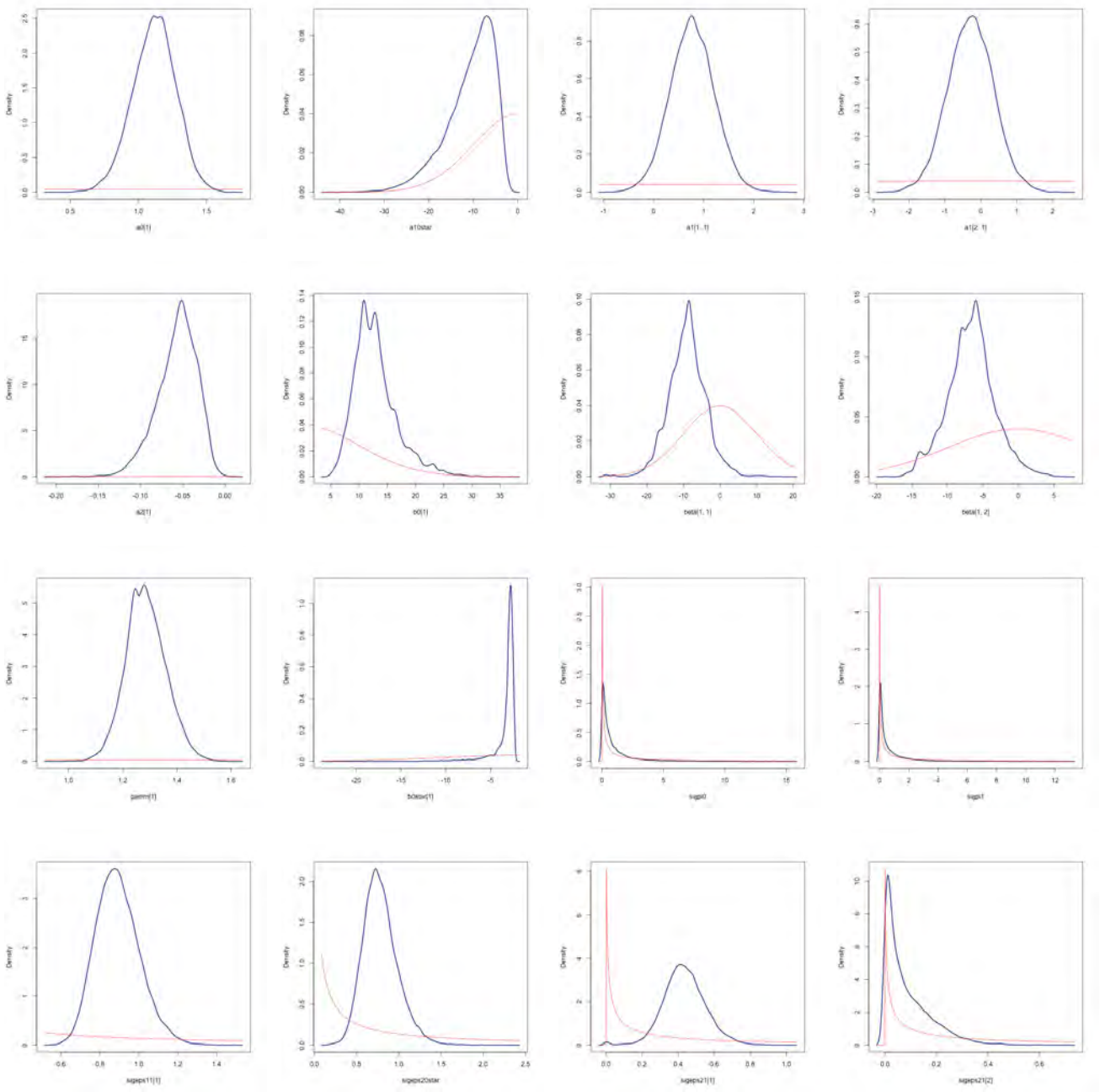


Figure 9.2.6: Posterior distributions (navy) and prior distributions (red) for the Level 4 parameters in the final adopted Model A from Section 9.2.3.1. Top row from left to right: $\alpha_{10}, \alpha_{10}^{*}, \alpha_{111}, \alpha_{112}$. Second row from left to right: $\alpha_{21}, \beta_0, \beta_1, \beta_2$. Third row from left to right: $\gamma, \beta_0^{*}, \sigma_{\pi 0}, \sigma_{\pi 1}$. Bottom row from left to right: $\sigma_{11}, \sigma_{20}, \sigma_{211}, \sigma_{212}$.

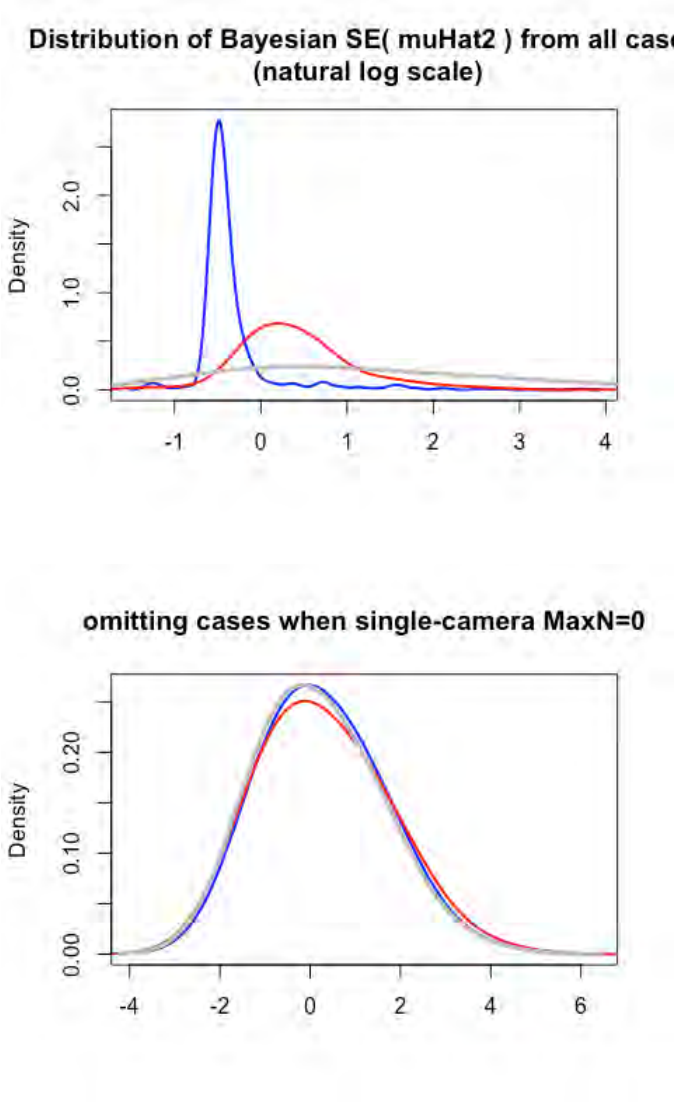


Figure 9.2.7: Distribution of Bayesian SE($\hat{\mu}_{ij2}$) on the natural log scale for Models A (blue), B (red), and 0 (gray). Here, "Bayesian SE($\hat{\mu}_{ij2}$)" refers to the posterior standard deviation of μ_{ij2} for each (i,j)th observation. Top panel: each curve is the histogram made up of the 653 values of $ln[SE(\mu_{ij2})]$ (all cases in the training data). Bottom panel: cases of $y_1 = 0$ omitted.

REFERENCES FOR CHAPTER 9.2

Bacheler, Nathan M., Kyle W. Shertzer, Zebulon H. Schobernd, and Lewis G. Coggins. 2023. "Calibration of Fish Counts in Video Surveys: A Case Study from the Southeast Reef Fish Survey." *Frontiers in Marine Science* 10 (April). https://doi.org/10.3389/fmars.2023.1183955.

de Valpine, P., C. Paciorek, D. Turek, et al. 2024. *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling*. R package version 1.3.0. https://cran.r-project.org/package=nimble.

Miller, Timothy J, Chhandita Das, Philip J Politis, et al. 2010. *Estimation of Albatross IV to Henry B. Bigelow Calibration Factors*. NOAA Northeast Fisheries Science Center reference document 10-05. https://repository.library.noaa.gov/view/noaa/3726.

Ponisio, Lauren C., Perry de Valpine, Nicholas Michaud, and Daniel Turek. 2020. "One size does not fit all: Customizing MCMC methods for hierarchical models using NIMBLE." *Ecology and Evolution* 10 (5): 2385–416. https://doi.org/10.1002/ece3.6053.

10. Abundance Estimation

10.1. Overview of Survey-Methods Approach

Lynne Stokes and Ellen Taylor

10.1.1 Introduction

A probability sampling methodology was used for data collection to enable estimation of greater amberjack (GAJ) abundance across the Gulf of America and South Atlantic. A probability sample is one in which locations of the universe of shelf waters are chosen for observation in such a way that each sampled location's probability of selection can be calculated. The theory of probability sampling then allows estimators of parameters of interest to be produced from the sample, and their properties evaluated.

In this project, the goal was to estimate abundance of Greater Amberjack (GAJ) and related totals (such as total within size classes) in all regions of the Gulf of America and South Atlantic. To do so, probability sample designs were executed independently in each region. Data was collected at each site in order to assess the number of GAJ present. One or both of the following methods was used to collect data at each site: deploying cameras to collect video that allowed visual counts to be made, or deploying echosounders on vessels that allowed acoustic counts of potential GAJ to be enumerated, as described in Chapter 9. Two separate estimates of abundance were made for each region, with each estimate using information from one of these data sources.

The sample designs and estimation procedures were implemented similarly across regions. In this chapter, we describe the typical sample design and give generic estimators for that design so that all regions' methods can be expressed as similarly as possible. However, every region's sample design had unique features required by their varying levels of habitat mapping information available and the characteristics of their region's habitat. In some cases, accommodations to designs were also required when data collection was shared with other ongoing surveys. In this chapter, we will describe the general method used for sampling and estimation of GAJ from both the camera and acoustic data. When these differences did occur, they are discussed in each region's section of Chapter *11.*

10.1.2 Sample design

Across the Gulf and South Atlantic, most data collections for camera and acoustic measurements were obtained from stratified two-stage random sample designs, where strata were defined by some combination of region, depth, and habitat type. Habitat type typically included the categories artificial and natural reefs, whose sizes and locations were known or estimated from previous studies. In the Western Gulf, counts of artificial reefs by type of structure were also known. All areas not known to be in one of these categories were designated as their own habitat type, referred to as unconsolidated or uncharacterized bottom.

In most regions, the primary sampling units (PSUs) for the two-stage design were defined as the cells of a grid defined across the sea floor. Because the grid cells were too large to be viewed by a camera, smaller units within each PSU, known as secondary sampling units (SSUs), were randomly chosen from each PSU for collecting counts of GAJ. The form of the SSUs varied depending on the gear and type of habitat. For natural reef or other continuous habitats, the SSUs were typically areal units chosen at randomly selected locations within the grid cell. For some artificial habitats, the SSUs were the structures themselves. One or more SSU's were chosen for sampling from each sampled PSU.

For camera samples, data were obtained by counting GAJ from video recorded at the sampled SSU sites. The area of the SSUs varied from region to region due to gear type, water characteristics and presence or absence of bait. SSU area was assessed by biologists in each region and calculated either theoretically based on characteristics of the gear and the way it was deployed, or by measurements made by the camera. In most regions, sensitivity analyses were conducted to evaluate the effect of the assessment of SSU area on abundance estimates. Discussion of findings from these investigations are in the regional sections of Chapter 12.

Data collection for the acoustic survey typically took place at a subset of the chosen camera SSU sites. However, the areas of the acoustic SSUs varied from site to site and were larger than those of the camera, as the vessel carrying the acoustic equipment traversed a back-and-forth pattern that included the camera's SSU site, but also some of the surrounding area. The area of the acoustic SSUs were measured by processing of the acoustic data.

In the Western Gulf, the artificial reef habitat was divided into several strata based on the type of structure (e.g., oil platforms, wrecks). The number of such structures was known for each grid cell. The SSUs were thus selected by randomly choosing structures within the sampled PSU grid cell. In that case, the SSU size was the number of structures observed by the camera and the habitat size within the PSU was the number of like structures present in the entire PSU, rather than their area. An exception to this rule was when the sampled structure was so large that it could not be observed completely by camera, such as for oil platforms. In that case, the collective platform areas were taken as habitat size within the PSU and the area observed was the SSU size.

10.1.3 Estimation

10.1.3.1 Overview

The general approach for estimation of GAJ abundance in all regions and for all habitats was to first estimate a density of GAJ over the habitat in the stratum, and then multiply this estimated density by the size of the habitat in the stratum. This provided an estimate of GAJ abundance for the habitat in the stratum. Finally, these abundance estimates were summed over all strata in the region. The regional chapters report both the stratum estimates as well as their sums, along with standard errors for each.

This type of estimator is known as a stratified ratio estimator in the sampling literature, since the estimated density (count of GAJ per habitat unit) is a ratio in which both numerator and denominator are random quantities observed from the sample. The density was defined as either GAJ per square kilometer or GAJ per structure, depending on the form of the SSU for the habitat being assessed. Likewise, the magnitude of habitat in the stratum will also be in the same units as the denominator of the density; i.e., either habitat area or total count of habitat structures in the stratum. We note that the amount of habitat in the stratum is treated as known in our estimates, though it will also be from habitat mapping programs prior to the current study. Some of the mapping programs, such as the one providing habitat size in West Florida, are also based on sample designs and have been ongoing for years. We justify treating these estimates of habitat size as known because the sample sizes producing their estimates are an order of magnitude larger than those of the current survey. Thus, their uncertainty will likely have a negligible effect on the standard error of the abundance estimator.

Calculation of estimators and their standard errors for all regions were carried out by the R package *survey* (Lumley 2010, 2026). This software handles analysis of survey data from complex probability samples, such as the stratified random multi-stage designs used in this project. Standard errors were calculated using the delta method, as implemented in the software.

All regions used sample designs with independent sample selection from one stratum to another. Therefore, the standard error for an estimator of total abundance over any set of aggregated strata can be computed as the square root of the summed estimated variances.

In some regions the estimates from the uncharacterized bottom were omitted from the regional sum due to the unreliability of the estimate for that stratum. This was due to high variability of habitat in the stratum, which sometimes consisted of unmapped areas containing unknown reef habitats in addition to areas without structure. When such reefs were encountered in samples of the stratum, it produced highly variable sample data, such as data containing a single large non-zero observation in its sample. This variability results in estimators with poor statistical properties.

10.1.3.2 Estimators and their standard errors

In this section we express the form of the most common estimator calculated among the regions. We first define the notation used for the estimators. Let $h = 1, \ldots, H$ denote the indices of the strata. The number of PSUs (i.e., grid cells) in the hth stratum's universe is denoted by N_h, and the number of those that are selected to be sampled is denoted by n_h. We denote the size of the habitat in the ith PSU of the hth stratum by X_{hi} and the size of the habitat in the jth sampled SSU from the ith sampled PSU as x_{hij}. The size of habitat in the entire h^{th} stratum universe is denoted by $X_{h0} = \sum_{i=1}^{N_h} X_{hi}$.

When the SSUs are areas, then X_{h0}, X_{hi}, and x_{hij} denote the total area of habitat in the hth stratum, in the ith PSU of the hth stratum, and the jth SSU of the ith PSU of the hth stratum, respectively. When the SSUs are structures, then X_{h0} and X_{hi} denote the number of structures in the stratum and PSU, respectively, and x_{hij}= 1, since a single structure is the SSU. The number of GAJ assessed to be present[2] in the jth sampled SSU of the ith PSU of the hth stratum is denoted by y_{hij}. The total number of GAJ in the ith PSU of the hth stratum is denoted by t_{hi} and the total number in the stratum is denoted by $t_h = \sum_{i=1}^{N_h} t_{hi}$.

The ratio estimator used for most regions and habitats for amberjack abundance t_h is then

$$\hat{t}_h = X_{h0}\hat{\delta}_h, \qquad \text{(Equation 10.1.1)}$$

where $\hat{\delta}_h$ is an estimator of GAJ density in the stratum. The density estimate is a ratio of weighted sums; i.e.,

$$\hat{\delta}_h = \frac{\sum_{i=1}^{n_h} w_{hij} y_{hij}}{\sum_{i-1}^{n_h} w_{hij} x_{hij}}.$$

The weights w_{hij} are called sample weights since they are based on the sample design. w_{hij} is calculated as the reciprocal of the probability of selection of the $(h, i, j)^{th}$ SSU. It can be interpreted as the number of unobserved SSUs that each observed (sampled) SSU represents in the universe.

These probabilities varied somewhat from region to region because the method of selection of the SSUs were not always the same. These differences occurred because the data collection operations for this project were sometimes conducted in collaboration with other survey programs with specified protocols that required specific sample sizes for particular habitats. However, a simple and common sample design used was to select a

[2] This assessment could be based on a calibrated value, a max N value, or an acoustically measured value.

simple random sample of PSUs from the universe of PSUs, and at the second stage to select a single SSU at random from all possible SSUs (areas or structures) in the sampled PSU. In that case,

$$w_{hij} = \left(\frac{N_h}{n_h}\right)\left(\frac{X_{hi}}{x_{hij}}\right).$$

The standard error of $\hat{t}_h$ was estimated using a common approximation for assessing variability of multi-stage estimators, as shown in Lohr (2022), Eq. (5.33). This approach approximates the variance arising from sampling in all stages with a component reflecting only the variance of PSU-to-PSU total estimates, which incorporates variability arising from subsequent stages because the PSU total is itself an estimate. This standard error estimator is the one provided in most survey analysis software, including the R package *survey* used in this project.

10.1.3.3 Estimation of abundance for acoustic counts

The sample design and estimators for estimates of abundance from acoustic data were similar to those of the camera-based abundance estimation. The sample design was identical, except that sample sizes were typically smaller, and sometimes much smaller, than those of the camera design. The PSUs that were included in the acoustic sample were a subset of the camera-measured PSUs, and its data was collected as closely as possible to the same time interval.

As noted previously, the SSUs for the acoustic sample were larger than those of the camera sample because of the ability of the equipment to collect data over larger areas. This made the probabilities of selection of the SSUs, and therefore the sampling weights, different from those of the camera estimates.

A major difference in the two methods, however, was the method of obtaining counts of GAJ from the sample. The acoustic data collection methods could not distinguish GAJ from other similar large-bodied fish, referred to as "potential *Seriola*" in Chapter 9. In order to obtain a count comparable to the camera-based GAJ, it was necessary to reduce the count of these potential *Seriola* to actual GAJ. The method used for this was to adjust the acoustic count downward by multiplying it by a fraction that represented the proportion of potential *Seriola* that are actually *Seriola*. This fraction is unknown, and undoubtedly varies from one region and habitat to another. Thus, it was necessary to estimate it from the camera data.

The video collected from the camera sample was processed to count both GAJ and the total of other large-bodied fish that would have been identified as potential *Seriola* by acoustic measurement. Thus, the fraction of potential *Seriola* that are actual *Seriola* at site (*h,i*), denoted by $\pi_{GAJ,hi}$ could be computed for each PSU sampled. However, for many PSUs, the number of fish observed was small. Therefore, some estimates of the fraction made from individual PSUs were highly variable. An alternative approach would be to aggregate camera counts across multiple PSUs, such as those in a stratum, and use the aggregated data to estimate a fraction. But this approach would produce a biased estimate if the true fraction varied across PSUs within the stratum.

To mitigate this problem, a method known as empirical Bayes estimation (Morris, (1983)) was used to estimate the fraction. The empirical Bayes model assumes that the proportion of GAJ across Stratum h has a beta distribution; i.e.,

$$\pi_{GAJ,hi} \sim Beta(\alpha, \beta)$$

and the number of GAJ and potential *Seriola* observed in a PSU ($y_{hi.}$ and $z_{hi.}$, respectively) have a binomial distribution,

$$y_{hi.}|z_{hi.} \sim Binomial\left(z_{hi.}, \pi_{GAJ,hi}\right).$$

Under this model, the empirical Bayes estimators of proportion for each PSU is estimated by estimating the parameters of the beta prior from the set of observed proportions and producing estimates of the proportions from the resulting posterior. These estimators have the form of weighted averages between the PSU estimates and the aggregated proportion estimate from the stratum; i.e.,

$$\hat{\pi}_{GAJ,hi} = \hat{\theta} * (\text{estimated proportion of GAJ among large} - \text{bodied fish in } PSU(h,i))$$
$$+\left(1-\hat{\theta}\right) * (\text{estimated proportion of GAJ among large} - \text{bodied fish in stratum } h),$$

where $\hat{\theta}$ is a weight between 0 and 1. The size of the weight $\hat{\theta}$ depends on the relative sizes of the estimated variances of the two proportion estimates. For example, the larger the number of potential *Seriola* ($z_{hi.}$) observed by camera in a PSU, the higher the weight on the PSU estimate. The aggregated proportion's weight will be larger when the variance of the beta prior is estimated to be small, indicating that the PSU proportions have little variability.

The advantage of empirical Bayes approach is that the resulting estimators have lower mean squared error than either the individual or aggregate estimates. The R software package glmmTMB (McGillycuddy 2025) was used to produce these estimates for all regions.

Once these estimates of proportion were produced, they were used as multipliers to adjust the acoustic estimates of potential *Seriola* to obtain predictions of GAJ estimates. Then these values were used as observed GAJ in the estimator shown in (Equation 10.1.1) to produce abundance estimates from acoustic data, just as they were for camera data.

Literature Cited

Lohr, S. (2022). Sampling Design and Analysis, 3rd Edition., Chapman and Hall/CRC Press.

Lumley T (2010). *Complex Surveys: A Guide to Analysis Using R: A Guide to Analysis Using R*. John Wiley and Sons.

Lumley T (2026). "survey: analysis of complex survey samples." R package version 4.5.

McGillycuddy M, Warton DI, Popovic G, Bolker BM (2025). "Parsimoniously Fitting Large Multivariate Random Effects in glmmTMB." *Journal of Statistical Software*, 112(1), 1–19. doi:10.18637/jss.v112.i01.

Morris, C. N. (1983). Parametric Empirical Bayes Inference: Theory and Applications. *Journal of the American Statistical Association*, *78*(381), 47–55. https://doi.org/10.2307/2287098

10.2. Overview of Bayesian Approach

Absolute Abundance Modeling for Natural and Artificial Habitats in Alabama-Mississippi Region Utilizing Camera-to-Acoustic Calibration Tool of Section 9.1

Grace S. Chiu,[1,2] *Anton H. Westveld,*[3,2] *Mark A. Albins,*[4] *Allison L. White*[5]

[1]*William & Mary's Batten School of Coastal & Marine Sciences, Virginia Institute of Marine Science; Department of Statistics, University of Washington; Department of Statistics and Actuarial Science, University of Waterloo*

[2]*Department of Statistical Sciences and Operations Research, Virginia Commonwealth University*

[3]*Research School of Finance, Actuarial Studies and Statistics, Australian National University*

[4]*Dauphin Island Sea Lab; Stokes School of Marine and Environmental Sciences, University of South Alabama*

[5]*Cooperative Institute for Marine and Atmospheric Studies, University of Miami; Southeast Fisheries Science Center, NOAA*

SUMMARY

Chapter 10.2 here develops a series of models that are used collectively to make integrative statistical inference for the region-wide absolute abundance of GAJ in the Alabama-Mississippi (AL-MS) region at non-platform artificial and natural structures. *Integrative* estimation and uncertainty bounds — hence, the coefficients of variation (CVs) — reflect data generating processes at the *system level* as a whole; thus, the statistical uncertainty in, say, a calibrated MaxN as a model output from Chapter 9.1 is accounted for when this value is used as an input value for the abundance interpolation model developed in Chapter 10.2, so that the uncertainty in the abundance model output inherently captures the uncertainty carried over from the calibration procedure. This uncertainty is further captured by the CV produced for the regional absolute abundance estimate that pools all location-specific abundance estimates across the region's sampling universe.

Conventional sampling design-based abundance estimation provides a list of numerical estimates and CVs at the stratum level, which does not facilitate any interpolation within or between the disjoint strata. In contrast, an abundance interpolation model both facilitates the prediction of abundance under different environmental conditions, and provides insights into the spatial distribution of abundance as a natural phenomenon. A model output with obvious practical utility is a map of the estimate or CV of any modeled quantity, e.g., density (abundance per ha).

For abundance interpolation, our methodology in Chapter 10.2 requires the knowledge of each target structure's areal footprint. Because platforms are unaccompanied by this information, the three data points observed at a platform are omitted from modeling. Modeling is also unavailable for non-structures (unconsolidated / uncharacterized bottom) in AL-MS because no GAJ was detected by the ROV at any of the 22 data points at non-structures, so that relationships cannot be established for a quantity that shows no gradient.

The part of the AL-MS region with existing habitat mapping data is in AL only, and it is partitioned into 446 grid cells with qualifying artificial or natural structures; cell-level mapping data are used as covariates in our models.

The AL-MS survey data at target structures comprise ROV and/or SBRUV and/or acoustic abundance data, alongside covariate data, all observed at the reef level. We only consider the 113 reef-level data points that involve either the ROV alone, or the ROV paired with the echosounder. Moreover, we omit the 4 reef-level data

points from MS, none of which is accompanied by information on the target's areal footprint. Further filtering results in a final dataset with 95 reef-level data points from AL only.

Three key assumptions throughout Chapter 10.2 are i) all fish observed at a target structure is regarded as a randomly drawn sample from the entire population of fish that live at the target; ii) the ROV's field of view (effective sampling area, or ESA) is taken to be 1.926405 ha throughout the regional survey; and iii) data on the structures' areal footprints are regarded as known. Because of assumption *i*, our model does not account for cross-habitat contamination in the sampling process. The ESA and footprint information in assumptions *ii–iii* are estimates that in fact vary across samples, although information about the uncertainty / variability in these estimates were not available at the time of modeling.

Our series of models are developed and fitted in five phases:

1. the original camera-to-acoustic cross calibration model described in Chapter 9.1;
2. the subsequent updated calibration model (also described in Chapter 9.1) with augmented data that include all deployments of {ROV, SBRUV, echosounder} from the AL-MS survey;
3. a reef-level Bayesian hierarchical model that relates the following quantities at each observed reef:
    - the observed ROV MaxN,
    - the calibration estimate of GAJ density per hectare sampled — and the associated uncertainty — both from applying the calibration tool of Phase 2 above,
        - "GAJ density per hectare sampled" refers to the number of GAJ in the hectare that is centered at the target structure, which includes the target and its surrounding non-habitat,
    - reef-level covariate data (externally imputed when missing),
    - the unknown GAJ density per hectare sampled,
    - the unknown probability of GAJ detection by the ROV camera,
    - the unknown probability of GAJ presence per hectare sampled;
4. various versions of a cell-level Bayesian hierarchical spatial model that relates the following cell-level quantities at each sampled cell:
    - the estimate of cell-level GAJ density per habitat-hectare, derived by aggregating reef-level estimate(s) within the cell, and scaling the result by the total footprint of the target(s) sampled from that cell,
    - the estimate of the probability of cell-level GAJ presence based on the reef-level model,
    - cell-level mapping data,
    - pairwise distance between the cell and other sampled cells;
5. various versions of a cell-level Bayesian hierarchical model that interpolates the results from Phase 4 to the entire universe of cells in AL with artificial and natural structures.

We compare three sets of final estimates of GAJ density and population size (absolute abundance), with CVs and/or Bayesian confidence intervals (CIs):

a. Version A based on a model variant from Phases 4–5,

b. Version B based on another model variant from Phases 4–5, and

c. naive (unrealistic) extrapolation of Phase 3 reef-level inference across the AL universe.

Resulting interactive maps of the estimates and CVs of cell-level densities and population sizes are produced based on Versions A and B, to facilitate the identification of "hot spots" — see Supplement A. Additionally, CIs and CVs of universe-level population sizes are reported for each of Versions A and B, as well as the naive extrapolation of the inference from Phase 3.

Roughly, the cell-level model-based population size estimates (CVs) from Versions A and B are around

- 35 000 (1.2) at artificial reefs —
    - 33 000 (1.3) at low-relief structures,
    - 1 800 (0.6) at high-relief structures,
- 1.9 million (1.2) at natural reefs.

With major caveats, the naive extrapolation of reef-level model-based population size estimates (CVs) are around

- 5 300 (0.5) at artificial low-relief reefs,
- 650 (0.7) at artificial high-relief reefs,
- 0.95 million (0.5) at natural reefs.

Major caveats include:

- The inherent uncertainty that is ignored by Assumptions *i–iii* above is unaccounted for. Had the relevant information on this uncertainty been available at the time of modeling, incorporating it into the final absolute abundance estimates would have been straightforward under the Bayesian inference paradigm.
- The CVs for population size estimates based on the naive extrapolation of reef-level inference from Phase 3 is unrealistically small, because it blatantly disregards any spatial variability in both the density of GAJ and the probability of GAJ presence that is evident in Phases 4–5.
- Proper uncertainty propagation would require all Phases 1–5 (at least Phases 3–4) to be incorporated into a single integrated modeling framework, but the scope of such an endeavor would be impractically large for the timeline of this project that was made available for modeling. Although we have made every effort to carry over uncertainty from one phase to the next in an ad hoc manner, concrete information is unavailable on the reliability of the CVs from Phases 4–5, except for our belief that they are much more reliable than the CVs from naively extrapolating the reef-level inference from Phase 3.

Additional details of our methods and associated caveats are discussed below.

10.2.1. INTRODUCTION

The main objective of Chapter 10.2 is to develop a statistical modeling inference approach for regional GAJ absolute abundance that addresses the following methodological challenges in a unified, holistic manner:

a. a camera-based MaxN at a target reef must be calibrated to the actual abundance (count) scale;

b. an observation by the acoustic echosounder at a target reef, while being a measure of actual abundance, must be corrected to "remove" the non-GAJ individuals from it;

c. covariate data must be used to make inference for habitat-specific GAJ density across the regional sampling universe, despite observing a very small number of reefs from the universe, so that

d. reef-level covariate data are missing at a vast majority of the sampling universe;

e. cell-level covariate data, available throughout the sampling universe within AL, must be reconciled with reef-level abundance data that are only available in a small number of cells;

f. the prevalence of reef-level data that are 0 must be explicitly modeled through the probability of GAJ absence; consequently, the model must allow for cell-level GAJ absence;

g. areal footprint data are required to convert habitat-specific GAJ density to a habitat-specific GAJ population size (absolute abundance);

h. prior to aggregation across the sampling universe, densities and population sizes should appear on a map to identify GAJ "hot spots."

The scope and complexity of a single, fully integrative model to address all of a–h above at the system level would be vast, given all the "moving parts" in the system that are simultaneously related to every other "moving part" in an integrative model.

Throughout the GAC project, the statistical modelers in the project (GSC and AHW) have relied on the synthesis data and calibration experimental data from Alabama to trial and prototype our inference approaches towards achieving this ambitious objective. As is the case with model development for any complex system, the overall system is partitioned into various components of manageable scope (thus, less complexity), then prototypes are developed for each individual component. Once individual components are in their final prototyping stage, they may be progressively combined towards reconstituting the single complex system. A final integrative model for the single system still requires rounds of refinement to allow "moving parts" from individual prototype components to integrate with each other to form a fully functioning single system. The original plan from the early stages of the GAC project was to finalize the integrative inference approach for AL-MS, then adapt this region's approach to other regions.

As it turned out, survey and mapping data were still being finalized for various regions and habitat types in the last two months of 2025, including the survey data required to update the calibration model of Chapter 9.1. By the time data had been finalized, many of our prototype components for AL-MS proved to be unsuitable for integrative final inference: because each prototype component was necessarily established for the data structure available at the time of prototyping, any slight changes to the data could render a prototype component obsolete, which in turn could render all downstream prototype components also obsolete.

With very limited time left to finalize the statistical modeling for the GAC project, we arrived at the following *roadmap* (comprising part of the Figure 1 flowchart) to mitigate the chain reaction of obsolescence that would have resulted from formal integrative modeling of the entire system:

- We regard the calibration modeling of Chapter 9.1 as a standalone project component that is not part of the overall integrative abundance modeling framework for any of the regions.

- For abundance modeling, we focus on the AL-MS region only, in hopes of arriving at an inference model and resulting abundance estimation that is as adequate as possible for this region within the remaining time frame.

- For model-based interpolation of GAJ density across the AL sampling universe, we rely on a cell-level model to leverage the completeness of cell-level covariate data across the universe. Thus, once a cell-level model is deemed adequate, it is considered the *final* cell-level model and is fitted to sampled cells; this model fit can "fill in the gaps" spatially by being applied to the covariate data on the unsampled cells to infer the corresponding cell-level GAJ densities.

- To facilitate cell-level inference for AL-MS, reef-level GAJ densities must be aggregated to the cell-level. Thus, a reef-level model is required; once a reef-level model is deemed adequate, it is considered the *final* reef-level model and is fitted to the observed reefs. Note that a large number of cells in the universe were unsampled, and so reef-level covariate data are missing entirely from them. As such, unlike the final cell-level model, the final reef-level model cannot be used to interpolate reef-level densities across the AL sampling universe. This explains why a reef-level model alone is insufficient for reliable inference of GAJ population size in AL-MS. Also note that for the sampled cells, a large majority of them had only one reef of a given habitat type observed, and occasionally, two were observed. In either case, aggregation across a sampled cell involves pooling the inferred reef-level population size across observed reefs of the same habitat type within the cell, dividing this total by the total areal footprint of those observed reefs, then multiplying the result by the total areal footprint of habitat within that cell.

- Modeling absolute abundance requires input data on the actual abundance (count) scale. Thus, reef-level MaxN and acoustic counts are fed into the calibration tool from Chapter 10.1, whose output is in turn fed into the reef-level model as response data. (Note that the calibration tool requires any acoustic count to be paired with a MaxN, although a MaxN can be unpaired.) Because areal footprint data were not available at the time of calibration modeling, the output from the calibration tool — and hence, from the reef-level model — is in the form of GAJ density per *hectare centered at the target reef*, instead of GAJ density per *hectare of habitat*. However, to facilitate inference of region-wide population size through the cell-level model, the latter type of density must be modeled to establish a relationship between cell-level density and covariates. Thus, the reef-level model output is aggregated to a cell-level density per *habitat-hectare* prior to use for cell-level modeling.

Because we prototype and diagnose the roadmap components individually, it is impossible to formally propagate inference uncertainty from one component to the next, with one exception: extending the fitted cell-level model for spatial interpolation is a fully integrative process. For the rest of the roadmap, we gave our best effort within the project time frame to incorporate elements in each model component that would allow some form of ad hoc uncertainty propagation. Details of this effort and other parts of our modeling roadmap are given in Section 10.2.2.

The remainder of Chapter 10.2 is structured as follows. In Section 10.2.3, we describe the data on which the calibration model is developed and trained. Section 10.2.4 describes the modeling at the reef level, of which the output is used as input for cell-level modeling that is described in Section 10.2.5. The section also presents the naive estimation of habitat-level GAJ abundance based on reef-level modeling. In Section 10.2.6, we describe the process of taking the adopted model from Section 10.2.5, which is fitted only to grid cells that are sampled, for spatial interpolation across the universe of grid cells. Section 10.2.7 presents the cell-level model-based estimation of habitat-level GAJ abundance. We provide concluding remarks in Section 10.2.8. Supplementary details appear in Section 10.2.9.

10.2.2. OVERVIEW OF INFERENCE APPROACH FOR ABSOLUTE GAJ ABUNDANCE IN AL-MS

Our inference model for the population size (absolute abundance) of GAJ in the AL-MS region mainly relies on ROV MaxN data and acoustic echosounder data (if available alongside the ROV) that were observed at non-platform artificial and natural structures from March 20 to November 29, 2023. Also used for modeling are mapping data across the 446 grid cells, each measuring just under 343 hectares, that make up the AL sampling universe; these were observed from 2011 to 2022. The key elements of the mapping data are the longitude and latitude of cell centroids, the number of reefs for each habitat type within the cell — categorized as a high- / low-relief artificial / natural structure — and the areal footprint of each mapped structure.

> *Three key assumptions throughout Chapter 10.2 are i) all fish observed at a target structure is regarded as a randomly drawn sample from the entire population of fish that live at the target; ii) the ROV's field of view (effective sampling area, or ESA) is taken to be 1.926405 ha throughout the regional survey; and iii) data on the structures' areal footprints are regarded as known. Because of assumption i, our model does not account for cross-habitat contamination in the sampling process. The ESA and footprint information in assumptions ii–iii are estimates that in fact vary across samples, although information about the uncertainty / variability in these estimates were not available at the time of modeling.*

Under the above assumptions, our modeling approach for region-wide inference of GAJ density (and hence, population size) comprise the following components, or *phases* in the Figure 10.2.1 flowchart, for each habitat type:

1. a cross-calibration tool, developed in Chapter 10.1, that converts any MaxN value (observed by a drop camera, SBRUV camera, trap camera, or ROV camera) to an estimated actual abundance per hectare centered at the target reef ($\hat{\phi}_0^{\text{sw}}$) — the tool is based on an integrative model that relates data from various cameras and the acoustic echosounder that were deployed simultaneously in a calibration experiment;

2. an updated calibration tool that additionally accounts for all AL-MS survey data points that involve the ROV, SBRUV, and acoustic gears simultaneously (also see Chapter 10.1);

3. a reef-level model that posits a relationship among the ROV MaxN (y) recorded at a target reef, the corresponding calibrated value on the actual abundance scale ($\hat{\phi}_0^{\text{sw}}$, by applying the Phase 2 calibration tool to the ROV MaxN or the accompanying acoustic count, if available), reef-level covariates ($\mathbf{x}_1$), the probability of GAJ presence per hectare centered at the target reef (p_0^{sw}), GAJ density per hectare centered at the target reef (φ_0^{sw}), regression coefficients, and various nuisance model parameters;

    - for naive extrapolation of reef-level results directly to the habitat level, φ_0^{sw} is aggregated across all observed targets of the same habitat type to produce a naive habitat-level density, which can be multiplied by the total habitat-level footprint across the universe for a "naively modeled" habitat-level GAJ population size for the AL universe;

4. a cell-level model that posits a relationship among the estimated cell-level GAJ density per habitat-hectare ($\hat{\tilde{\varphi}}$) that is a quantity derived from the output of the Phase 3 reef-level model, the estimated cell-level probability of GAJ presence per habitat-hectare ($\tilde{p}$) that is also a derived quantity from the Phase 3 output, the corresponding true unknown parameters (φ and p), cell-level covariates ($\mathbf{x}_2$), regression coefficients, and various nuisance model parameters;

5. a spatial interpolation model to facilitate the inference for cell-level GAJ density per habitat-hectare across the entire AL universe, adapted from the Phase 4 model that is established for sampled cells only;

6. the final inference for region-wide GAJ population size is based on first multiplying cell-level densities from Phase 5 by the total cell-level areal footprints, then totaling the products across all cells in the AL universe.

In the next section, we provide additional details about the survey and mapping data used in Phases 3–6.

10.2.3. DATA

10.2.3.1 AL-MS grid cells of mapped habitat

Figure 2 and Figure 3 show the locations of all 446 grid cells that comprise the universe of mapped habitat in AL. Each grid cell is approximately 343 hectares in area. As no mapping data exist for MS, the modeling and spatial interpolation in Chapter 11.2 ignore MS. That is,

> *MS is excluded entirely from our model-based inference for GAJ abundance in the AL-MS region.*

Mapping data that are candidate covariates for the cell-level model are

- longitude and latitude of cell centroids,
- the number of reefs for each habitat type within the cell — categorized as
    - either a high- or low-relief structure, and
    - either an artificial or natural structure,
- and the total areal footprint of each habitat type within the cell.

10.2.3.1.1 Cell-level covariate selection

We conduct exploratory data visualization of cell-level candidate covariates to identify the need for variable transformations, and any subsequent collinearity among candidate covariates, all before the start of modeling:

- Using the original cell-level reef count and areal footprint variables, we define new cell-level variables by aggregating them into *artificial*, *natural*, (without breakdown by relief) and *structure* (of either type). The original and newly defined variables are considered alongside each other as candidate covariates.

- Square-root transformation is applied to all cell-level reef counts and total areal footprints prior to visualization. The resulting variables are highly collinear with each other. We decide to keep cell-level total areal footprint for *structure* and discard all others, because it is a continuous variable, and it encompasses all habitat types.

- Even under a square-root transformation, this variable exhibits extreme right-skewness. As such, despite its lack of interpretability, we decide to use a tenth-root transformation on the variable (see Figure 4, top two panels).

- The final set of cell-level covariates, collectively denoted by the 4-dimensional vector $\mathbf{x}_2$, includes the constant 1 (for the intercept), cell-level longitude and latitude, and the tenth-root of the total areal footprint of structure.

10.2.3.2 AL-MS reef-level survey data

The entire survey dataset comprise data points observed at 202 target locations. As already justified in Section 10.2.1, the following data points are discarded from modeling:

- 47 observations that did not involve the ROV;

- 18 of the remaining observations taken under 0 visibility;

- 6 of the remaining observations taken in MS and at platforms, which have no areal footprint information,

- 21 of the remaining observations taken at non-structures,

- 15 of the remaining observations that were identified as potential outliers based on exploratory data visualization.

While omitted data points are unmodeled at the reef level, their home grid cells — if inside AL — are modeled at the cell-level via spatial interpolation.

The cells from which the final 95 data points originate appear in Figure 10.2.5.

10.2.3.2.1 Application of calibration tool

The calibration tool from Chapter 9.1 is applied to all ROV and acoustic counts from the final 95 data points (start of Phase 3 in the Figure 10.2.1 flowchart). The calibration values appear in Figure 10.2.6, and are computed as follows.

10.2.3.2.1.1 ROV unpaired with echosounder

For each of the 22 unpaired MaxNs, the calibration tool takes the MaxN as input, and produces the calibrated value $\hat{\phi}_0^{\text{sw}}$, as well as its standard error, $SE(\hat{\phi}_0^{\text{sw}})$. Both quantities are computed according to the Bayesian hierarchical calibration model-based approach described in Chapter 10.1.

Recall from Section 10.2.1 that $\hat{\phi}_0^{\text{sw}}$ is an estimate of GAJ density per hectare sampled (or swept), being the hectare centered at the target reef; the superscript “sw” refers to “swept.”

Note that $SE(\hat{\phi}_0^{\text{sw}}) = 81.346$ is a constant associated with the specific camera type (ROV in this case). Also note that the calibration formula has a positive intercept, so that $\hat{\phi}_0^{\text{sw}} > 0$ for these data points.

10.2.3.2.1.2 ROV paired with echosounder

For each of the remaining 73 sets of paired counts, let

- y and y_+ respectively denote the ROV's MaxN for GAJ and for GAJ-like fish identified as acoustically indistinguishable from GAJ (see Chapter 10.1 for the species list),
- ROV camera ratio $= r = \begin{cases} y/y_+ & \text{if} \quad y_+ > 0 \\ 0 & \text{if} \quad y_+ = y = 0 \end{cases}$,
- N_{EK} and A_{EK} respectively denote the acoustic count and associated surrounding area (in hectares) swept by the echosounder.

The calibration tool takes y and r as input, and produces

- $\hat{r}_p$: the linear-regression prediction for the pooled camera ratio, one that would have been computed by pooling all cameras' MaxN for GAJ, and all cameras' MaxN for GAJ-like fish, had ROV and other camera types been deployed alongside the ROV,
- $\hat{\sigma}_{\hat{r}_p}$: the prediction standard error (SE) of $\hat{r}_p$ based on the linear regression.

The final calibrated value and its SE are computed as

$$\hat{\phi}_0^{\text{sw}} = \hat{r}_p N_{EK}/A_{EK}\,, \quad \hat{\sigma}_{\hat{\phi}_0} = SE(\hat{\phi}_0^{\text{sw}}) = \hat{\sigma}_{\hat{r}_p} N_{EK}/A_{EK}\,.$$

Note that in 6 cases, $\hat{\phi}_0^{\text{sw}} = \hat{\sigma}_{\hat{\phi}_0} = 0$ due to $N_{EK} = 0$.

10.2.3.2.2 Reef-level covariate selection

Each reef-level observation in the final dataset is associated with the following candidate explanatory variables:

- visibility index (1 / 2 / … / 5),
- target's habitat type (artificial-low / artificial high / natural-low / natural-high),
- target's longitude and latitude,
- target's depth (m) from the surface,
- target's areal footprint (hectares), and
- temperature, salinity, and dissolved oxygen concentration (DO) during ROV deployment, where each of these variables is further broken down into a surface reading, middle reading, and bottom reading.

Other reef-level variables are *a priori* regarded as either being inapplicable due to their having little to no variability across the dataset, or being irrelevant / redundant based on context. The same reasoning applies to the omission of all cell-level covariates for reef-level modeling.

Exploratory data visualization of reef-level candidate covariates (prior to and after the removal of potential outliers) allows us to identify the need to redefine habitat strata, the need for variable transformations, and any collinearity among candidate covariates, all before the start of modeling:

- Of all 16 natural reefs observed, only 2 are of high relief. Thus, we combine the two relief types for all natural reefs to form a single "natural" habitat stratum to ensure explanatory power of the variable habitat.
- Square-root transformation is applied to each of the target's depth, its areal footprint, and the three DOs.
- Similar to its cell-level counterpart, reef-level areal footprint data exhibit severe right-skewness (but less so than the cell-level variable). Although a tenth root transformation reduces skewness better than a square-root (see Figure 10.2.4, bottom panels), the latter is kept for interpretability, in that area is on the scale of a squared measurement.
- The target's latitude and square-root depth are highly collinear — only the latter is kept for modeling due to its interpretability as a covariate.
- The three temperatures are highly collinear — only the middle temperature is kept for modeling due to its being *in the middle*.
- The middle salinity and the square-root of the bottom DO are highly collinear — the latter is chosen arbitrarily for modeling.

The final variables considered for reef-level modeling are:

- i denoting the target habitat stratum,
    - $i = 1$ for artificial-high,
    - $i = 2$ for artificial-low, and
    - $i = 3$ for natural,
- j denoting the sampled cell from the ith stratum,
- k denoting the observed target within the (i, j)th cell,
- y_{ijk} = ROV's MaxN at the (i, j, k)th target,
- $\hat{\phi}^{\text{sw}}_{0,ijk}$ = calibration tool output for the (i, j, k)th target, with standard error $SE(\hat{\phi}^{\text{sw}}_{0,ijk})$, and
- $\mathbf{x}_{1ijk}$ = the 11-dimensional vector of covariates comprising
    - the constant 1 (for the intercept),
    - visibility index (no missing data),
    - the target's longitude and its square-root transformed depth and areal footprint (no missing data for any of these),

- the three square-root transformed DOs (all missing for 20 cases),
- the middle temperature (missing for the same 20 cases), and
- the bottom salinity (missing for the same 20 cases) and surface salinity (missing for 14 of these 20 cases).

10.2.3.2.3 Missing data imputation for reef-level covariates

Covariate selection as described in Section 10.2.3.2.2 is based on pairwise relationships between variables prior to missing data imputation. The cases of $\mathbf{x}_{1ijk}$ with missing data are imputed externally using the multivariate imputation by chained equations (MICE) approach, implemented as the `mice::mice()` function in R (van Buuren and Groothuis-Oudshoorn 2011). We apply `mice::mice()` once (i.e., `mice(..., m=1)` and keeping other arguments at their default values) to the entire set of candidate covariates after variable transformation, but prior to variable selection. This ensures that missing data imputation accounts for the inter-predictability of candidate covariates post-variable transformation.

Note that proper uncertainty propagation would require the data imputation framework to be fully integrated with the overall abundance modeling framework. As explained in Section 10.2.1, the scope and complexity of such an integrative model would be impractical.

> *While not ideal, we disregard the uncertainty associated with the imputed covariate data due to practicality.*

10.2.4. PHASE 3 MODELING: REEF-LEVEL

At the (i, j, k)th target structure, let

- $\hat{\phi}^{\text{sw}}_{0,ijk}$ = calibration tool output,
- $\hat{\sigma}_{\hat{\phi}_{0,ijk}} = SE(\hat{\phi}^{\text{sw}}_{0,ijk})$,
- $x_{11,ijk}$ = visibility index,
- $\mathbf{x}_{1ijk}$ = 11-dimensional covariate vector (including $x_{11,ijk}$),
- y_{ijk} = ROV's MaxN,
- μ_{ijk} = expected value of y_{ijk},
- $\varphi^{\text{sw}}_{0,ijk}$ = GAJ density per sampled hectare centered at the target,
- $p^{\text{sw}}_{0,ijk} = P(\mu_{ijk} > 0)$ = probability that the ROV detects GAJ presence, and
- $v = 0.9693403$ = scaling factor that translates $p^{\text{sw}}_{0,ijk}$ to the probability of GAJ presence in the hectare centered at the target.

See Supplement 10.2.B for the rationale behind our value of v. The subscript '0' for some of the parameters is to distinguish reef-level quantities from cell-level quantities in Section 10.2.5.

We posit the following zero-inflated hierarchical model at the reef level:

$$
\begin{aligned}
y_{ijk} \mid \mu_{ijk} &\sim \text{Pois}(\mu_{ijk}) \\
\mathbb{1}\{\mu_{ijk} > 0 \mid p^{\text{sw}}_{0,ijk}\} &\sim \text{Bern}(p^{\text{sw}}_{0,ijk}) \\
\text{logit} p^{\text{sw}}_{0,ijk} \mid \boldsymbol{\alpha}_i, \kappa_i &\sim N\left(\mathbf{x}'_{1ijk}\boldsymbol{\alpha}_i, \frac{\kappa_i^2}{x_{11,ijk}}\right) \\
\text{If } \mu_{ijk} = 0: \quad & \varphi^{\text{sw}}_{0,ijk} = 0 \\
\text{If } \mu_{ijk} > 0: \quad & \begin{cases} \mathbb{1}\{\varphi^{\text{sw}}_{0,ijk} > 0 | \{\mu_{ijk} > 0\}, p^{\text{sw}}_{0,ijk}\} \sim \text{Bern}(v p^{\text{sw}}_{0,ijk}) \\ \ln \mu_{ijk} \mid \boldsymbol{\gamma}_i, \chi_i \sim N\left(\mathbf{x}'_{1ijk}\boldsymbol{\gamma}_i, \frac{\chi_i^2}{x_{11,ijk}}\right) \end{cases} \\
\text{If } \varphi^{\text{sw}}_{0,ijk} > 0: \quad & \begin{cases} \ln \varphi^{\text{sw}}_{0,ijk} \mid \boldsymbol{\beta}_i, \sigma_{\delta,i} \sim N(\mathbf{x}'_{1ijk}\boldsymbol{\beta}_i, \sigma^2_{\delta,i}) \\ \text{If } \hat{\phi}^{\text{sw}}_{0,ijk} > 0: \quad \ln \hat{\phi}^{\text{sw}}_{0,ijk} \mid \omega_0, \omega_1, \varphi^{\text{sw}}_{0,ijk} \sim \\ \qquad N\left(\ln \varphi^{\text{sw}}_{0,ijk}, \omega_0 + \omega_1 \frac{(\hat{\sigma}_{\hat{\phi}_{0,ijk}} + 10^{-12})^2 / (\varphi^{\text{sw}}_{0,ijk})^2}{x_{11,ijk}}\right) \end{cases}
\end{aligned}
$$

where $\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \boldsymbol{\gamma}_i$ are regression coefficient vectors, and $\kappa_i^2, \chi_i^2, \sigma^2_{\delta,i}, \omega_0, \omega_1$ are regression error variance parameters.

> *Our reef-level model assumes that, whenever GAJ is present within the ROV's field of view, then the ROV always sees at least one GAJ (thus, $\mu > 0$). This assumption is the basis for the above relationship between p_0^{sw} and φ_0^{sw}.*

Our hierarchical model is formulated in a way that MaxN data (y) inform the probability of GAJ presence (as reflected by p_0^{sw}), and MaxN and calibrated data ($\hat{\phi}_0^{\text{sw}}$) simultaneously inform GAJ density under GAJ presence (i.e., $\{\varphi_0^{\text{sw}} | \varphi_0^{\text{sw}} > 0\}$. Both y and $\hat{\phi}_0^{\text{sw}}$ appear in the model because often $y > 0$ but $\hat{\phi}_0^{\text{sw}} = 0$ due to a 0 acoustic count, in which case it is more reasonable to have p_0^{sw} be directly informed by $y > 0$ (visual evidence of presence) instead of by $\hat{\phi}_0^{\text{sw}} = 0$.

We model p_0^{sw} through a logit link, with overdispersion via the variance parameter κ, scaled by the inverse of visibility (high index value refers to high visibility) to reflect the reliability of y when visibility is high. The same scaling is applied to the variance parameters for $\ln\mu$ and $\ln\hat{\phi}_0^{\text{sw}}$, both of which are directly related to y.

Note that the regression of μ cannot take place when $\mu = 0$, which is controlled by the parameter p_0^{sw}. Similarly, the regression of φ_0^{sw} cannot take place when $\varphi_0^{\text{sw}} = 0$, which is controlled by both μ and p_0^{sw}. Finally, $\hat{\phi}_0^{\text{sw}}$ is only modeled when both $\hat{\phi}_0^{\text{sw}}, \varphi_0^{\text{sw}} > 0$, because on the natural log scale, it is assumed to have mean $\ln\varphi_0^{\text{sw}}$. Note that the regression of μ cannot take place when $\mu = 0$, which is controlled by the parameter p_0^{sw}. Similarly, the regression of φ_0^{sw} cannot take place when $\varphi_0^{\text{sw}} = 0$, which is controlled by both μ and p_0^{sw}. Finally, $\hat{\phi}_0^{\text{sw}}$ is only modeled when both $\hat{\phi}_0^{\text{sw}}, \varphi_0^{\text{sw}} > 0$, because on the natural log scale, it is assumed to have mean $\ln\varphi_0^{\text{sw}}$.

Moreover, when modeling the relationship between the $\hat{\phi}_0^{\text{sw}}$ data and the unknown GAJ density φ_0^{sw}, we wish to constrain the inference for φ_0^{sw} to reflect the reliability of the calibration tool (in addition to the reliability of y,

which is addressed by x_{11} in the variance statement for $\ln\hat{\phi}_0^{\text{sw}}$). Similar to Gneiting et al. (2005), our variance for $\ln\hat{\phi}_0^{\text{sw}}$ is modeled as a bias-corrected value of a theoretically derived variance for $\ln\hat{\phi}_0^{\text{sw}}$ (see Supplement 10.2.C). This formulation is our attempt to handle uncertainty propagation from Phases 1–2 (application of calibration tool) to Phase 3 here. By doing so, $\hat{\phi}_0^{\text{sw}}$ is not regarded by the reef-level model as unconstrained data.

10.2.4.1 Final Phase 3 model

Throughout Chapter 11.2, we implement our Bayesian models using the NIMBLE software (Ponisio et al. 2020; de Valpine et al. 2024); it produces Markov chain Monte Carlo (MCMC) draws from the joint posterior distribution (akin to the frequentist sampling distribution).

After numerous rounds of model fitting, diagnostics, and model refinement, our final reef-level model has the following constraints:

- No habitat effects:
    - $\kappa_i = \kappa_1$ for all $i = 2,3$,
    - $\chi_i = \chi_1$ for all $i = 2,3$,
    - $\alpha_{i\ell} = \alpha_{1\ell}$ for all $i = 2,3$ and $\ell = 1,2,8$ (intercept, longitude, square-root middle DO),
    - $\gamma_{i\ell} = \gamma_{1\ell}$ for all $i = 2,3$ and $\ell = 5,10$ (surface salinity, visibility index).
- No numerical covariate effects:
    - $\alpha_{i\ell} = 0$ for all ℓ for all $i = 4,6,7,9,10$ (middle temperature, bottom salinity, square-root top and bottom DO, visibility index),
    - $\beta_{i\ell} = 0$ for all ℓ for all $i = 4,6,7,8,9,10,11$ (middle temperature, bottom salinity, all square-root DO's, visibility index, square-root depth),
    - $\gamma_{i\ell} = 0$ for all ℓ for all $i = 4,6,7,11$ (middle temperature, bottom salinity, square-root top DO, square-root depth).

with prior distributions

$$\begin{gathered}\alpha_{i\ell} \sim N(0, 30^2) \text{ if unconstrained above,} \\ \beta_{i\ell} \sim N(0, 10^2) \text{ if unconstrained above,} \\ \gamma_{i\ell} \sim N(0, 20^2) \text{ if unconstrained above,} \\ \ln(\chi_1), \ln(\sigma_{\delta,i}) \sim N(0, 5^2) \text{ if unconstrained above,} \\ \kappa_1^{-2} \sim \text{Gamma}(0.5,1)\,, \\ \omega_0^{-1} \sim \text{Gamma}(1.1,1)\,, \\ \omega_1^{-1} \sim \text{Gamma}(5,1)\,.\end{gathered}$$

Figure 10.2.7 suggests that our prior distributions are not restrictive for these data, except for κ_1 and some α's — all are parameters at the logit level of the model. Their posterior and prior distributions do not differ greatly, suggesting rather weak identifiability for these parameters. Though, the α's in question are those that depend on i (habitat) that are required to remain in the model in order to address the interaction between the numerical covariate and habitat. One example is $\alpha_{i,\ell=3}$ associated with the target's square-root footprint; the

inference shows a rather strong signal for $i = 2$ (artificial-low) only, despite the extreme skewness of the covariate for this stratum (see Figure 10.2.4). Thus, the weak identifiability of $\alpha_{i,\ell=3}$ for $i = 1,3$ (artificial-high and natural) suggests that square-root footprint has little explanatory power for p_0 in those two habitat strata.

> *At this stage, we keep all weakly identifiable parameters in the model. If given more time for model development, we would consider alternative formulations of the reef-level model to improve identifiability.*

10.2.4.2 Aggregation to cell-level densities and probabilities

Reef-level inference for $\varphi^{\text{sw}}_{0,ijk}$ can be aggregated over k to estimate an underlying cell-level density for habitat i. We do so as follows.

Let $A_{0,ijk}$ denote the areal footprint of the (i,j,k)th target. Then, we take the true reef-level density to be

$$\varphi_{0,ijk} = \frac{\varphi^{\text{sw}}_{0,ijk}}{\min\{1, A_{0,ijk}\}}.$$

This definition aligns with our key assumption in Section 10.2.2, whereby all fish observed at a target are assumed to be associated with that target only. Thus, $\varphi^{\text{sw}}_{0,ijk}$ is the mean GAJ count within a hectare that is centered at the target. The density is therefore this mean count divided by the footprint if the footprint is under 1 hectare, or simply equal to the mean count if the footprint is greater than 1 hectare.

Now, we aggregate $\varphi_{0,ijk}$ over k by defining

$$\tilde{\varphi}_{ij} = \frac{\sum_{k=1}^{n_{ij}} A_{0,ijk}\, \varphi_{0,ijk}}{\sum_{k=1}^{n_{ij}} A_{0,ijk}}$$

where the tilde ("~") denotes a cell-level quantity extrapolated from the n_{ij} observed reefs within the (i,j)th cell. Recall that $n_{ij} = 1$ for many sampled cells, and is occasionally equal to 2. Because any given cell typically has many more than 2 reefs, the value of $\tilde{\varphi}_{ij}$ necessarily differs from the underlying (true) cell-level density, which we denote by φ_{ij}.

To carry over the reef-level inference from Phase 3 into the cell-level inference of Phase 4, we take posterior summaries of $\tilde{\varphi}_{ij}$, namely,

- $\hat{\tilde{\varphi}}_{ij}$ = posterior median of $\tilde{\varphi}_{ij}$ — see maps in Figure 10.2.5 and histograms in Figure 10.2.8, and
- $\tilde{p}_{ij}$ = posterior probability that $\tilde{\varphi}_{ij} > 0$ — see the histograms in Figure 10.2.8.

These quantities are used as model input in Phase 4 (see Section 10.2.5).

10.2.4.3 Naive extrapolation to habitat-level population size

Similar to the definition of cell-level density $\tilde{\varphi}_{ij}$ using reef-model inference, we define a habitat-level density as follows:

$$\tilde{\varphi}_i = \frac{\sum_{j \text{ sampled}} \sum_{k=1}^{n_{ij}} A_{0,ijk}\, \varphi_{0,ijk}}{\sum_{j \text{ sampled}} \sum_{k=1}^{n_{ij}} A_{0,ijk}}.$$

Next, let A_{ij} denote the total areal footprint of habitat i structures within any jth cell from the AL universe. Now, we let

$$\widetilde{\aleph}_i = \tilde{\varphi}_i \sum_{j \text{ in universe}} A_{ij}$$

represent the ith habitat-level GAJ population size.

Further, let $\widetilde{\widetilde{\aleph}}_i$ denote the posterior median of $\widetilde{\aleph}_i$, and $\widetilde{w}_i$ denote the half-width of the 68% credible interval (CI) for $\widetilde{\aleph}_i$, the latter being akin to a frequentist standard error. Then, we define a "coefficient of variation" (CV) for $\widetilde{\aleph}_i$ as

$$CV(\widetilde{\aleph}_i) = \frac{\widetilde{w}_i}{\widetilde{\widetilde{\aleph}}_i}.$$

Table 10.2.1 shows the 68% and 80% CIs for $\tilde{\varphi}_i$ and $\widetilde{\aleph}_i$ for each i, as well as the CVs for $\widetilde{\aleph}_i$.

10.2.5. PHASE 4 MODELING: CELL-LEVEL

> *Phase 4 modeling is developed for sampled cells only.*

Recall that the reef-level model from Phase 3 and the cell-level model in Phase 4 are not integrated into a single model, but instead are fitted separately. As stated in Section 10.2.4.2, we now carry over the Phase 3 inference into Phase 4 by taking two sets of posterior summaries associated with $\tilde{\varphi}_{ij}$, namely $\hat{\tilde{\varphi}}_{ij}$ = posterior median of $\tilde{\varphi}_{ij}$, and $\tilde{p}_{ij}$ = posterior probability that $\tilde{\varphi}_{ij} > 0$ (see Figure 10.2.8). These summaries are fed into Phase 4 as input (i.e., data) for the cell-level model.

First, some important points must be made to distinguish between quantities from the reef-level model in Phase 3 and the cell-level model in Phase 4, and between a model parameter and its posterior summary:

- Each (i, j)th combination has an underlying cell-level parameter φ_{ij}, which is the underlying (true) GAJ density in the (i, j)th cell; it is a parameter in the cell-level model, but not a parameter in the reef-level model. The same applies to the cell-level parameter p_{ij}, which is the underlying (true) probability of GAJ presence in the (i, j)th cell.

- To make proper cell-level inference for φ_{ij} and p_{ij} to facilitate spatial interpolation across the AL universe of sampled and unsampled cells, we need a model in Phase 4 to be fitted to cell-level data that are not part of Phase 3. That is, a Phase 4 posterior distribution for φ_{ij} and p_{ij} is required.

- The reef-level based quantity, $\tilde{\varphi}_{ij}$, is derived from reef-level model parameters. Thus, $\tilde{\varphi}_{ij}$ itself can be regarded as an estimate of φ_{ij}. Because reef-level model parameters are associated with a posterior distribution, any derived quantity based on the Phase 3 inference is also associated with the Phase 3 posterior distribution. This posterior cannot replace the Phase 4 posterior, which is only available after a Phase 4 model has been fitted.

- One way to feed Phase 3 inference results into Phase 4 is to use numerical summaries of the Phase 3 posterior distribution. Hence, we take $\hat{\tilde{\varphi}}_{ij}$ and $\tilde{p}_{ij}$. Note that $\tilde{p}_{ij}$ is a summary of the posterior distribution of $\tilde{\varphi}_{ij}$, and can be regarded as a Phase 3 estimate for p_{ij}.

Feeding Phase 3 posterior summaries into Phase 4 as cell-level model input is the reason why these phases are not fully integrated into a single modeling framework. For full integration of the two, each saved MCMC draw from the Phase 3 posterior should be fed into Phase 4 as model input. This requires fitting a Phase 4 model as many times as there are saved Phase 3 MCMC draws (tens of thousands). Doing so while relying on off-the-shelf MCMC software (NIMBLE in this case) is infeasible given the available project timeline, especially when a Phase 4 model is still under development, which requires the trialing of different model formulations and diagnosis of each trialed formulation.

Nevertheless, to account for the fact that Phase 3 posterior summaries are not true data, but instead are modeled output, we formulate Phase 4 in a way that attempts to capture the features of the Phase 3 posterior distribution.

First, we inspect the posterior distribution of $\tilde{\varphi}_{ij}$ and $\mathbb{1}\{\tilde{\varphi}_{ij}\}$ for several (i, j) combinations to ensure that we propose Phase 4 model statements that are sensible for Phase 3 results. Figure 10.2.9 showcases the posterior distribution of these two quantities for selected cells from each habitat stratum. For the artificial-low stratum, cells A-16 and B-274 show rather different posterior behavior: the former has a zero-inflated posterior for $\tilde{\varphi}$, explaining the high posterior incidence of GAJ absence. In contrast, the latter posterior for $\tilde{\varphi}$ is entirely above 0 (100% GAJ presence *a posteriori*). Meanwhile, the showcased cell from either the artificial-high stratum (A-6) or natural stratum (G-62) show an extremely skewed posterior for $\tilde{\varphi}$, but they are not zero-inflated (low posterior incidence of GAJ absence). The apparent gradient in posterior behavior in the artificial-low stratum as inferred in Phase 3, coupled with the extreme skewness and potential lack of gradient in behavior in the other two habitat strata, imply that a Phase 4 cell-level model must be flexible enough to capture these features.

Next, we consider incorporating Phase 3 inference uncertainties into Phase 4 variance parameters that are associated with the model input. We do so in a fashion that is similar to what is done at the end of Phases 1–2 and the beginning of Phase 3 (as described immediately before the start of Section 10.2.4.1). Specifically, we attempt to constrain the variance for $\hat{\tilde{\varphi}}$ and $\tilde{p}$ through theoretically derivations using the delta method, as well as sampling theory. See Supplement 10.2.C for details.

10.2.5.1 Formulations of cell-level models

We trialed two major formulations of a cell-level model, the first assumes spatial random effects (cell effects) for p_{ij}, while the second assumes them for φ_{ij}. Further, the first attempts to fully incorporate sampling theory-based variance parameters, although the resulting model turns out to be non-viable (the MCMC algorithm fails to run). This is due largely to the volatile theoretical behavior of $\tilde{\varphi}$, which is based on 1–2 observed targets from a cell with a population of structures that consists of only 1 reef to more than 12 000 reefs. Thus, $\tilde{\varphi}$ has a relatively small variance when the cell has few reefs, but its variance is vast when the cell is full of reefs.

As such, we proceed to consider the two model formulations without incorporating sampling theory, although the variance parameters are still constrained by theory through the delta method. This way, the Phase 3 results that are used as Phase 4 model input are not regarded as unconstrained data.

The resulting first formulation appears in Supplement Section 10.2.9.4. The model inference is unsatisfactory, as the fitted values of φ do not align with $\hat{\tilde{\varphi}}$ (Figure 10.2.10, right panel). In contrast, the fitted values from the second formulation are in good agreement with $\hat{\tilde{\varphi}}$ (Figure 10.2.10, left panel; note that what may appear to be a positive "bias" of cell-level estimates at low densities is immaterial to the inference of the habitat-level GAJ population size, because these estimates correspond to very small areal footprint in the corresponding habitat-cell combinations).

Below, we present the model statements of this second formulation.

10.2.5.2 A viable cell-level model with spatial autocorrelation

For the (i, j) sampled cell, in addition to $\hat{\tilde{\varphi}}_{ij}, \varphi_{ij}, \tilde{p}_{ij}$, and p_{ij}, let

- $\mathbf{x}_{2ij}$ = 4-dimensional covariate vector (see Section 10.2.3.1.1),
- $d_{j,j'}$ = distance between the jth and j'th sampled cells,
- $\mathbb{R}$ = spatial autocorrelation matrix,
- ρ = exponential decay parameter in the spatial correlation function, and
- ξ_{ij} = expected value of φ_{ij} when $\varphi_{ij} > 0$.

We posit the following zero-inflated hierarchical model, with spatial random effects among cell-level mean densities:

$$\text{Data informed:} \quad \begin{cases} \text{logit } \tilde{p}_{ij} \mid \zeta, p_{ij} \sim N\left(\text{logit } p_{ij}, \dfrac{\zeta^2}{p_{ij}(1-p_{ij})}\right) \\ \text{logit } p_{ij} \mid \boldsymbol{\psi}, \nu \sim N(\mathbf{x}'_{2,ij}\boldsymbol{\psi}, \nu^2) \end{cases}$$

$$\mathbb{1}\{\varphi_{ij} > 0 \mid p_{ij}\} \sim \text{Bern}(p_{ij})$$

$$\varphi_{ij} \big| \{\varphi_{ij} > 0\}, \xi_{ij} \sim \text{Exp}(\text{mean} = \xi_{ij})$$

$$\text{If } \varphi_{ij} > 0: \quad \begin{cases} \ln \xi_{ij} = \mathbf{x}'_{2,ij}\boldsymbol{\lambda}_i + \epsilon_j + \theta_{ij} \\ \{\epsilon_j \mid \eta, \rho\} \sim MVN(\mathbf{0}, \eta^2 \mathbb{R}), \quad R_{jj'} = \text{Cor}(\epsilon_j, \epsilon_{j'} \mid \rho) = e^{-d_{jj'}/\rho} \\ d_{jj'} = \text{distance between } j, j' \\ \theta_{ij} \mid \chi \sim N(0, \chi^2) \end{cases} \tag{1}$$

$$\text{Data informed:} \quad \begin{cases} \text{If } \hat{\tilde{\varphi}}_{ij}, \varphi_{ij} > 0\text{:} \\ \hat{\tilde{\varphi}}_{ij} \mid \varphi_{ij} \sim \text{Exp}(\text{mean} = \varphi_{ij}) \end{cases}$$

where $\boldsymbol{\psi}$ and $\boldsymbol{\lambda}_i$ are regression coefficient vectors, and ζ, ν, η are regression error variance parameters.

The logit-level model parameters' lack of dependence on habitat (i.e., their lack of the subscript i) is based on the diagnosis from the unsatisfactory model (see Supplement Section 10.2.9.4). This model assumption also reflects the lack of gradient in $\tilde{p}_{ij}$ as seen in Figure 10.2.8.

When GAJ is present, the cell-level densities φ_{ij} and $\hat{\varphi}_{ij}$ are modeled as exponentially distributed, instead of being log-normal. This model assumption is based on the histograms of $\tilde{\varphi}$ (partially shown in in Figure 10.2.9) and $\ln\tilde{\varphi}$ (not shown). We model $\hat{\varphi}_{ij}$ to have mean φ_{ij}, and φ_{ij} to have a mean that exhibits spatial autocorrelation through a spatial random effect (ϵ_j). The mean of φ_{ij} has an additional source of random error (θ_{ij}), allowing habitat-cell interaction on the overdispersion for φ_{ij}.

10.2.5.2.1 Final Phase 4 model

After numerous rounds of model fitting, diagnostics, and model refinement, our final Phase 4 model has the following constraints:

- $\chi_i = \chi$ for all i,
- $\psi_{\ell=2} = \lambda_{i,\ell=2} = 0$ for all i (no cell latitude effects on either GAJ density or presence probability).

Prior distributions are

$$\ln\zeta, \ln\rho, \ln\chi \sim N(0,1)\,,$$
$$\ln\eta \sim N(0, 2^2)\,,$$
$$\nu^{-2} \sim \text{inverse-Gamma}(1.1,1)\,,$$
$$\psi_\ell, \lambda_{i\ell} \sim N(0, 10^2) \text{ if unconstrained above.}$$

Figure 10.2.11 suggests that our prior distributions are not at all restrictive for these data, as the posterior distribution differs noticeably from the prior distribution for all parameters.

Figure 10.2.12 shows the posterior medians for φ in a map.

This model is now applied for spatial interpolation across the AL universe of grid cells.

10.2.6. PHASE 5: ADOPTION OF PHASE 4 FOR SPATIAL INTERPOLATION

For spatial interpolation, we refit the final Phase 4 model (Section 10.2.5.2.1) to all cells, as follows.

The spatial correlation structure from Phase 4 is tailored to the 56 cells that were sampled out of the universe of 446. To allow spatial autocorrelation to be extended to all 446, the Phase 4 model is re-expressed to include all unsampled cells, whereby $\hat{\tilde{\varphi}}_{ij}$ and $\tilde{p}_{ij}$ are specified as *missing data* for all j that is unsampled. The Bayesian hierarchical structure of the universe-wide model allows missing data to be regarded as unknown model parameters and thereby imputed as part of model inference. As spatial autocorrelation is assumed across all 446 in the universe, the resulting inference for the spatial autocorrelation parameters are affected by the imputed values, and thus may differ substantially from those parameter values in Phase 4 which are tailored to the sampled cells only.

In fact, although the autocorrelation structure in Phase 4 is one of the least complex structures for spatial phenomena, it turns out to be non-viable at the universe level — the model does not converge. MCMC results (not shown) suggest that η and ρ are jointly unidentifiable. As such, we reformulate the Phase 4 model in two

simpler, spatially uncorrelated versions — Models A and B below — and fit them to the universe data. Again, we do so by relying on the internal missing data imputation mechanism of Bayesian model inference.

> *Interpolation across the AL universe via missing data imputation of Bayesian model inference means that Phase 4 uncertainty is properly propagated into Phase 5. However, recall that Phase 3 uncertainty is only partially accounted for in Phase 4 via the delta method.*

10.2.6.1 Model A: Uncorrelated spatial random effects

Letting j extend to all cells in the AL-universe, Model A here replaces Equation 10.2.1 with

$$\text{If } \varphi_{ij} > 0: \quad \begin{cases} \ln\xi_{ij} = \mathbf{x}'_{2,ij}\boldsymbol{\lambda}_i + \epsilon_j + \theta_{ij} \\ \epsilon_j \mid \eta \sim N(0, \eta^2) \\ \theta_{ij} \mid \chi \sim N(0, \chi^2) \end{cases} \qquad (2)$$

thus removing all spatial correlation among the cell effects ϵ_j. All parameter constraints prior distributions remain the same as that in Section 10.2.5.2.1.

The results from Model A are shown as maps in Figure 10.2.3 (for ϵ_j) and Figure 10.2.13 (for φ_{ij} and CV(φ_{ij}), where CV is defined similarly as in Section 10.2.4.3).

The need for a random cell effect can be assessed by the intraclass correlation coefficient,

$$r = \frac{\eta^2}{\eta^2 + \chi^2},$$

whose 10th, 50th, and 90th posterior percentiles are {0.01, 0.19, 0.78}. While the posterior median for r appears low, rather high values (above 60%) are within 80% credibility. Thus, cell effects are not entirely irrelevant.

Now, the inference for cell-level density φ_{ij} can be directly translated to that for cell-level population size $\aleph_{ij} = A_{ij}\varphi_{ij}$, where A_{ij} is the total footprint of habitat type i inside cell j. The results are mapped in Figure 10.2.14 (for $\aleph_{ij}$ and CV($\aleph_{ij}$)).

10.2.6.2 Model B: Overdispersion without cell effects

Model B here further simplifies Model A by replacing Equation 10.2.2 with

$$\text{If } \varphi_{ij} > 0: \quad \begin{cases} \ln\xi_{ij} = \mathbf{x}'_{2,ij}\boldsymbol{\lambda}_i + \theta_{ij} \\ \theta_{ij} \mid \chi \sim N(0, \chi^2) \end{cases}$$

The results from Model B are shown as maps in Figure 10.2.15 (for φ_{ij} and CV(φ_{ij})) and Figure 10.2.16 (for $\aleph_{ij}$ and CV($\aleph_{ij}$)).

10.2.7. PHASE 6: HABITAT-SPECIFIC TOTAL GAJ ABUNDANCE

Based on the Watanabe-Akaike information criterion (WAIC), Models A and B have equal predictive performance, with WAIC = 266. Thus, we proceed to produce inference for habitat-level population size using both models. In either case, the ith habitat level population size parameter is

$$\aleph_{i+} = \sum_{j \text{ in universe}} \aleph_{ij} .$$

Table 10.2.2 shows the 68% and 80% CIs for $\aleph_{i+}$ and CV($\aleph_{i+}$), for each i.

Posterior quantile values from Model A tend to be slightly higher than those for Model B. CVs are comparable between the two, but that for artificial-low is 13% higher for Model A.

10.2.8. FINAL REMARKS

A great deal of effort has been devoted to modeling this substantially complex system. The complexity is due not only to the fact that making model-based inference for the total number of individuals of a fish species that live in the entire Gulf region has never been attempted prior to the GAC project. It is also due to the extreme sparsity and inherent uncertainty in the data, in that model-based inference is attempted for a sampling universe with limited habitat mapping data, neither camera data nor acoustic data reflect the underlying quantity of interest, both sampling gears can perform poorly under low visibility and other environmental conditions, the sampling approach relies on an imprecise definition of "fish population" that associates highly mobile individuals to static objects that often cannot be comprehensively observed with the sampling gears, fish data are observed at 0, 1, or 2 reefs from inside a grid cell that may contain 1 to tens of thousands of reefs, and only about one tenth of all grid cells are associated with fish data.

With very limited time available for modeling, our modeling effort has been devoted to the AL-MS region only, while handling the above challenges with as much rigor as it is practical under these circumstances. Usable data towards modeling originate from AL reef structures only, including habitat mapping data that are required for spatial interpolation across the universe. As such, all inference in Chapter 10.2 refers to reef structures across the mapped universe in AL. Survey data observed at non-structures are all 0s, and therefore cannot be modeled.

The enormous amount of moving parts in the AL-MS system have been handled in 6 phases, most of which are standalone, unconnected components. Proper uncertainty propagation from one phase to the next is attempted in a rather ad hoc manner, and uncertainty propagation is simply ignored from the step of imputing missing covariate data that is part of Phase 3; thus the final uncertainty estimates in the form of CVs are not entirely reliable. However, the model-based spatial interpolation results from Phases 4–6 are considered a great deal more reliable than those from Phase 3, which entirely ignores the spatial features and their influences on GAJ density and GAJ presence/absence.

An ideal abundance interpolation model would be a single, truly integrative model constructed at the system level, so that CVs would be at the system level, being rigorously propagated across all sources of uncertainty associated with all data from the calibration experiment and the regional survey (MaxN and acoustic counts, habitat footprints, camera ESA, etc.). Future iterations of the work presented in Chapter 10.2 may include incorporating the uncertainty associated with the ROV's field of view and the areal footprint of habitat structures; developing and refining formulations of a truly integrative model that can address weak parameter identifiability; developing bespoke MCMC algorithms to better allow proper uncertainty propagation (although

such an endeavor may require effort on the scale of a PhD dissertation); and adapting the finalized modeling framework to other regions in the GAC project.

10.2.9. SUPPLEMENTS

10.2.9.1 SUPPLEMENT A: Interactive maps from AL-MS abundance modeling for Artificial & Natural habitats only

All interactive maps can be downloaded as a single HTML file. Once downloaded, the HTML file is best viewed using the Chrome web browser.

10.2.9.2 SUPPLEMENT B: Derivation of v

Because the ROV's field of view is approximated by $A_{ROV} = 1.926405 (> 1)$, we assume that, if GAJ is truly present within the ROV's field of view (i.e., if $\mu > 0$), then

- the ROV always sees at least one GAJ,
- the probability for $\{\varphi_0^{sw} > 0\}$ is v times p_0^{sw}, where $v < 1$,
- the ROV's detection intensity can be represented by the Gaussian density function, so that in the case of $1 = P(\mu > 0) \approx$ area under the $N(0,1)$ curve between ± 3, then

$$v = P(\varphi_0^{sw} > 0) = P(\varphi_0^{sw} > 0 | \mu > 0)$$

= area under the $N(0,1)$ curve between $\pm 3 \times \frac{\text{radius}_{\text{hectare}}}{\text{radius}_{\text{ROV}}}$

= area under the $N(0,1)$ curve between $\pm \frac{3}{\sqrt{A_{ROV}}}$

= 0.9693403.

Thus, we take $v = 0.9693403$ in general.

10.2.9.3 SUPPLEMENT C: Derivation of variance terms using the delta method

10.2.9.3.1 Phase 3 variances

We theoretically derive $\text{Var}(\ln\hat{\phi}_0^{sw})$ for Phase 3 using the delta method based on $\text{Var}(\hat{\phi}_0^{sw}) = \hat{\sigma}_{\hat{\phi}_0}^2$. Thus,

$$\text{Var}(\ln\hat{\phi}_0^{sw}) \approx \frac{\hat{\sigma}_{\hat{\phi}_0}^2}{[E(\hat{\phi}_0^{sw})]^2}.$$

We further assume that $E(\hat{\phi}_0^{sw}) \approx \varphi_0^{sw}$ to get

$$\text{Var}(\ln\hat{\phi}_0^{sw}) \approx \frac{\hat{\sigma}_{\hat{\phi}_0}^2}{(\varphi_0^{sw})^2}.$$

For computational numerical stability, we add the offset value 10^{-12} to $\hat{\sigma}_{\hat{\phi}_0}$. Finally, like Gneiting et al. (2005), we take $\omega_1 \frac{(\hat{\sigma}_{\hat{\phi}_0} + 10^{-12})^2/(\varphi_0^{sw})^2}{x_{11}}$ as the regression error variance with bias, where ω_0 is the bias correction.

10.2.9.3.2 Phase 4 variances

In Phase 4, $\tilde{p}_{ij}$ and $\hat{\tilde{\varphi}}_{ij}$ are "data". Similar to the Phase 3 variances above, we assume

$$\text{logit}\,\tilde{p}_{ij} \sim N\left(\text{logit}\,p_{ij}, \zeta_{0i} + \frac{\zeta_1}{p_{ij}(1-p_{ij})}\right),$$
$$\hat{\tilde{\varphi}}_{ij} \big| \hat{\tilde{\varphi}}_{ij} > 0 \sim \text{Exp}(\text{mean} = \tilde{\xi}_{ij}),$$
$$\text{Var}(\ln\tilde{\xi}_{ij}) = \tau_{0i} + \tau_1 \frac{C_{ij}}{\varphi_{ij}^2}.$$

Our rationale is as follows.

First, by the delta method, we have

$$\text{Var}(\text{logit}\,\tilde{p}_{ij}) \approx \frac{1}{p_{ij}(1-p_{ij})},$$

and applying the reasoning of Gneiting et al. (2005) yields $\zeta_{0i} + \zeta_1/[p_{ij}(1-p_{ij})]$.

Next, we wish to transfer the reef-level model uncertainty that is inherent in $\hat{\tilde{\varphi}}_{ij}$ directly to Phase 4. Thus, we take

$$\text{Var}(\ln\tilde{\xi}_{ij}) \approx \text{Var}(\ln\hat{\tilde{\varphi}}_{ij} | \hat{\tilde{\varphi}}_{ij} > 0) \approx \frac{\text{Var}(\hat{\tilde{\varphi}}_{ij} | \hat{\tilde{\varphi}}_{ij} > 0)}{[E(\hat{\tilde{\varphi}}_{ij} | \hat{\tilde{\varphi}}_{ij} > 0)]^2}$$

by the delta method.

Dropping the condition "$|\hat{\tilde{\varphi}}_{ij} > 0$" from the notation to reduce clutter, we further take

$$\text{Var}(\hat{\tilde{\varphi}}_{ij}) \approx \text{Var}(\tilde{\varphi}_{ij}), \qquad E(\hat{\tilde{\varphi}}_{ij}) \approx \tilde{\varphi}_{ij} \approx \varphi_{ij}.$$

Next, let $\tilde{\tau}_i$ denote the theoretical variance of $\{\varphi_{0,ijk}^{\text{sw}} : k = 1, \ldots, n_{ij}\}$ for all j within i. Also let M_{ij} denote the total number of reefs in the (i,j) cell. Based on random sampling theory and the fact that all $A_{0,ijk} < 1$ in our data, we have

$$\text{Var}(\tilde{\varphi}_{ij}) = \frac{\text{Var}(\sum_{k=1}^{n_{ij}} \varphi_{0,ijk}^{\text{sw}})}{(\sum_{k=1}^{n_{ij}} A_{0,ijk})^2},$$
$$\text{Var}\left(\sum_{k=1}^{n_{ij}} \varphi_{0,ijk}^{\text{sw}}\right) = \sum_{k=1}^{n_{ij}} \sum_{k'=1}^{n_{ij}} \text{Cov}\,(\varphi_{0,ijk}^{\text{sw}}, \varphi_{0,ijk'}^{\text{sw}}),$$
$$\text{Cov}(\varphi_{0,ijk}^{\text{sw}}, \varphi_{0,ijk'}^{\text{sw}}) = -\frac{\tilde{\tau}_i}{M_{ij}-1} \quad \text{for } k \neq k'$$
$$\Longrightarrow \quad \text{Var}(\tilde{\varphi}_{ij}) = C_{ij}\tilde{\tau}_i$$
$$\text{where} \quad C_{ij} = \begin{cases} 0 & \text{if } M_{ij} = 1 \\ \dfrac{n_{ij}(M_{ij}-n_{ij})}{M_{ij}-1} & \\ \dfrac{}{(\sum_{k=1}^{n_{ij}} A_{0,ijk})^2} & \text{if } M_{ij} > 1 \end{cases}.$$

Gathering terms, we have

$$\text{Var}(\ln\tilde{\xi}_{ij}) \approx \frac{C_{ij}}{\varphi_{ij}^2},$$

and applying the reasoning of Gneiting et al. (2005) yields $\tau_{0i} + \tau_1 \frac{c_{ij}}{\varphi_{ij}^2}$.

10.2.9.4 SUPPLEMENT D: An unsatisfactory cell-level model

The first formulation of our cell-level model is:

$$\text{Data informed:} \quad \begin{cases} \text{logit } \tilde{p}_{ij} \sim N\left(\text{logit } p_{ij}, \zeta_{0i} + \dfrac{\zeta_1}{p_{ij}(1-p_{ij})}\right) \\ \text{logit } p_{ij} = \mathbf{x}'_{2,ij}\boldsymbol{\psi}_i + \epsilon_j + \theta_{ij} \\ \{\epsilon_j\} \sim MVN(\mathbf{0}, \eta^2 \mathbb{R}), \quad R_{jj'} = \text{Cor}(\epsilon_j, \epsilon_{j'}) = e^{-d_{jj'}/\rho} \\ d_{jj'} = \text{distance between } j, j' \\ \theta_{ij} \sim N(0, \nu_i^2) \end{cases}$$

$$\begin{gathered} \mathbb{1}\{\varphi_{ij} > 0\} \sim \text{Bern}(p_{ij}) \\ \varphi_{ij} | \varphi_{ij} > 0 \sim \text{Exp}(\text{mean} = \xi_{ij}) \\ \text{If } \varphi_{ij} > 0: \qquad \ln \xi_{ij} = \mathbf{x}'_{2,ij}\boldsymbol{\lambda}_i \end{gathered}$$

$$\text{Data informed:} \quad \begin{cases} \hat{\tilde{\varphi}}_{ij} | \hat{\tilde{\varphi}}_{ij} > 0 \sim \text{Exp}(\text{mean} = \tilde{\xi}_{ij}) \\ \text{If } \hat{\tilde{\varphi}}_{ij}, \varphi_{ij} > 0: \qquad \ln \tilde{\xi}_{ij} \sim N\left(\ln \xi_{ij}, \tau_{0i}\right) \end{cases}$$

where the variance term for $\ln \tilde{\xi}_{ij}$ is reduced from the formulation $\tau_{0i} + \tau_1 C_{ij}/\varphi_{ij}^2$ (derived in Supplement C) which proved to be non-viable (see Section 5.1).

The finalized version of the model that yields the results in the right panel of Figure 10 has the following parameter constraints and prior distributions:

- $\zeta_{0i} = 0$ for all i,
- $\nu_i = \nu$ for all i,
- $\tau_{0i} = \tau_0$ for all i,
- $\boldsymbol{\psi}_i = \boldsymbol{\psi}$ for all i,
- $\psi_2 = \psi_4 = 0$ (no effects on presence probability from cell latitude and tenth-root of footprint of structures),
- $\lambda_{i2} = 0$ for all i (no cell latitude effects on GAJ density when GAJ is present),

prior distributions are

$$\begin{gathered} \ln \zeta_1, \ln \tau_0, \ln \eta, \ln \rho \sim N(0,1), \\ \nu^{-2} \sim \text{inverse-Gamma}(1.1,1), \\ \lambda_{i\ell}, \psi_\ell \sim N(0, 10^2) \text{ if unconstrained above.} \end{gathered}$$

TABLES FOR CHAPTER 10.2

Table 10.2.1: Reef-level inference for habitat-level density and population size. Percentages refer to posterior quantiles. Thus, 50% refers to $\hat{\tilde{\varphi}}_i$ or $\hat{\tilde{\aleph}}_i$, and 16% and 84% together form a 68% CI for $\tilde{\varphi}_i$ or $\tilde{\aleph}_i$.

		Density			Population size	
	Artificial-low	Artificial-high	Natural	Artificial-low	Artificial-high	Natural
10%	280.4348	276.4691	779.2520	3438.670	268.7924	499479.4
16%	308.8624	333.6561	904.0531	3787.247	324.3915	579473.5
50%	431.7573	667.2890	1486.1402	5294.175	648.7603	952575.5
84%	699.4837	1235.7613	2537.9273	8577.015	1201.4478	1626742.5
90%	893.0201	1470.4159	3027.8269	10950.142	1429.5868	1940754.8
CV				0.45	0.68	0.55

Table 10.2.2: Phase 5 modeling results: cell-level inference for universe-wide population size. Percentages refer to posterior quantiles. Thus, 50% refers to the posterior median of $\aleph_{i+}$, and 16% and 84% together form a 68% CI for $\aleph_{i+}$. The last column under each model refers to inference for pooling both strata of artificial structures.

	Model A				Model B			
	Artificial-low	**Artificial-high**	**Natural**	**Artificial**	**Artificial-low**	**Artificial-high**	**Natural**	**Artificial**
10%	14302.42	981.8884	620746.3	16659.77	13316.54	969.3831	609994.0	15574.99
16%	16746.17	1139.0792	782192.6	18921.48	15416.95	1097.1067	774172.7	17791.97
50%	34487.17	1888.1659	1935971.4	36828.32	29956.74	1835.4807	1935177.9	32500.73
84%	106936.98	3438.7575	5585995.1	108626.49	85007.31	3327.9991	5514164.3	87022.50
90%	155179.45	4274.1440	8071892.0	157611.17	125447.30	4074.0756	7705205.2	128499.97
CV	1.31	0.61	1.24	1.22	1.16	0.61	1.22	1.07

FIGURES FOR CHAPTER 11.2

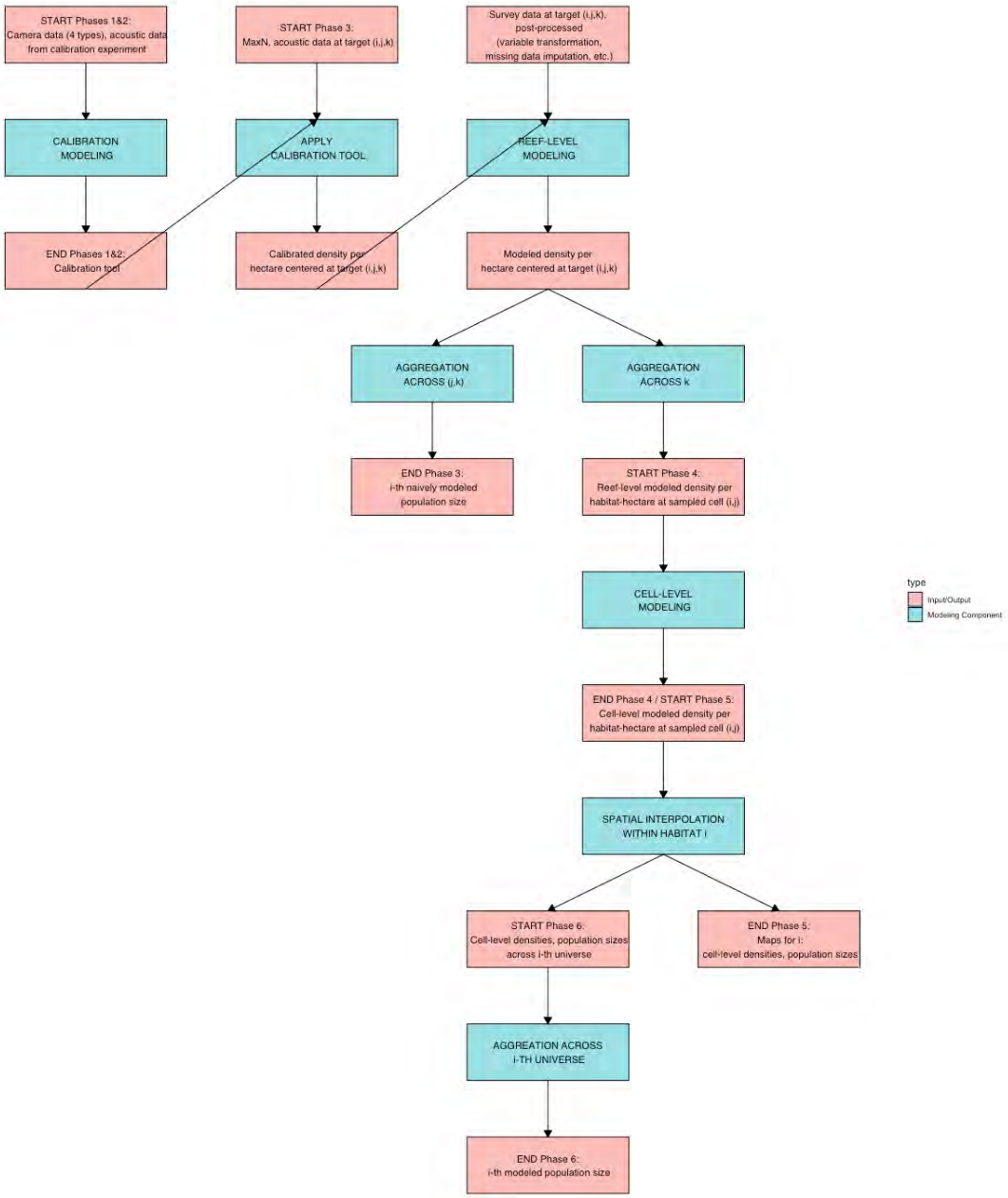


Figure 10.2.1: Roadmap of modeling components for the AL-MS region, where i denotes habitat type, j denotes sampled cell, and k denotes observed target structure.

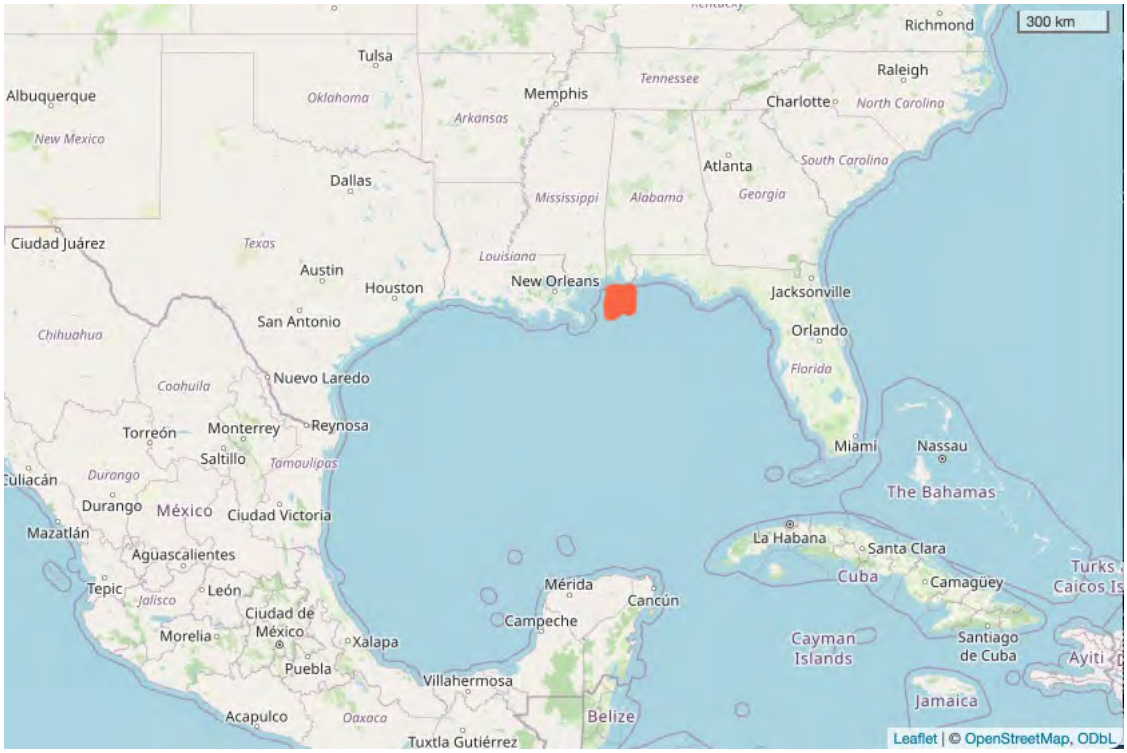


Figure 10.2.2: Location of the universe of mapped habitat in the AL-MS region, shown in red.

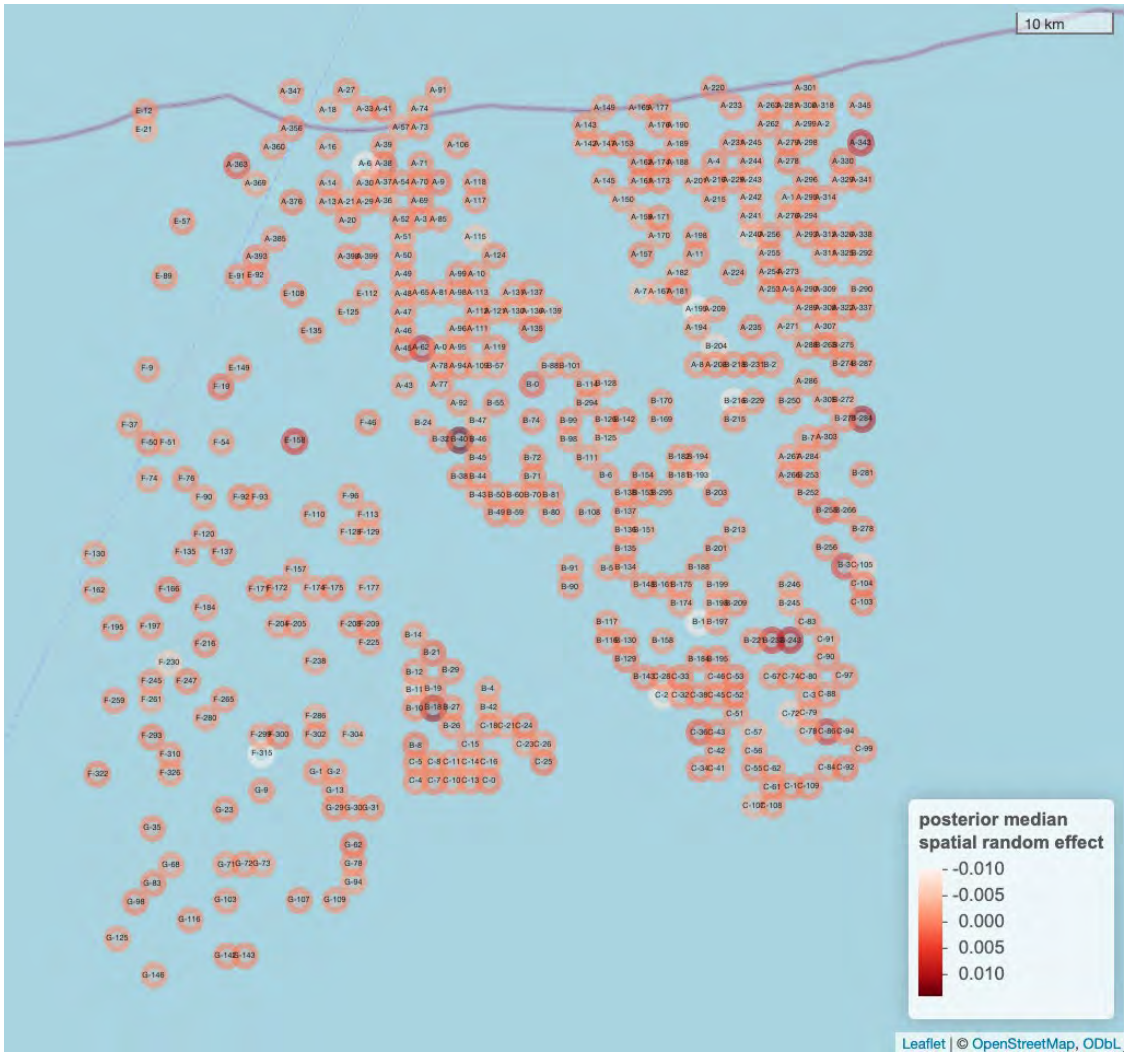


Figure 10.2.3: Map of estimated spatial random effects (posterior medians) from Model A. Each bubble on the map represents a grid cell in the universe of mapped habitat in the AL-MS region, labeled with the cell ID (which can be viewed in high resolution from the interactive maps under Supplement 10.2.A).

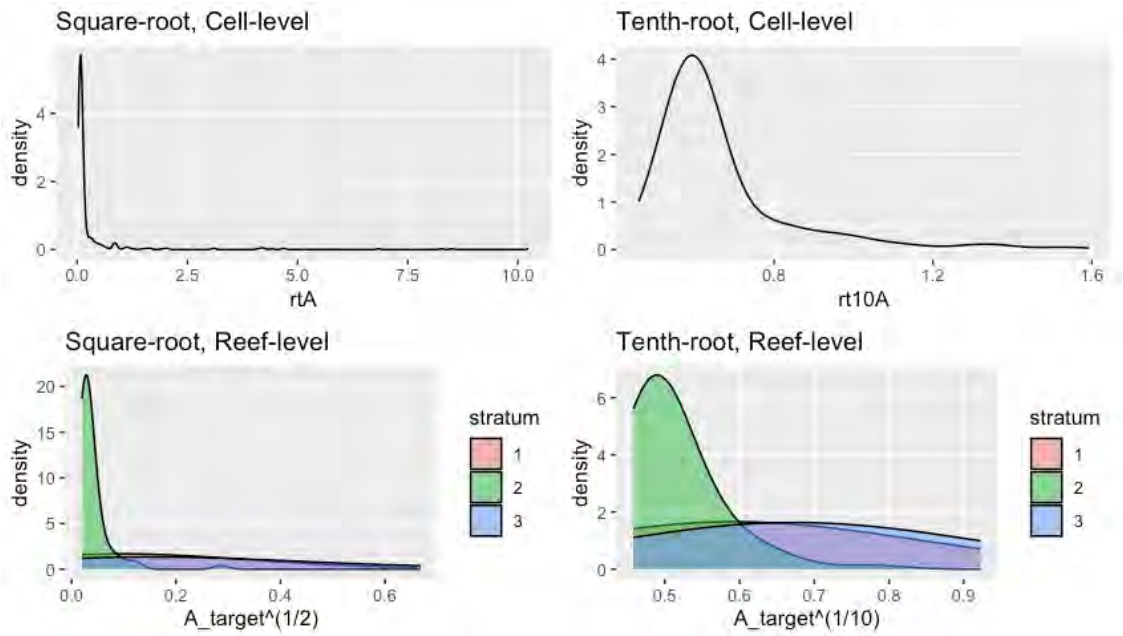


Figure 10.2.4: Distribution of areal footprint, after power transformation. Top panels: Cell-level total areal footprint of structure (either artificial or natural). Bottom panels: Reef-level areal footprint, broken down by habitat type (stratum): '1' for artificial-high, '2' for artificial-low, and '3' for natural.

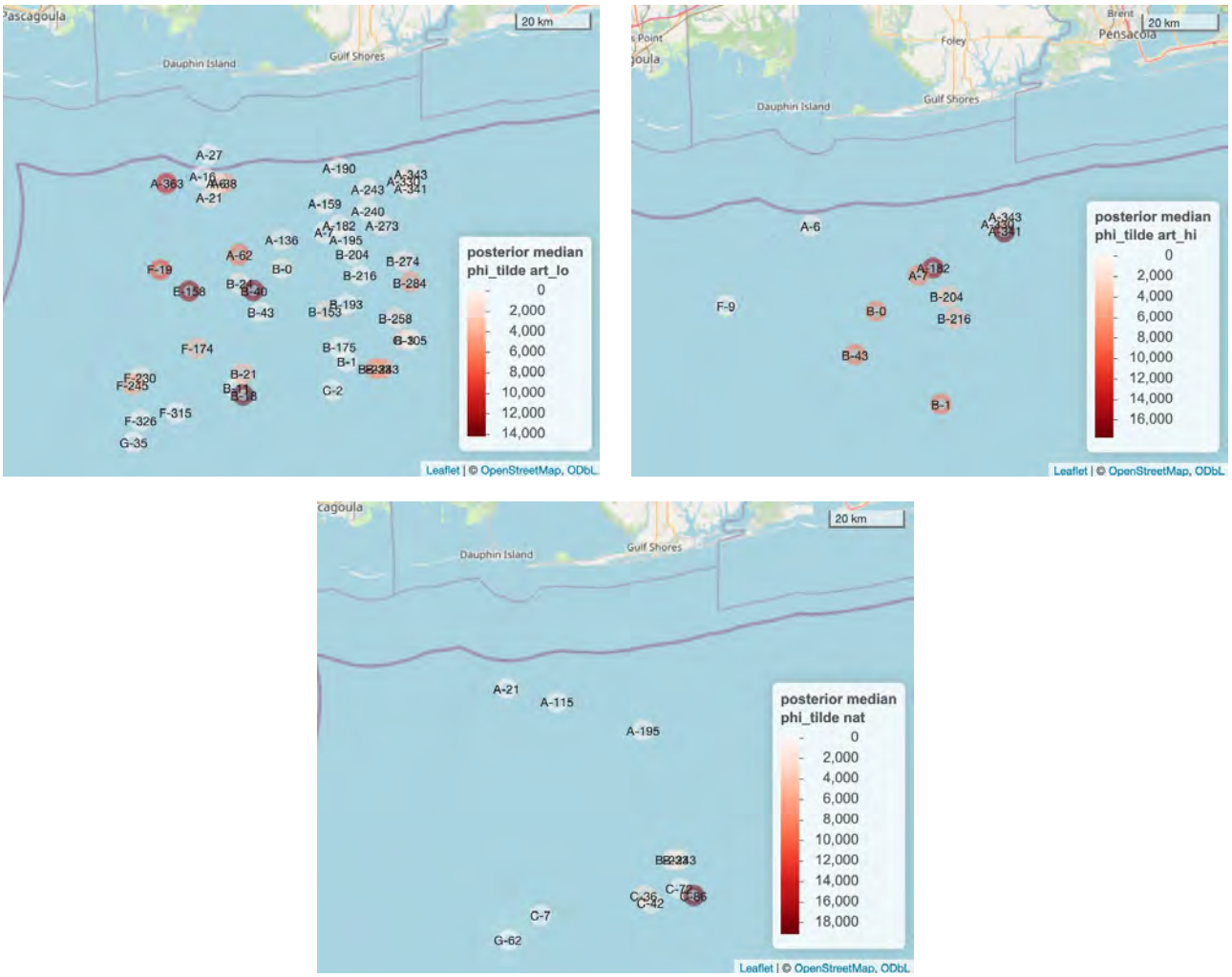


Figure 10.2.5: Reef-level modeling results: estimated cell-level GAJ densities (posterior medians) by habitat type. Bubbles with cell IDs correspond to grid cells from which reef-level data originate. Top-left panel: Low-relief artificial structures (67 reefs from 48 cells). Top-right panel: High-relief artificial structures (12 reefs from 12 cells). Bottom panel: Natural structures (16 reefs from 11 cells). See Supplement 10.2.A for interactive versions of these maps. See Figure 10.2.8 for the same estimates presented as histograms.

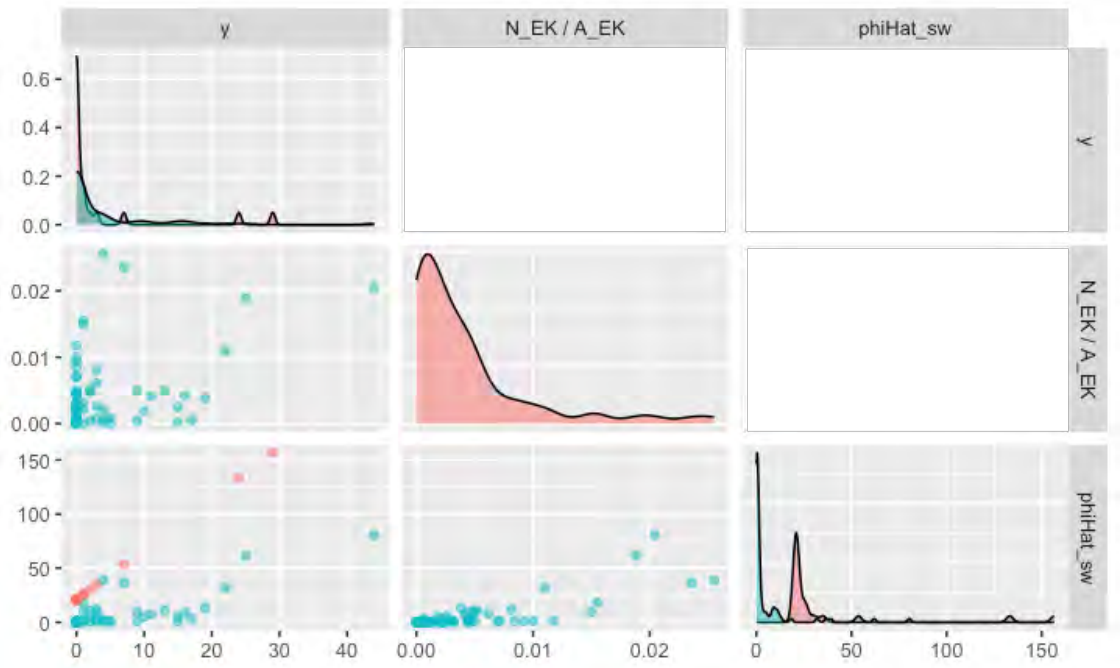


Figure10.2. 6: Calibrated abundance data by applying the calibration tool of Chapter 10.1 to the 95 reef-level data points. Variables from left to right / top to bottom: MaxN (y), acoustic-based density (N_{EK}/A_{EK}), and calibration tool output ($\hat{\phi}_0^{\text{sw}}$). Color denotes whether the ROV was paired with the echosounder (blue for YES, red for NO).

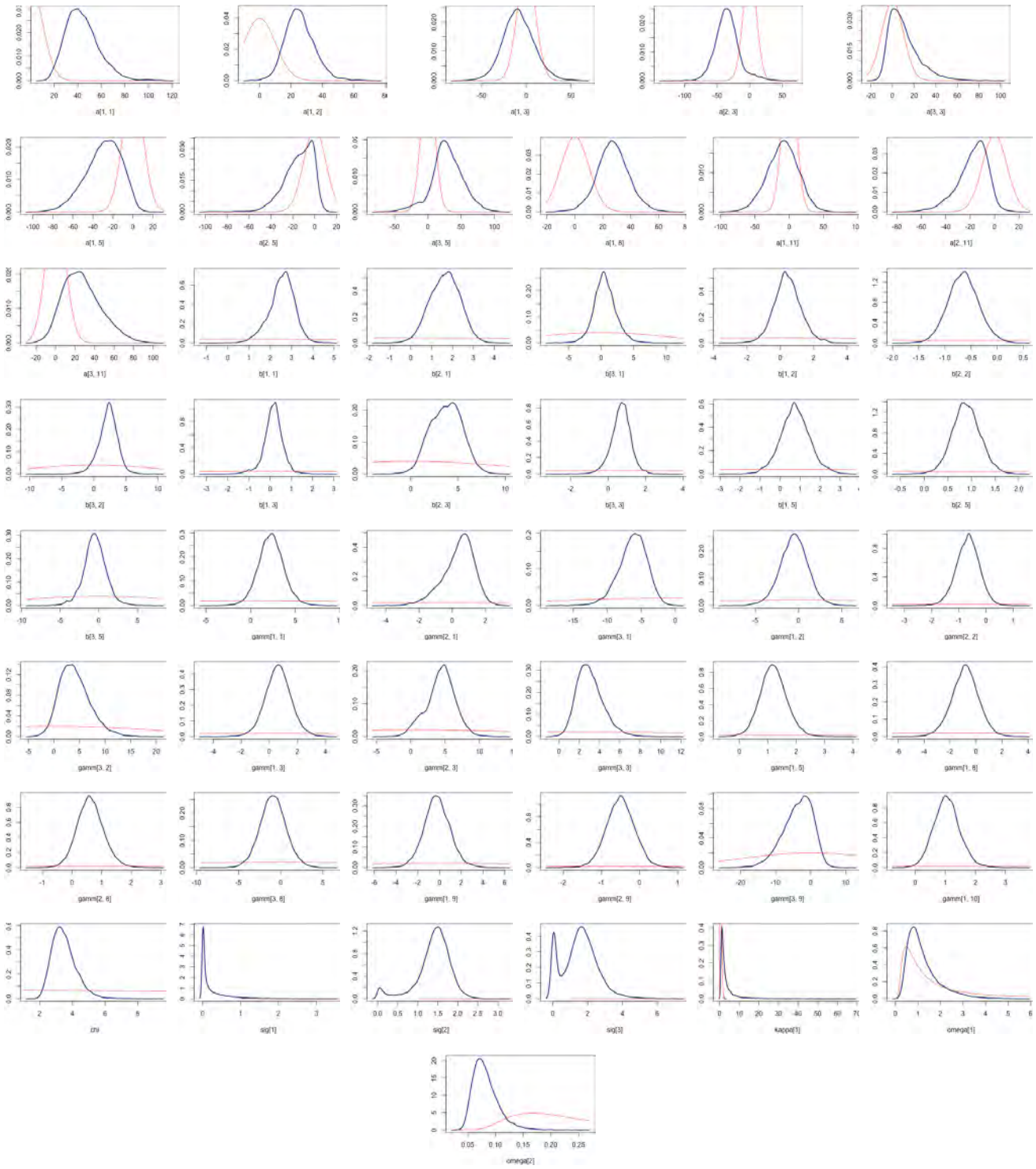


Figure 10.2.7: Reef-level modeling results: posterior and prior distributions, ordered according to the list of prior distributions given in Section 10.2.4.1. Posteriors are shown in blue, and priors, in red. Bimodality of posteriors does not suggest MCMC non-convergence, as evidenced by the convergence metric $\hat{R}$ (Gelman and Rubin 1992) being under 1.001 for the parameters in question.

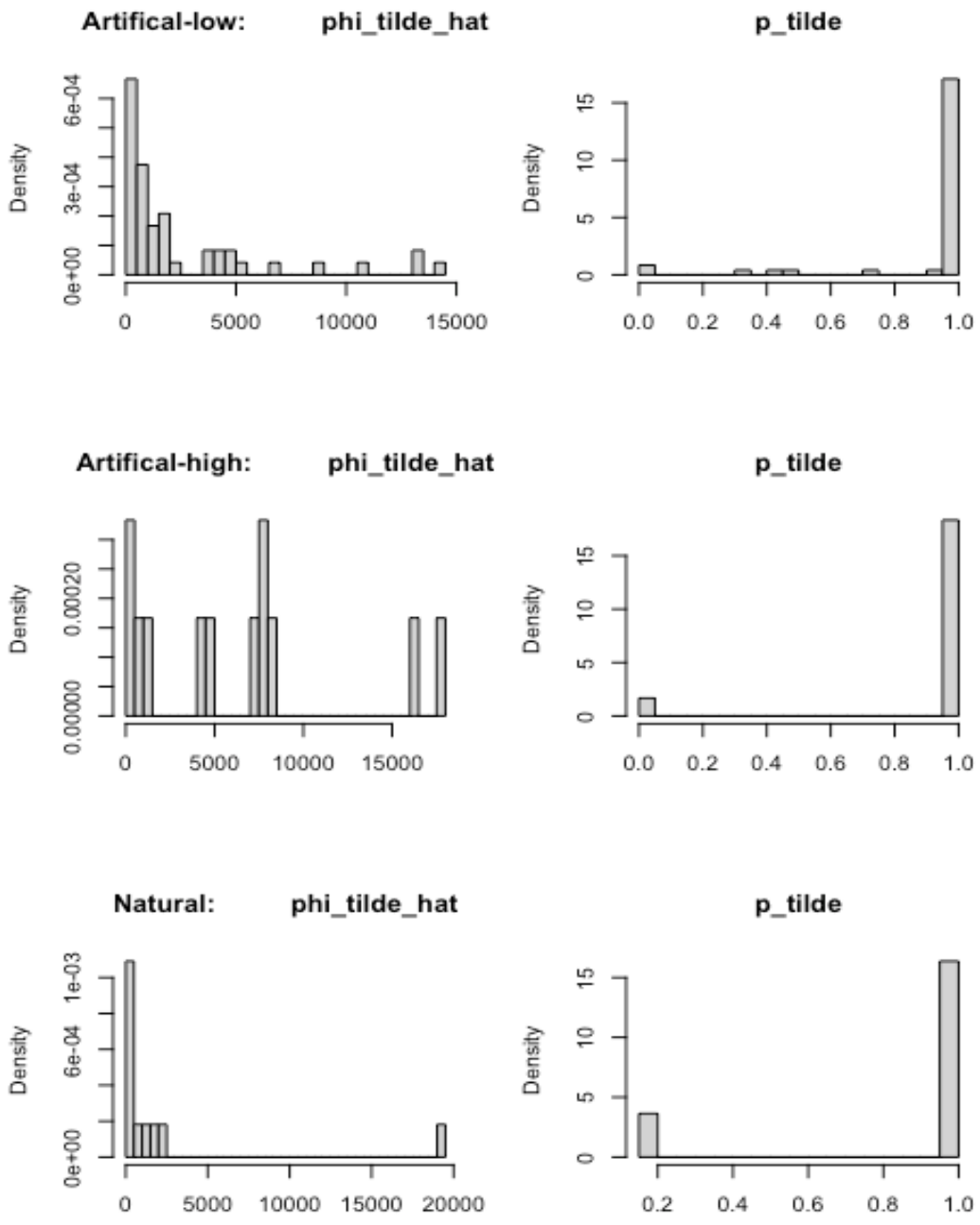


Figure 10.2.8: Reef-level modeling results: histograms of cell-level estimates for density and presence probability. Left column: $\hat{\tilde{\varphi}}_{ij}$ = posterior median of $\tilde{\varphi}_{ij}$ (inferred density); see maps for $\hat{\tilde{\varphi}}_{ij}$ in Figure 10.2.5. Right column: $\tilde{p}_{ij}$ = posterior probability for $\{\tilde{\varphi}_{ij} > 0\}$ (inferred GAJ presence). Top row: Low-relief artificial structures (48 sampled cells). Middle row: High-relief artificial structures (12 sampled cells). Bottom row: Natural structures (11 sampled cells).

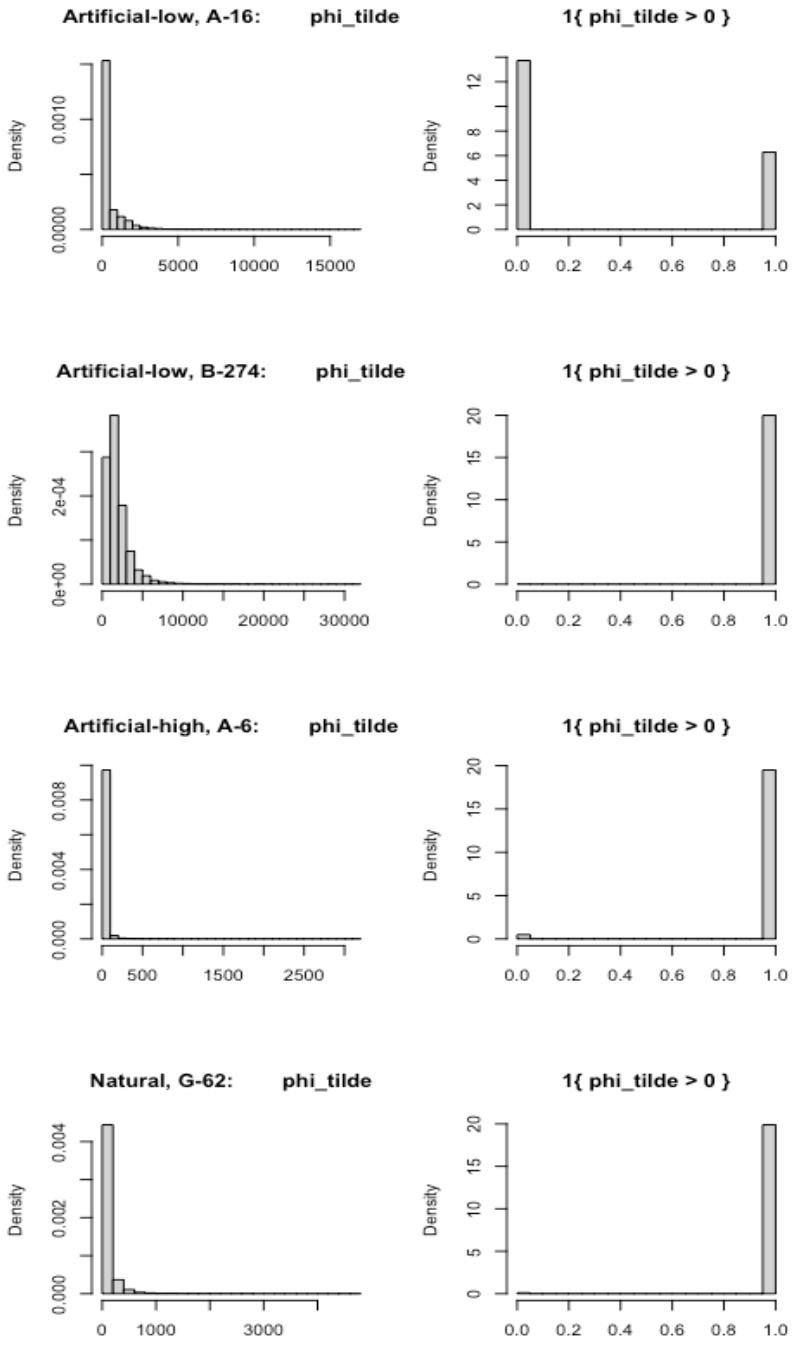


Figure 10.2.9: Reef-level modeling results: posterior distribution of cell-level density and presence/absence for selected habitat-cell combinations. Left column: inference for density ($\tilde{\varphi}_{ij}$). Right column: inference for GAJ presence ($\mathbb{1}\{\tilde{\varphi}_{ij} > 0\}$) — a value of '0' corresponds to inferred absence, and a value of '1', inferred presence. Each row corresponds to the same combination of habitat stratum and grid cell ID as indicated in the histogram title on the left.

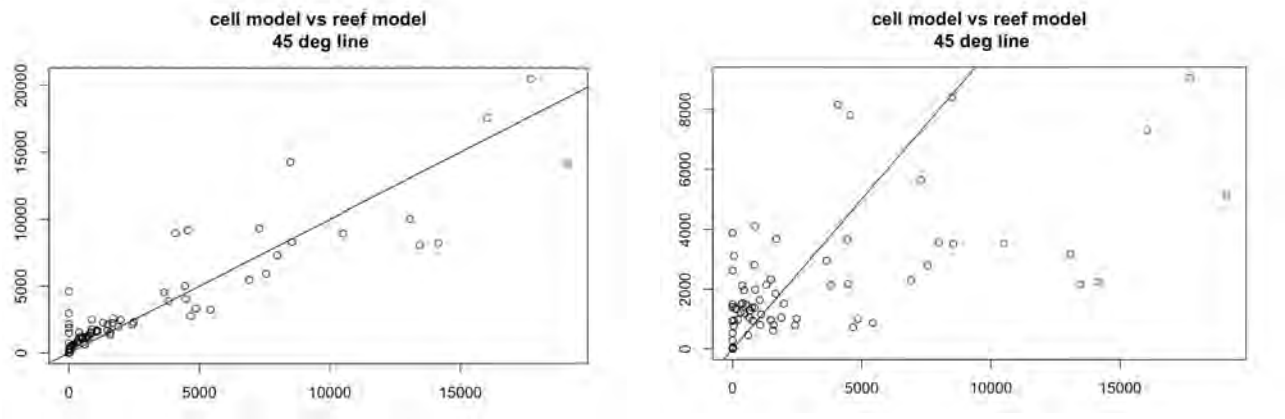


Figure 10.2.10: Cell-level density estimates for sampled grid cells: Phase 4 vs. Phase 3. Vertical axis is the posterior median of φ from Phase 4 inference. Horizontal axis is $\hat{\bar{\varphi}}$ from Phase 3 inference. Left panel: Results from the final Phase 4 model (Section 10.2.5.2.1). Right panel: Results from the unsatisfactory model (described in Supplement Section 10.2.9.4).

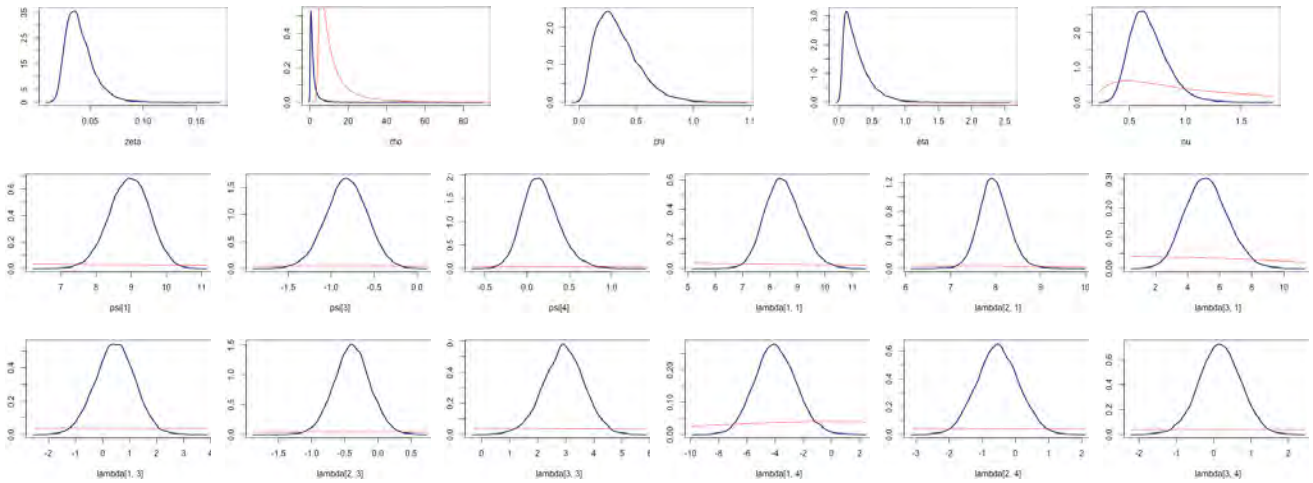


Figure 10.2.11: Phase 4 cell-level modeling results for sampled cells: posterior and prior distributions, ordered according to the list of prior distributions given in Section 10.2.5.2. Posteriors are shown in blue, and priors, in red.

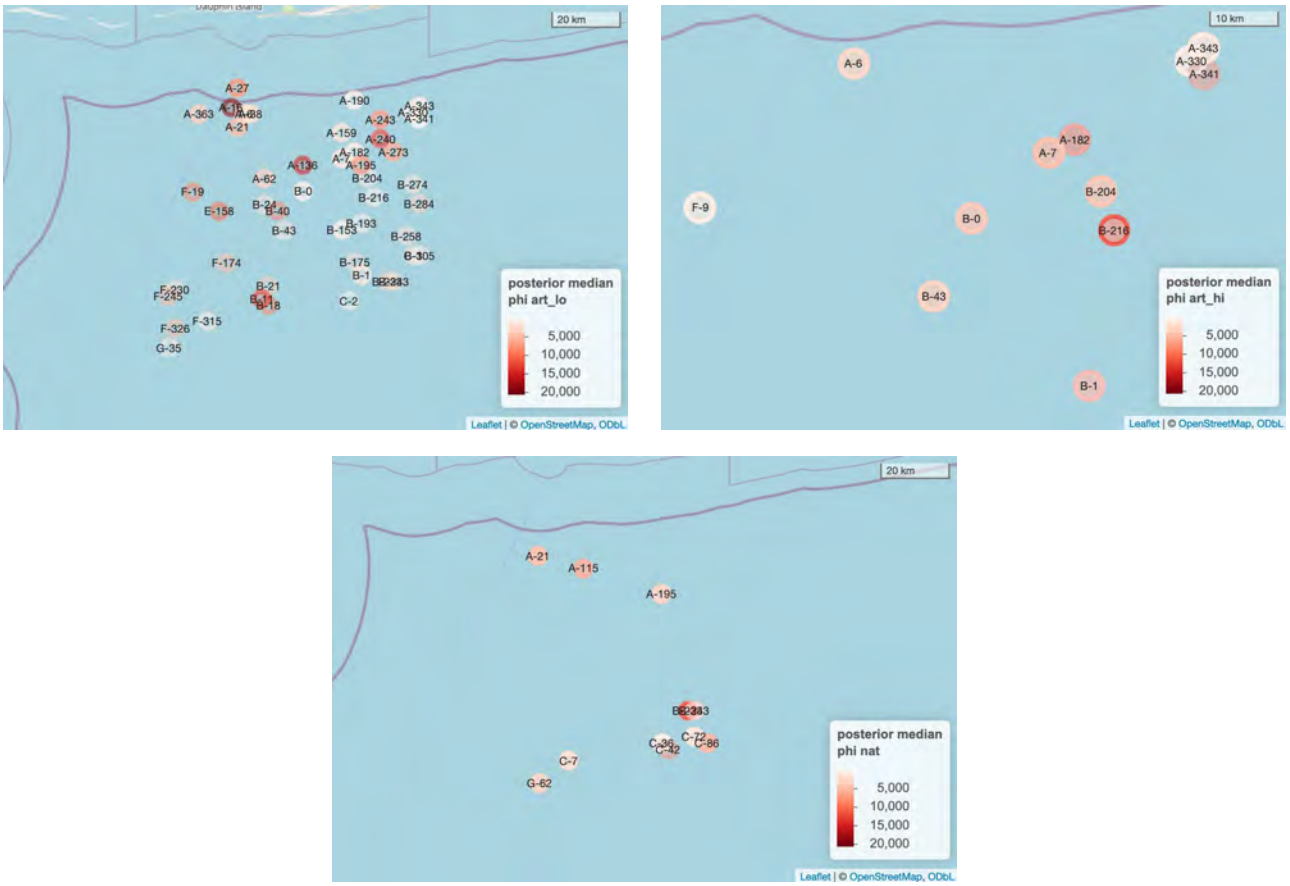


Figure 10.2.12: Phase 4 cell-level modeling results for sampled cells: estimated cell-level GAJ densities (posterior medians) by habitat type. Top-left panel: Low-relief artificial structures. Top-right panel: High-relief artificial structures. Bottom panel: Natural structures. See Supplement 10.2.A for interactive versions of these maps.

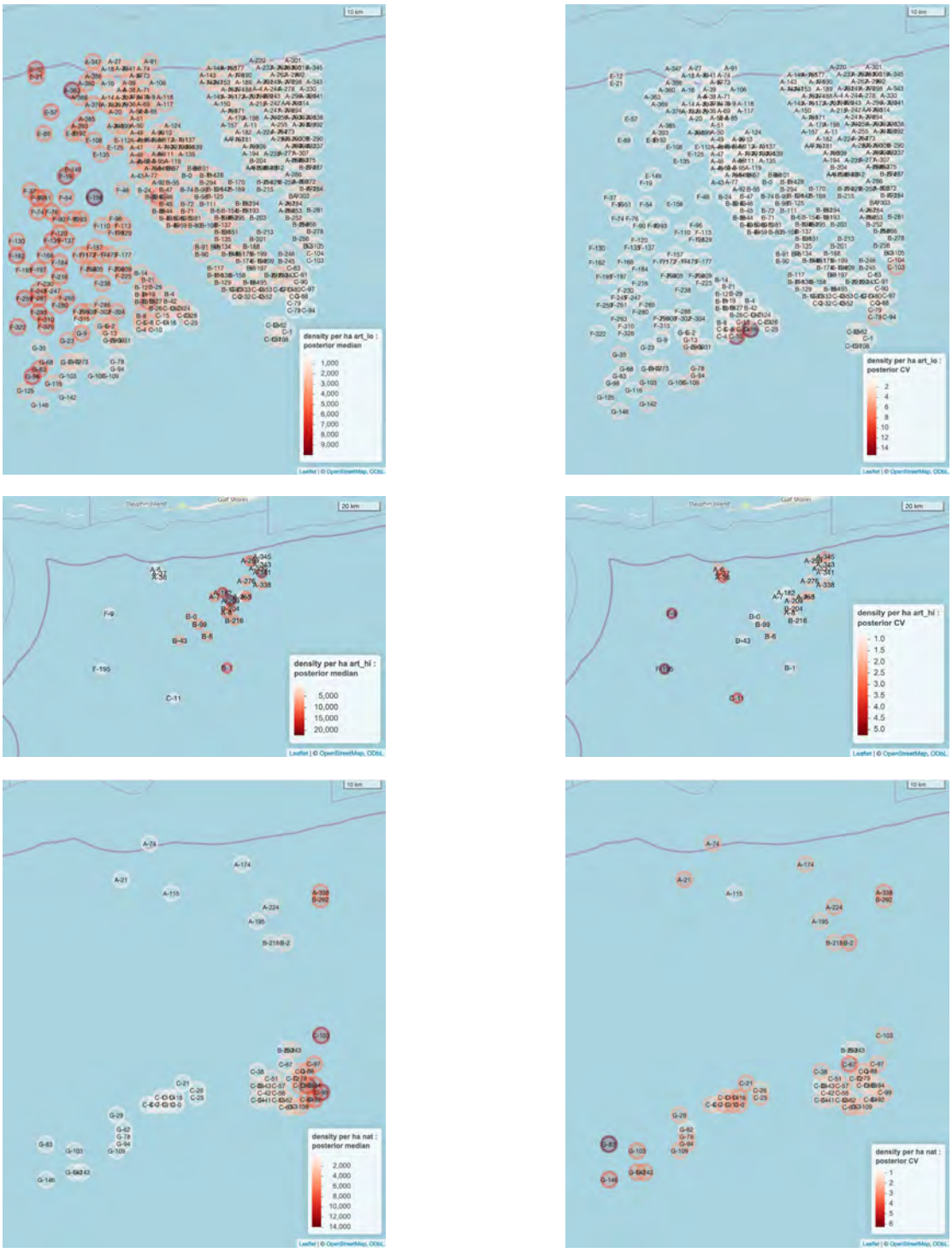


Figure 10.2.13: Phase 5 cell-level Model A results for AL universe: estimated cell-level GAJ densities and CVs, by habitat type. Top row: Low-relief artificial structures. Middle row: High-relief artificial structures. Bottom row: Natural structures. Left column: density estimates (posterior medians). Right column: coefficients of variation (CVs). See Supplement 10.2.A for interactive versions of these maps.

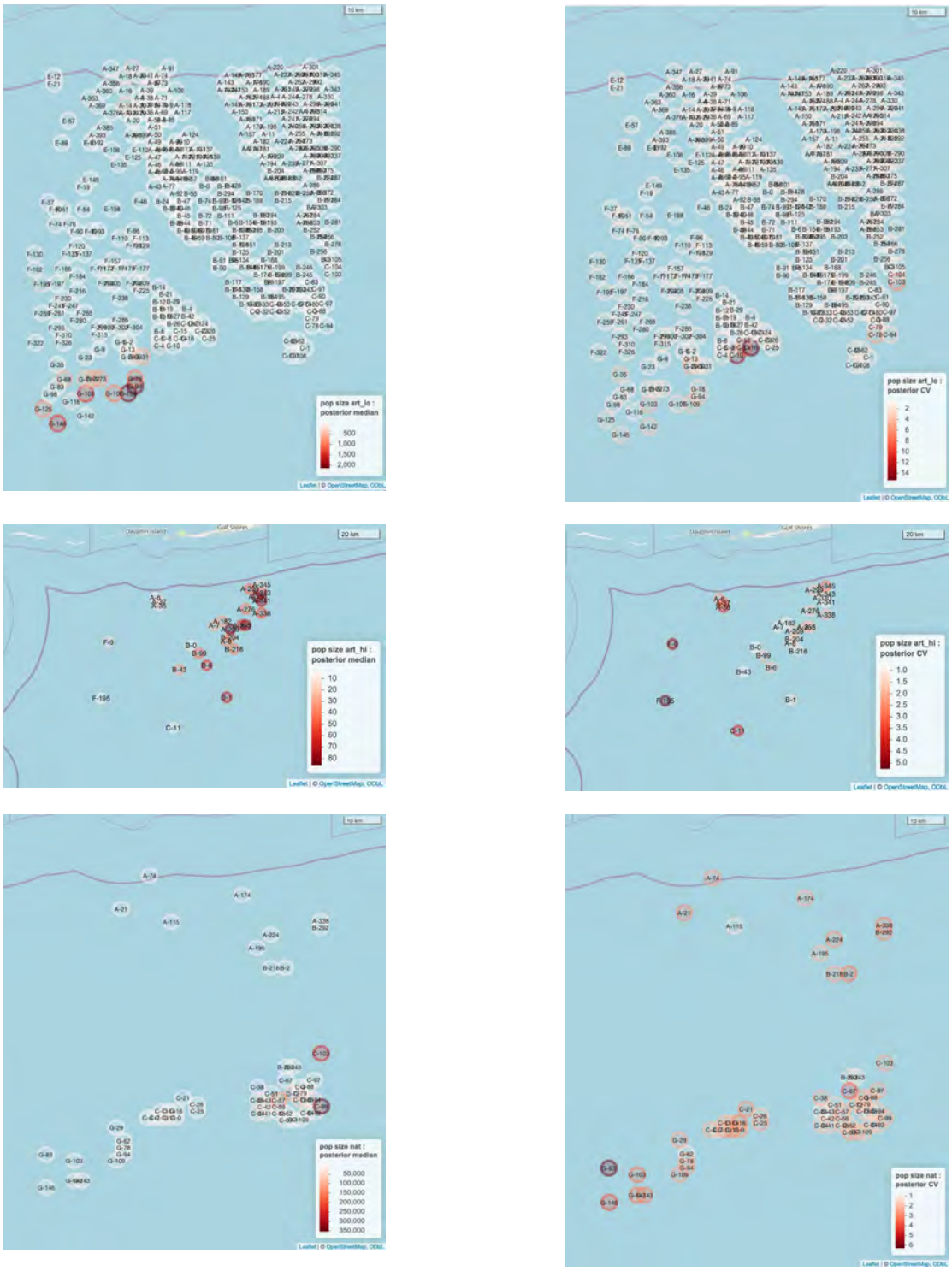


Figure 10.2.14: Phase 5 cell-level Model A results for AL universe: estimated cell-level GAJ population sizes and CVs, by habitat type. Top row: Low-relief artificial structures. Middle row: High-relief artificial structures. Bottom row: Natural structures. Left column: population size estimates (posterior medians). Right column: coefficients of variation (CVs). See Supplement 10.2.A for interactive versions of these maps.

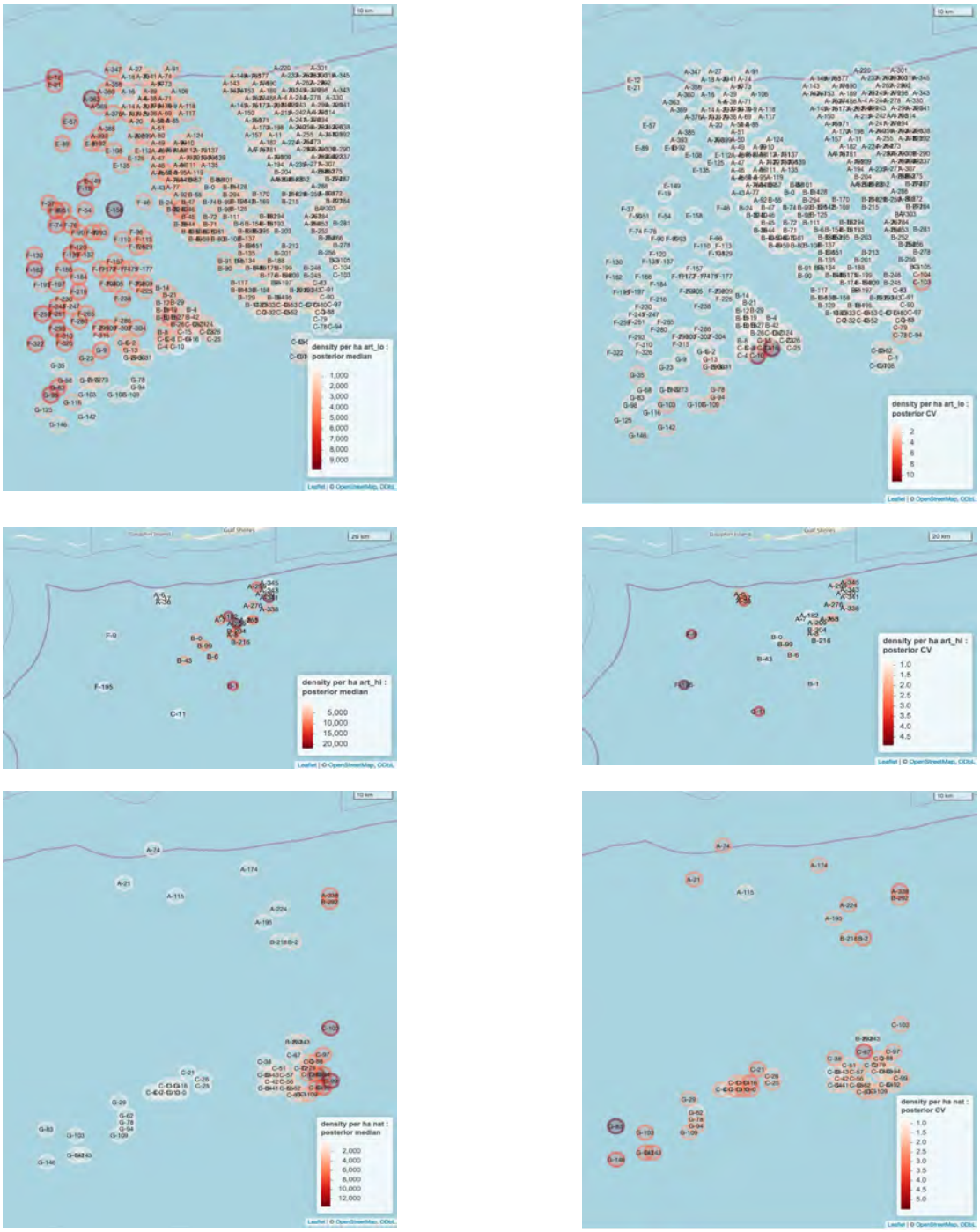


Figure 10.2.15: Phase 5 cell-level Model B results for AL universe: estimated cell-level GAJ densities and CVs, by habitat type. Top row: Low-relief artificial structures. Middle row: High-relief artificial structures. Bottom row: Natural structures. Left column: density estimates (posterior medians). Right column: coefficients of variation (CVs). See Supplement 10.2.A for interactive versions of these maps.

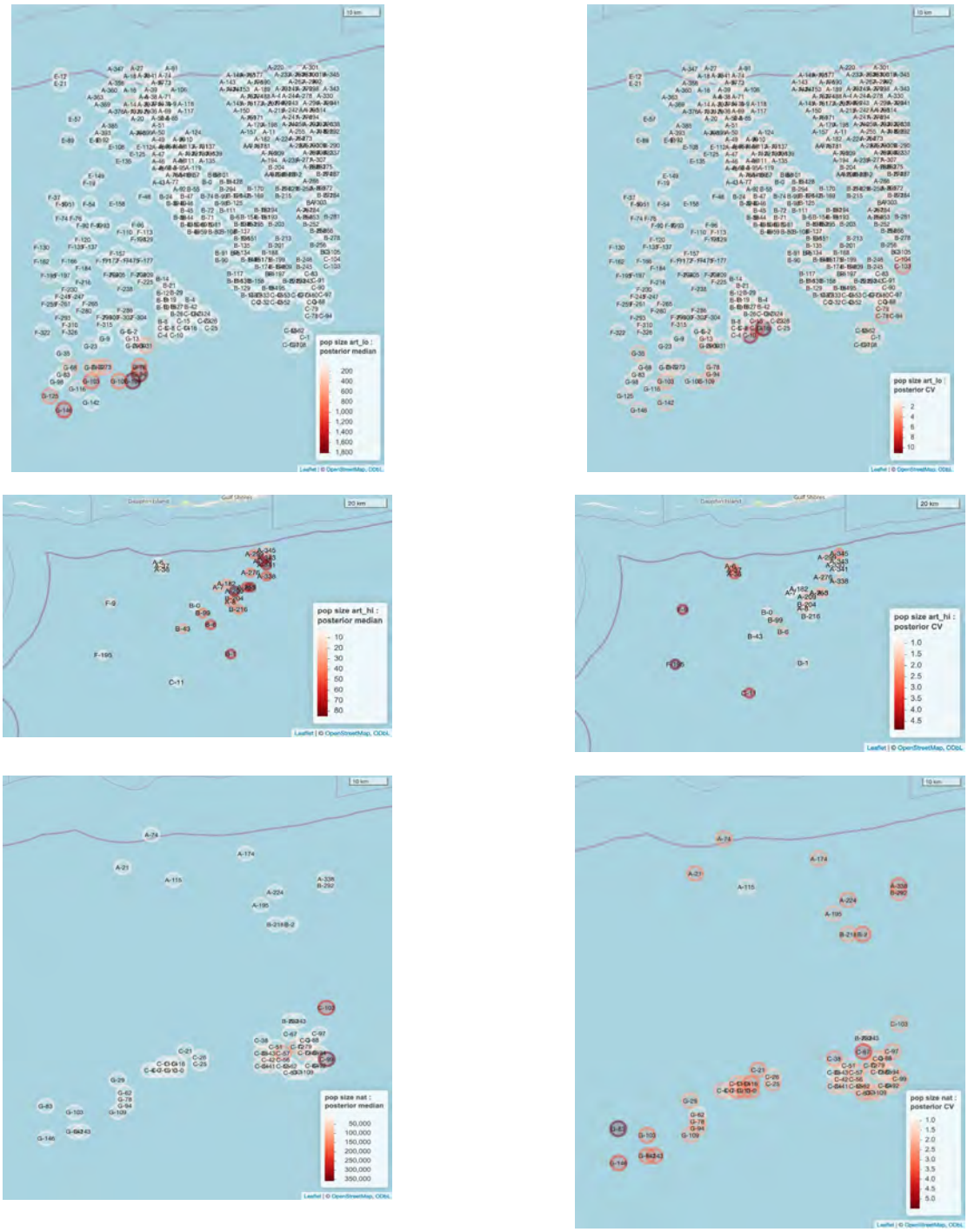


Figure 10.2.16: Phase 5 cell-level Model B results for AL universe: estimated cell-level GAJ population sizes and CVs, by habitat type. Top row: Low-relief artificial structures. Middle row: High-relief artificial structures. Bottom row: Natural structures. Left column: population size estimates (posterior medians). Right column: coefficients of variation (CVs). See Supplement 10.2.A for interactive versions of these maps.

REFERENCES FOR CHAPTER 11.2

de Valpine, P., C. Paciorek, D. Turek, et al. 2024. *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling*. R package version 1.3.0. https://cran.r-project.org/package=nimble.

Gelman, Andrew, and Donald B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 7 (November). https://doi.org/10.1214/ss/1177011136.

Gneiting, Tilmann, Adrian E. Raftery, Anton H. Westveld, and Tom Goldman. 2005. "Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation." *Monthly Weather Review* 133 (May): 1098–118. https://doi.org/10.1175/MWR2904.1.

Ponisio, Lauren C., Perry de Valpine, Nicholas Michaud, and Daniel Turek. 2020. "One size does not fit all: Customizing MCMC methods for hierarchical models using NIMBLE." *Ecology and Evolution* 10 (5): 2385–416. https://doi.org/10.1002/ece3.6053.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "Mice : Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45. https://doi.org/10.18637/jss.v045.i03.

11. Regional Estimates

11.1. Western Gulf of America

Greater amberjack (*Seriola dumerili*) abundance, distribution, and size-structure on natural and artificial habitats of the western Gulf of America

Mischa K. Schultz, Lynne Stokes, Ellen Taylor, Kevin Boswell, Creed C. Branham, Maelle M. Cornic, Brondum M. Krebs, Daniel Lippi, Sean P. Powers, Jay R. Rooker, Mathew Streich, Tara Topping, R.J. David Wells, Allison White, Stephen R. Midway, and Michael A. Dance

INTRODUCTION

The greater amberjack (*Seriola dumerili*) is a large, predatory carangid that inhabits shelf waters ranging from Nova Scotia to Brazil in the Atlantic Ocean, including the Gulf of America (hereafter, Gulf) and Caribbean Sea, and is largely found in association with hard-structured habitats (Bachelor et al. 2022). In U.S. waters, the species has long been managed as two distinct stocks, separated at the Florida Keys due to jurisdictional differences and minimal genetic exchange (SEDAR 2020; Hargrove et al. 2018). Following population declines of red drum (*Sciaenops ocellatus*) in the 1980s and 1990s, greater amberjack emerged as a commercially important alternative. The resulting increase in harvest pressure contributed to stock declines in both the Gulf and Atlantic. Although historically important in the commercial industry, recreational landings dominate overall removals in both regions, underscoring the species' economic importance, though regulations in the Gulf have constrained recreational opportunities. Management efforts have supported stock recovery in the Atlantic, with increases in abundance and landings over the past two decades (SEDAR 59), but the Gulf stock remains overfished and subject to overfishing (SEDAR 70). Indices of abundance for greater amberjack rely primarily on fishery-dependent data (SEDAR, 2020) and closures, while beneficial for the stock, will result in a loss of fishery-dependent data necessary for assessing stock productivity. In response, there is a critical need for a comprehensive assessment of greater amberjack abundance and distribution in both regions (GARP 2020).

Understanding the biology and habitat use of greater amberjack is critical to interpreting stock trends and informing management. Greater amberjack display ontogenetic shifts in habitat use, occupying a range of natural and artificial structures throughout their life cycle. In the northern Gulf, greater amberjack encounter a network of differing habitats that support their life cycle, as well as diverse fish assemblages (Ajemian et al. 2015; Garner et al. 2019; Rooker et al. 1997; Streich et al. 2017). Among these habitats, multiple reef banks formed by salt domes create reef structures off the coasts of Texas and Louisiana. In addition to these naturally occurring habitats, human activities—particularly oil and gas extraction—have led to the installation of thousands of artificial structures across the region (Kaiser et al., 2020). Petroleum platforms provide extensive vertical relief and surface area, forming habitats that attract dense reef fish communities, including high concentrations of greater amberjack (Gallaway et al. 2021; Cowan & Rose 2016; Reynolds et al. 2018). However, characterizing the abundance of greater amberjack across these habitats is logistically challenging due to the vast spatial scale, deep reef locations, and turbid conditions influenced by Mississippi River discharge (Gardner et al., 2022). These factors hinder the development of habitat-specific abundance indices, particularly for species like greater amberjack that display broad habitat associations.

Quantifying greater amberjack abundance is further complicated by reliance on traditional extractive sampling methods such as trawling, fish traps, and longlines (Farrell et al. 2014; Bacheler & Smart 2016; Karnauskas et al.

2017; French et al. 2021). Most abundance and distributional data come from fishery-dependent sources such as catch-per-unit effort (Karnauskas et al. 2017; Mammel et al. 2022; Wang et al. 2023) which are vulnerable to bias from changing regulations and closures. While effective, these methods can impact fish populations and may be impractical in complex habitats or marine reserves (McGeady et al. 2023). As a result, non-extractive approaches, such as underwater visual surveys (stationary baited remote underwater video and roving underwater video surveys), have become increasingly valuable for monitoring reef-associated species (Bolser et al. 2020; Jabado et al. 2021; Tothill et al. 2024). Advancements in underwater video technology, including high-resolution imaging and automated fish identification, have improved the efficiency and accuracy of visual surveys. However, underwater video surveys are subject to inherent biases, including variable detection rates due to habitat complexity and environmental conditions (Rooper et al. 2010), behavioral responses to the camera (Somerton et al. 2017), and limited spatial coverage or replication that may hinder accurate estimates of size and species composition.

Hydroacoustic sampling combats some of the issues encountered when using video surveys by enabling sampling the entire water column, with comparatively low effort (Egerton et al. 2018). The quality of hydroacoustic samples does not depend on visibility or detection probability in different habitats. However, species-specific identities, size, and orientation in the water column are not fully captured by hydroacoustic equipment, thus additional information is needed to accurately interpret acoustic signals (Fernandes et al. 2017; Scoulding et al. 2023). Stereo video systems, which provide high-resolution imagery and allow for precise length measurements and species identification, help overcome these limitations. As a result, integrating hydroacoustic data with underwater stereo video has proven to be an effective approach for estimating fish densities with improved accuracy (Boldt et al., 2018; Rasmuson et al., 2022). Compact stereo-drop cameras have been suggested to increase the number of video samples necessary for characterizing species proportions across hydroacoustic surveys (Rasmuson et al. 2022; Scoulding et al. 2023). Combining data from multiple gears, such as hydroacoustics and stereo drop cameras, presents the opportunity for improved inference of greater amberjack abundance across US waters, particularly in complex regions such as the US Gulf.

To address persistent data gaps in greater amberjack distribution and habitat use in the northern Gulf, we aimed to answer the following objectives: i) obtain habitat and depth zone specific estimates of greater amberjack density in the western GOM, ii) assess relative depth use of greater amberjack at different habitats to better understand vertical habitat partitioning, iii) characterize size distribution of greater amberjack on different habitats and depth zones and iv) use previously published von Bertalanffy growth parameters to estimate age structure and construct a length-based catch curve from video-derived data. This non-extractive approach offers a valuable alternative to traditional gear-based methods, which often suffer from size-selectivity biases (Langlois et al. 2015). Collectively, these objectives aim to improve our understanding of greater amberjack spatial ecology and demographic structure, contributing important data to support more effective management of this ecologically and economically valuable species in the Gulf.

METHODS

Study Design

A stratified-random sampling design was developed to estimate greater amberjack abundance in shelf waters (20 – 150 m) of the Western Gulf of America (Louisiana and Texas regions west of the Mississippi River). Surveys consisted of hydroacoustic and video surveys conducted at randomly selected sites. This approach focused on distributing sampling effort geographically and across a range of depths and habitats.

The shelf was first divided into three depth zones representative of shallow (20-40 m), mid-(40-70 m), and outer-shelf (70-150 m) regions. Known habitats that were used in the sampling effort included artificial reefs (AR), standing gas platform (P), and natural reef (NR). The rest of the shelf that did not fall into any of these categories was classified as uncharacterized bottom (UCB). Thus, there was a maximum of 15 unique strata (5 habitat types x 3 depth zones) in the survey design for each region (Louisiana and Texas). The artificial reef habitat stratum included structures deployed and maintained by the state artificial reef programs (LDWF, TPWD) and available through online databases (e.g., https://www.wlf.louisiana.gov/page/artificial-reefs). Collectively, this stratum was considered the "known" artificial reefs. Locations of additional artificial structures were obtained using the NOAA wrecks and obstructions layer (https://www.nauticalcharts.noaa.gov/data/wrecks-and-obstructions.html). However, because the status of these structures is not updated regularly, there was some concern that inclusion of these structures in the artificial reef stratum could bias estimates, as it was unknown what percentage of these sites still exist. As a result, these structures were separated from the artificial reef layer, and treated as a separate stratum, which is hereafter referred to as wrecks (artificial reefs of unknown status). Locations of standing oil and gas platforms were extracted from the BOEM platform database (https://www.data.boem.gov) which is updated regularly. Decommissioned platforms were removed from the stratum, so that only existing platforms remained. Geographic location and extent of natural reef habitat was obtained from Gardner et al. (2022) which represented the most recent and best available dataset for the western Gulf of America. Finally, the remaining seafloor that did not contain mapped (known) benthic structure was collectively grouped as uncharacterized bottom.

Sample locations were selected randomly from grids (5 x 5 nautical miles) overlaid across the shelf of each state (Figure 11.1.1). Each grid cell containing potential habitat (by structure or area) for a specific habitat stratum was included in the sampling universe of that habitat stratum (Table 11.1.1). A total of 165 cells were randomly chosen in Louisiana: artificial reefs (n = 30), platforms (n = 45), wrecks (n = 30), uncharacterized bottom (n = 45), and natural reef (n = 15) (Table 11.1.2). In Texas, a total of 161 cells were randomly chosen: artificial reefs (n = 30), platforms (n = 39), wrecks (n = 30), uncharacterized bottom (n = 52), and natural reef (n = 10) (Table 11.1.2). The number of samples per stratum was decided based on available sampling time, likelihood of and expected variance in amberjack presence, dividing effort across each of the habitat strata, and preliminary discussion with regional sampling groups, the statistics team, and project management team. Sampling effort (number of sampling sites) for each habitat stratum was stratified across the three fixed depth zones in proportion to the available habitat in each depth zone (Table 11.1.3). For each habitat-depth stratum, cells were then randomly drawn (based on the number of samples outlined in Table 11.1.2) independently without replacement. Thus, while a cell could not be selected twice within a habitat-depth stratum, a cell could be selected a second time in a different stratum. After cell selection, a single sampling location within a cell was chosen. For cells containing multiple habitat features (e.g., multiple oil and gas platforms, etc.), a single specific sampling location or structure was selected randomly. Because of time and logistical constraints (i.e., weather), sampling crews had the ability to adjust randomly selected cells or locations within cells to adjacent cells or features to facilitate survey completion.

Camera and hydroacoustic surveys

Survey protocol included use of both hydroacoustics and fixed camera arrays to record fish communities and estimate the abundance of greater amberjack on the Louisiana and Texas shelves. Hydroacoustic surveys were performed using two Simrad EK80 transducers: ES 120-7C and ES 70-7C. For most surveys, transducers were coupled with a motion reference unit and mounted to a towfish which was pulled at a depth of 0.8 m approximately 10 m behind the vessel. On surveys conducted on the R/V Pelican cruise (LSU, TAMUG, and

TAMU-CC, August of 2023, Table 11.1.4), transducers were installed on a vertical pole mount fixed to port side of the vessel, and a global positioning system was used to determine the geographical coordinates. Transducers were calibrated during surveys using 32 mm and 22 mm diameter tungsten spheres on all vessels. Video surveys were conducted using SeaGIS underwater stereo-video systems, featuring two front-facing GoPro Hero 10 cameras and a rear-facing GoPro Hero 10 mounted to a stainless-steel frame. The cameras were installed in waterproof housing attached to a SeaGIS (http://www.seagis.com.au) steel base bar with 700 mm of camera separation and 8° inward convergence. Angling the cameras inward enabled the capture of 3D images for accurate fish length estimation. Cameras were set to record at 1080p resolution with a linear field of view (FOV) and 30 frames per sec. Trawl floats were attached to each corner of the camera frame for stability and to control the descent speed (0.6 m/s) of the camera array during vertical transects. A temperature-depth logger (TDL; Sensus Ultra, Reef Net Inc., Ontario, Canada) was affixed to the frame to record depth (±0.3 m) and temperature (±0.8°C) at least every 10 seconds, providing a time series of environmental data for each survey. Stereo-video systems were calibrated prior to and after each field sampling trip using a 3D calibration cube and the CAL software v3.91 (SeaGIS Pty LTd., Australia, https://www.seagis.com.au/bundle.html). Calibrations were checked regularly with scale bars of known lengths to ensure measuring accuracy.

Fishery independent underwater video surveys were conducted off the Louisiana and Texas coast in the Gulf of America from May to September 2023. Additional surveys took place in July and August of 2024. Video surveys were performed between the daylight hours of 0730 – 1800h to ensure adequate daylight conditions. Vertical transects were designed to obtain density estimates while minimizing attraction to gear, double counting, and the risk of gear entanglement in petroleum platforms or toppled platforms. This methodology achieved these goals while also allowing maximum visibility of the structure. Because of the unique nature of the different habitat strata, hydroacoustic transects and camera surveys were designed to efficiently survey each habitat (Figure 11.1.2). Hydroacoustic surveys at oil and gas platforms consisted of 4 perpendicular transects on all four sides of the structure, extending 100m beyond the platform. Hydroacoustic surveys at artificial reef and wreck sites consisted of 7 parallel transects of 200 m from the center of the reef, extending out to 70 m from the center point. Vertical transects with the camera array were conducted adjacent to the structure on the down current side. Hydroacoustic surveys at natural reefs and uncharacterized bottom focused on a larger footprint (0.7 km x 1 km) of the seafloor to capture habitat variability. Vertical transects with the camera arrays were conducted at five locations within the footprint (along the transect lines to align with hydroacoustic surveys). This design was implemented with the intent for camera surveys to obtain species composition information and hydroacoustic surveys to obtain a total count of individuals.

Video processing

Fish identification, enumeration, and measurement from video footage were conducted using EventMeasure software (SeaGIS Pty Ltd., Bacchus Marsh, Victoria, Australia). This process included a composite add-on program, VidComp, allowing simultaneous viewing of front and rear cameras for a comprehensive 360° perspective. Transect times were binned into 10 seconds intervals and matched with the associated depth and temperature for each observation. Therefore, each fish point count or length measurement was linked to its position (depth in meters) in the water column. Fish were identified to the lowest taxonomic level and counted during the descent of each transect. Fish were enumerated in two ways. First, we conducted a transect count, defined as the total number of individual greater amberjack observed along the descending transect (surface to seafloor) at each site. The continuous descent speed of the vertical transect minimized risk of double counting. MaxN, defined as the maximum number of an individual species observed at any one frame of the video, was

also calculated as a conservative estimate of the minimum number of fish present at each site and was used to calculate the species ratio to inform observations from acoustic samples.

During video processing, we recorded complementary data for each transect, including current direction, water clarity, transect quality, bottom type, and seafloor habitat characteristics. Current direction was classified relative to the front-facing cameras as: with the camera's field of view (away), perpendicular to the camera's field of view (sideways), or against the cameras field of view (toward). Each transect was assigned a quality code to indicate potential issues affecting visibility or analysis. Transects were categorized as dark video, obstructed view, out of focus, turbid water, suspended camera array if the camera did not reach the bottom, or no issues. Water clarity adapted from Bachelor et al., (2022) was categorized as poor visibility if the substrate was not visible, fair if the substrate was visible, but structure or horizon are not visible, and good if both the substrate and horizon are clearly visible. For all transects where the substrate was visible, bottom type was classified into three categories: sand/silt, gravel or rock. For transects occurring over natural reef habitats, seafloor characteristics were determined as the determined as the percentage of visible substrate that was hardbottom on each camera and substrate relief estimated in three categories: low (<0.3m), moderate (0.3 – 1.0m) or high (>1.0m).

Using the stereo-configuration and calibration procedure, fish lengths and distances can be measured with high accuracy (Reynolds et al., 2018; Tothill et al., 2024). Fish within 10m of the stereo-video were measured to decrease measurement error with increasing distance. Fork length (FL) measurements were made when both the caudle peduncle and snout were visible in both cameras and when the fish was oriented at an angle of less than 45° perpendicular to the camera. Measurements with a precision (mm) to length ratio greater than 15% were removed from the analysis (SeaGIS Pty Ltd., Bacchus Marsh, Victoria, Australia).

Camera-Based Estimates

For each sampling site, greater amberjack abundance was estimated from stereo camera video using habitat-specific calculations to account for differences in the spatial extent and structure of each habitat type. The total transect count and MaxN count of greater amberjack observed along the vertical transect at a site served as the bases for all estimates and were converted to either a density or number of fish per structure, depending on habitat type. Camera sampling area was calculated as the area of a circular sector, where the radius was defined by the maximum detection distance of the stereo camera system (13.9m) and the field of view (FOV) was expressed as a proportion of a full circle ($\pi r^2 \times FOV/360$), yielding an effective sampling area in m^2.

Site-level estimates differed by habitat type to reflect the underlying spatial structure of each reef category. For artificial reefs and platforms, total amberjack abundance per reef (fish km^{-2}) was estimated by standardizing the observed total count by the camera sampling area and scaling to the total sampled reef structure area:

$$\widehat{N}_{AR,P} = \left(\frac{n_{obs}}{A_{cam}}\right) \times A_{structure}$$

where n_{obs} is the total count observed, A_{cam} is the camera area sampled (km^2), and $A_{structure}$ is the total area of sampled structure (km^2). For natural reef and uncharacterized bottom sites, where no discrete structure boundary could be defined, abundance was expressed as a density standardized to the camera sampling unit (fish km^{-2}):

$$\widehat{N}_{NR,UCB} = \frac{n_{obs}}{A_{cam}}$$

For wrecks, where there was no information on area of structure, or structure components, abundance was expressed as:

$$\widehat{N}_W = n_{obs} \times 1$$

This approach preserves the ecological meaning of each estimate relative to the habitat being sampled while allowing site-level values to be aggregated or modeled across habitat types.

Species Ratio Apportionment for Hydroacoustic-Based Estimates

Relative greater amberjack apportionment was calculated using the acoustic and video data together to inform observations from acoustic samples. To apportion the acoustic counts from the echo integration to fish observed during video surveys, MaxN values for fish that were water column associated (> 0.5 m off the seafloor) and > 25 cm FL were summed at each site to calculate a total MaxN for that site. The fraction of greater amberjack to all other large-bodied fish that would have been identified as potential *Seriola* was estimated using an empirical Bayes estimation (see Chapter 10.1). Using a Bayes estimator for each depth-habitat combination allowed us to apply a species ratio to sites that had poor visibility or bad quality video surveys that would otherwise not have a representative ratio of the species within that depth zone. Species ratios were then applied to the total acoustic counts from the echo integration thus the resulting number represented the estimated greater amberjack seen at each site. Using the acoustic and video together allowed us to roughly correlate and apportion total counts of greater amberjack to use in subsequent analyses.

Depth Utilization Analysis

Greater amberjack depth utilization was investigated using the depth (m) of observation (OD) recorded by the TDL attached to the stereo-video system. To assess relative depth (RD) use at each habitat type, we calculated the ratio of OD to site (SD), where:

$$\mathrm{RD} = \frac{\mathrm{OD}}{\mathrm{SD}}$$

Relative depth was chosen as a metric because depth utilization is assumed to depend on the available water column at each site. The RD metric ranged from 0 to 1 where 0 represents the top of the water column and 1 represents the bottom. Because values were between zero and one, a Bayesian beta regression model was used to determine whether habitat influenced greater amberjack depth utilization:

$$\mathrm{Y_i} \sim \mathrm{Beta}(\mu_\mathrm{i}, \phi)$$

$$\mathrm{logit}(\mu_\mathrm{i}) = \alpha + \beta_{\mathrm{habitat[i]}}$$

where $\mathrm{Y_i}$ denotes the relative depth, is the dispersion parameter, and represent the coefficient of the fixed effect of habitat. Uninformative priors were placed on all parameters. We considered depth utilization to differ by habitat if the 95% credible intervals for the fixed effect did not overlap. Models were fit in package brms (Bürkner, 2017).

Bayesian Modeling of Amberjack Size Distribution by Depth Zone

Additionally, a Bayesian gaussian regression model was applied to assess whether greater amberjack size distribution varied across depth zones. Uninformative priors were used for all parameters except for a truncated normal prior on the intercept that reflects the size of greater amberjack at reef recruitment (250 – 350 mm; Dance et al., 2011; Manooch & Potts, 1997; Wells & Rooker, 2004). Size distribution was considered to vary by

depth zone if the 95% credible interval for the fixed effects did not overlap. Size distribution models were fit using the brms package (Bürkner, 2017) in R.

Length Assignment

We used the length-only dataset from the video transects to estimate the length structure of the abundance estimates. Individual fish in the video-derived length dataset were assigned to length categories. To align the length bins with the age-structure of greater amberjack, length bins were informed by the range of fork lengths associated with each age class in the aged specimen dataset in Chapter 8. Length bins were defined as 0-400, 400-600, 600-800, 800-1000, and 1000+ mm FL. Within each depth zone, the proportion of amberjack within each length bin was then calculated to characterize the length structure of the abundance estimates.

Catch Curve Analysis

The aged specimen dataset was used to construct the ALK. Aged amberjack specimens were assigned to 100-mm FL categories. The joint length-age distribution was tabulated as a contingency table of length category by age class. The ALK was then computed as the row-proportional form of this table — that is, for each length category, the proportion of individuals belonging to each observed age class was calculated. The resulting matrix (the ALK) expresses the conditional probability of a fish belonging to a given age class given its length category. For example, if 75% of aged fish in a given length bin were age 1, then any unaged fish in that bin has a 0.75 probability of being assigned age 1. Ages were then assigned to individual fish in the video-derived length sample using the alkIndivAge() function in the FSA package, which implements the Isermann and Knight (2005) algorithm. Rather than simply assigning mean ages by length bin, this method assigns discrete ages to each individual fish while preserving the expected proportions from the ALK. Once ages were assigned to the lengths, catch curves were computed only for ages 1+ assuming those individuals were fully recruited to the sampling gear. Then an instantaneous mortality rate (Z) was calculated using the natural logarithm of the number of greater amberjack in each age-class as a function of calendar age (Miller et al., 2022; Ogle, 2016). The instantaneous mortality rate was converted to an annual mortality rate using $A=1-e^{-Z}$ (Ogle, 2016).

Estimation of Absolute Abundance

Sample data for making both camera-based and acoustic-based estimates of abundance of *Seriola dumerili* were available for the Western Gulf. Estimation was carried out per state, and summed for a total estimate of abundance. The count values used for abundance for the camera-based estimates were observed estimates on sampled sites from vertical transects (described earlier). The count values to be used for abundance for the acoustic data were calculated from an echo integration procedure at the same sample sites as the camera data, but multiplied by a species-ratio factor that is derived from the camera data that attempts to estimate the proportion of large-bodied fish signals that are actually amberjack (See Chapter 10.1).

The sample design used to choose sample sites was a stratified multi-stage design. The strata were defined by region (LA, TX), habitat (Natural reef, artificial reef, platforms, wrecks, and unconsolidated bottom), and depth (20, 40, 70). The primary sampling units were grid cells, the union of which covered the entire Western Gulf. The second stage units were either randomly chosen areas within the grid cell for areal habitats (natural reef and unconsolidated bottom) or randomly chosen structures for enumerated habitats (artificial reefs, wrecks, and platforms). The areal coverage of the second stage units differed for camera and acoustic-based samples. The camera's areal unit was the area of habitat within the field of view of the camera, which was assessed to be

118.36 m^2 for artificial structures and 236.72 m^2 for natural reef and uncharacterized bottom, while the acoustic areal unit varied across sample units according to the area swept.

For the camera-based estimates, the design required analysis according to a three-stage design for the enumerated habitats. The third stage of sampling was necessary because it was typically not possible to observe the count of GAJ on an entire sampled structure due to the camera's limited field of view. Therefore, the fraction of the structure that was observed was taken as the third-stage selection probability. This fraction was calculated as the assessed areal field of view of the camera divided by the areal footprint of the sampled structure.

Estimates were produced using the stratified ratio estimator, as described in Chapter 10.1. For the areal habitats, the auxiliary variable used for the ratio estimator was its habitat area within the grid cell, while in the enumerated habitats, the auxiliary variable used was the count of structures. This was possible since the totals of structure count and areal coverage was known for the universe for each stratum. The sample weights used in the estimator were the reciprocal of the products of the two (for areal habitats) or three (for enumerated habitats) selection probabilities for the stages of the designs. These weights were adjusted for non-response using weighting class adjustment, where the weighting classes were the strata to which the missing observations belonged. The estimates and their standard errors were computed using the R package survey (10.32614/CRAN.package.survey), which appropriately handles variance estimation for complex sample designs. See Chapter 10.1 for the full estimation methodology.

RESULTS & DISCUSSION

A total of 326 sites were sampled across Louisiana (n = 168) and Texas (n = 158; Table 11.1.2); yielding usable data from 298 hydroacoustic surveys (LA = 152, TX = 146) and 734 camera surveys (412 from Louisiana and 322 from Texas). Greater amberjack (n =1,268; LA = 794, NTX = 474) were observed in 114 camera surveys across the western Gulf. The frequency of occurrence (FO) on survey sites in Louisiana was 35%, whereas the FO in Texas was 25%. Frequency of occurrence differed by habitat type, where in FO was highest on natural reefs (16 of 21 sites) in Louisiana and highest on platforms in Texas (22 of 39 sites; Figure 11.1.3). Of these observations, 242 greater amberjack were measured—112 from Louisiana and 130 from Texas—with fork lengths (FL) ranging from 314 to 1,593 mm (Figure 11.1.4; mean: 927.20 mm, s.e.: 21.80). Greater amberjack sizes (FL, mm) did not differ among Louisiana (mean: 913 mm, s.e.: 32.2), north Texas (mean: 981 mm, s.e.:27.4), and south Texas (mean: 866 mm, s.e.: 41.0). Greater amberjack were distributed across all habitat types in the western Gulf (Figure 11.1.5); however, densities varied by habitat.

Camera-based estimates

Site-level stereo camera estimates derived from MaxN (Table 11.1.5) were consistently lower than those derived from total counts (Table 11.1.6), which was expected because MaxN reflects the minimum number of fish present at a site. Still, patterns were generally consistent between the two approaches across subregions, habitats, and depth zones. Stereo-camera density estimates varied by depth zone, habitat, and state across the study area, but generally increased with depth and were highest on the mid to outer shelf, regardless of habitat or state (Table 11.1.4, Table 11.1.5).

Greater amberjack densities at natural reefs were highest in Louisiana, with both MaxN and total count estimates exceeding those at natural reefs in Texas. Natural reef habitat in both Louisiana and Texas is concentrated on the outer shelf, and because sample sizes were small (< 2 sites) in the 40 – 70 m depth zone,

the mid (40 – 70 m) and outer (70 – 150 m) depth zones for natural reef were combined during estimation (Tables 11.1.4, 11.1.5). Uncharacterized bottom site estimates were effectively zero in Texas across all depth zones, while estimates in Louisiana were elevated with high uncertainty in the 70 – 150 m depth zone. Greater amberjack density at artificial reefs was generally higher in Louisiana than in Texas, with the highest densities (MaxN >75 fish per structure) observed in 40 – 70 m (Louisiana) and 70 – 150 m (Texas and Louisiana) depth zones (Table 11.1.4). Greater amberjack densities at platforms were relatively low in the 20 – 40 m depth zone across the western Gulf, which could be partly influenced by reduced visibility at shallow water platforms. While amberjack densities at platforms were highest in the 70 – 150 m zone of Louisiana, overall mean densities were generally comparable between the two states (MaxN = 33 – 37 amberjack per platform; Total Count = 62 – 77 amberjack per platform) (Tables 11.1.4, 11.1.5). There were no greater amberjack observed on wrecks in Louisiana, and only a few sites in south Texas. It should be noted that sampling revealed a lack of structured habitat at most wreck sites in both states, indicating either the habitat is no longer present or the coordinates for these sites were inaccurate. Thus, the low numbers for wreck sites likely reflects a lack of habitat rather than low abundance on a particular type of reef.

Hydroacoustic-based estimates

Site-level acoustic estimates of greater amberjack also varied considerably across subregions, habitats, and depth zones, and followed patterns seen in the camera-based estimates (Table 11.1.6). Greater amberjack densities on natural reefs were considerably lower using hydroacoustic-based methods compared to camera-based methods. Hydroacoustic-based uncharacterized bottom estimates in Texas and Louisiana were again zero at the shallow and mid depth zones, with low densities observed in the deepest zone in both Texas (~6 fish per km^2) and Louisiana (~26 fish per km^2), indicating that amberjack occurrence over uncharacterized bottom was higher in Louisiana compared to Texas. Still, it should be noted that hydroacoustic estimates were considerably lower than camera-based estimates of density for uncharacterized bottom sites (Tables 11.1.4-11.1.6). Similar to camera-based approaches, hydroacoustic-based greater amberjack density at platforms and artificial reefs increased with depth in both states, with the highest estimates in the 70-150 m zone. Interestingly, platform densities derived from the hydroacoustic-based approach were greater than those with either camera-based approach. Hydroacoustic-based densities at Texas artificial reef sites were higher relative to camera-based densities in Texas, but the opposite patter was observed for Louisiana (Tables 11.1.4-11.1.6). Still habitat-specific abundance was comparable between the approaches. Densities on wreck sites in Louisiana and Texas were very low and amberjack were limited to a small number of sites, consistent with the video results and likely reflecting the general lack of structured habitat at most wreck sites in the region rather than low amberjack abundance on this reef type. Overall, hydroacoustic- and camera-based method produced similar spatial patterns of greater amberjack distribution. Differences in magnitude were observed with camera-based estimates generating higher abundance estimates for natural reef and uncharacterized habitats, but often lower estimates at artificial habitats (platform and artificial reefs).

Absolute Abundance Estimates

Absolute abundance estimates of greater amberjack across the study area were produced separately for camera-based (MaxN, Total Count) and hydroacoustic-based approaches (Table 11.1.7). Camera-based estimates for uncharacterized bottom were highly uncertain, and thus, the total absolute abundance is reported both including and excluding uncharacterized bottom for those approaches (Table 11.1.7, Figure 11.1.6). All approaches are presented in the tables; however, we opted to use the MaxN approach for overall camera-based estimates of absolute abundance to maintain continuity with other regions of the Gulf. Absolute abundance of

greater amberjack across the western Gulf estimated using the camera-based MaxN approach was just over 2.5 million with a standard error of approximately 1.7 million (Table 11.1.7). Excluding uncharacterized bottom, absolute abundance estimates from the camera-based MaxN approach were just over 821,000 amberjack with a standard error of approximately 286,000. Hydroacoustic-based estimates of absolute abundance for the western Gulf were nearly 616,000 amberjack with a standard error of just under 212,000. While camera-based estimates of abundance were generally higher than hydroacoustic-based estimates, substantial overlap in the estimates was observed between 800,000 and 1,125,000 individuals (Figure 11.1.6), suggesting this range may represent the most likely abundance of greater amberjack across the shelf habitats of the western Gulf.

With the exception of uncharacterized bottom, abundance estimates were highest on natural reefs in the western Gulf, with estimates of approximately 750,000 fish (MaxN), 1.4 million fish (total count), and 125,000 fish (hydroacoustic). Total abundance estimates of amberjack on platforms in the western Gulf was lowest using the MaxN approach (~19,000), with higher abundance estimates generated from the total count (~42,000) and hydroacoustic (~57,000) approaches. Abundance estimates at artificial reefs was variable ranging from just over 50,000 amberjack using the MaxN approach to over 85,000 using the hydroacoustic approach. Still, differences in abundance estimates between approaches for platforms and artificial reefs had little effect on overall abundance estimates. Across all approaches, while artificial habitats did have higher densities of amberjack, abundance estimates were typically far greater at natural reef and uncharacterized bottom habitats in the western Gulf relative to artificial habitats. In general, uncertainty was higher on uncharacterized bottom and to a lesser extent natural reef, likely due in part to the expansive areal coverage of these habitats relative to artificial reefs.

Depth Use

Greater amberjack relative depth use (standardized from 0 to 1, where 0 is the surface and 1 is the seafloor) varied with habitat association (Figure 11.1.7; mean = 0.69, sd = 0.21). Greater amberjack appeared deeper in the water column on natural reefs (0.83, 95% CRI = 0.80–0.86) and uncharacterized bottom (0.85, 95% CRI = 0.75–0.94) compared to the artificial structures. Mean relative depth use did not differ between artificial reefs (0.62, 95% CRI = 0.60–0.64) and petroleum platforms (0.62, 95% CRI = 0.58–0.65). Model results suggest greater amberjack are most likely to use the bottom half of the water column on all habitats but are observed higher in the water column on artificial structures, likely due to the increase in vertical habitat (see also Movement Chapter). The precision parameter was estimated at 3.93 (95% CRI = 3.55–4.31), indicating low variability in relative depth around the mean.

Length Distribution and Catch Curve Analysis

Greater amberjack length distribution varied by depth, with smaller individuals occupying shallow to mid-depth zones and the largest individuals found in deeper waters (Figure 11.1.8). Smaller greater amberjack (<936 mm FL) were significantly more likely to occupy shallow to mid-depth zones (Figure 11.1.9). However, there was no difference between size distribution in the shallow (584 mm, 95% CRI = 386–774) and mid (862 mm, 95% CRI = 786–936) depth zones. The largest greater amberjack were observed in the deep depth zone (1086 mm, 95% CRI = 991–1182), suggesting that fish length increases with depths (or farther from shore).

Fish length was also used to examine size structure of the greater amberjack population in the western Gulf by depth zone. The 20-40 m depth zone was characterized by smaller fish, with nearly 2/3 of the fish in this depth zone ranging from 400-600 mm FL (Table 11.1.8). Length structure was highly truncated in the deeper zones with 1000+ mm FL accounting for nearly 50% amberjack in the 40-70m depth zone and over 70% of the

amberjack observed in the 70-150 m depth zone. The finding that smaller fish, particularly less than 600 mm FL were relatively rare in the survey, suggests that our estimates of abundance are likely conservative. It seems likely that our survey methods better targeted the adult population while perhaps underestimating the abundance of smaller fish. The age classes used in the catch curve ranged from 1–11 (Figure 11.1.9). The catch curves to the natural log-transformed abundance at age for all study sites produced an estimated Z of 0.28, corresponding to annual mortality rate of 24.8% for greater amberjack in the western Gulf.

CONCLUSIONS

Greater amberjack were broadly distributed across artificial reefs, petroleum platforms, and natural reefs throughout the western Gulf at all sampled depths, with length structure differing by depth stratum. While higher abundances of greater amberjack were observed in Louisiana compared to Texas, patterns of distribution across habitats were similar for both states. In general, camera-based approaches (MaxN and total count) produced higher abundance estimates than the hydroacoustic approach; however, there was considerable overlap in the estimates between 800,000 and 1,125,000, suggesting the most likely abundance estimate falls within this range (Figure 11.1.6). Size structure of greater amberjack in the western Gulf was skewed towards larger individuals (> 1000 mm FL) which made up the majority of the observed fish in the deeper depth zones where abundance was highest. While an ontogenetic shift in which individuals move from shallower to deeper habitat likely explains the increase in size with depth that was observed, abundance was low in the shallower depth zone and smaller fish were relatively rare in the western Gulf surveys. Thus, abundance estimates in the western Gulf presented here generally reflect the abundance of larger (1,000 mm +), adult amberjack and as a result, total abundance estimates could be relatively conservative. It is possible that reduced visibility at nearshore sites in the western Gulf (where fish < 600 mm were most common) contributed to the lack of small fish in the survey; however, it should be noted that the abundance of large fish in Louisiana and Texas is supported by the regional size structure from the tagging components of this study.

Greater amberjack density on wrecks (NOAA Wreck layer) was lower than on other artificial structures; however, field sampling revealed an absence of structural complexity at most wreck sites in both states, suggesting that many of these sites were either non-existent or GPS locations were inaccurate. Regardless, low densities at wreck sites likely reflect a lack of available habitat rather than behavioral avoidance of a particular structure type. As a result, estimates of amberjack over artificial reef habitat in the western Gulf presented represents the fish on state managed reef sites. It is also important to mention that in this study, uncharacterized bottom includes unconsolidated substrate, unmapped natural reef habitat, artificial reefs and wrecks that are not maintained by the states of Louisiana and Texas, and pipeline infrastructure that rises above the substrate. Greater amberjack (and other species) were observed over uncharacterized bottom; however, sites with amberjack contained hardbottom habitat (Figure 11.1.10). Sites with hardbottom habitat were primarily in the 70-150 m depth zone near the shelf edge. This region is characterized by numerous hard bottom features or reef banks and the area in between these known banks contains scattered unmapped hardbottom habitat, which was observed in the surveys. Given the expanse of uncharacterized bottom in the western Gulf and uncertainty in the composition of the benthic habitat, this stratum has a significant effect on survey estimates and associated estimates of variance. High uncertainty in abundance estimates over uncharacterized bottom are thus not surprising given this stratum is essentially the portion of the shelf for which benthic habitat information is not available. Results from this study highlight the uncertainty in abundance estimates over uncharacterized bottom, the need for detailed maps of benthic habitat across the continental shelf of the Gulf, and the implications of unmapped habitats in future efforts and the design of fishery-independent surveys across the region.

Table 11.1.1. Potential sampling universe for each habitat-depth strata in the Western Gulf of America. Values represent the number of cells (5 x 5 nm) containing at least one unit of a particular habitat type from which sample sites were randomly drawn.

State	Depth (m)	Artificial Reefs	Wrecks	Platforms	Uncharacterized Bottom	Natural Reefs
Louisiana	20-40	14	80	75	303	0
	40-70	24	50	64	234	3
	70-150	17	26	57	230	57
Texas Total	20-40	16	54	7	320	6
	40-70	36	36	15	276	10
	70-150	23	10	17	204	37
North Texas	20-40	3	29	3	182	0
	40-70	26	20	10	144	7
	70-150	17	6	16	100	25
South Texas	20-40	13	25	4	138	6
	40-70	10	16	5	132	3
	70-150	6	4	1	104	12

Table 11.1.2. Sampling sites (n) selected for surveys in each habitat-depth strata in the Western Gulf of America.

State	Depth (m)	Artificial Reefs	Wrecks	Platforms	Uncharacterized Bottom	Natural Reefs
Louisiana	20-40	4	17	21	18	0
	40-70	14	9	15	14	1
	70-150	12	4	9	13	14
Texas Total	20-40	8	17	7	19	0
	40-70	14	11	16	16	1
	70-150	8	2	16	17	9
North Texas	20-40	2	9	3	11	0
	40-70	9	4	11	7	1
	70-150	5	1	15	6	6
South Texas	20-40	6	8	4	8	0
	40-70	5	7	5	9	0
	70-150	3	1	1	11	3

Table 11.1.3. Total number of sampling units (number of structures or areal coverage of seafloor) within each depth stratum in the Western Gulf of America.

State	Depth (m)	Artificial Reefs	Wrecks	Platforms	Uncharacterized Bottom (km^2)	Natural Reefs (km^2)
Louisiana	20-40	69	128	240	16,606.0	0
	40-70	237	70	165	12,395.2	9.1
	70-150	202	33	93	10,937.6	406.3
Texas Total	20-40	60	63	17	20,133.1	1.8
	40-70	145	41	27	16,785.4	19
	70-150	75	9	32	10,981.3	287.2
North Texas	20-40	19	34	5	11,793.8	0
	40-70	79	21	17	9,240.1	14.9
	70-150	45	5	31	4,988.1	268.3
South Texas	20-40	40	29	12	8,339.3	1.8
	40-70	36	20	10	7,545.3	4.1
	70-150	21	4	1	5,993.2	14.8

Table 11.1.4: Stratum level camera-based estimates of greater amberjack derived from MaxN across the Louisiana and Texas shelf in the western Gulf of America. Habitat types include platforms (P), uncharacterized bottom (UCB), artificial reefs (AR), wrecks (W), and natural reefs (NR). For each habitat, depth zones are represented as 20 (20-40m), 40 (40-70m), and 70 (70-150m). Density is given as the mean number of amberjack per structure for artificial reefs, platforms, and wrecks and fish per km2 for uncharacterized bottom and natural reef. Estimates are given for each state-habitat combination and totaled for the entire western Gulf. Estimates of standard error (SE) are given for each habitat stratum.

	Louisiana				Texas				Western Gulf Total	
Stratum Depth	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Estimate**	**Estimate SE**
AR	70.75	508	35,940	9,206	53.21	280	14,899	10,680	50,839	14,100
20	18.17	69	1,254	--	0.00	60	0	--		
40	80.93	237	19,179	--	19.72	145	2,860	--		
70	76.77	202	15,507	--	160.53	75	12,040	--		
NR	1,517.61	415.4	630,413	262,170	395.79	306.2	121,190	114,039	751,603	285,899
40, 70	1,517.61	415.4	630,413	--	395.79	306.2	121,190	--		
P	33.23	498	16,550	7,415	37.31	64	2,388	558	18,938	7,436
20	3.64	240	874	--	0	5	0	--		
40	23.57	165	3,889	--	39.80	27	1,075	--		
70	126.74	93	11,787	--	41.04	32	1,313			
UCB	43.35	39,938.8	1,731,374	1,709,492	0	47,899.8	0	0	1,731,374	1,709,492
20	0	16606.0	0	--	0	20,133.1	0	--		
40	0	12395.2	0	--	0	16,785.4	0	--		
70	158.30	10937.6	1,731,374	--	0	10,981.3	0	--		
W	0	231	0	0	0.10	104	10	7	10	7
20	0	128	0	--	0	63	0			
40	0	70	0	--	0.25	41	10			
70	0	33	0	--	--	--	--	--		
TOTAL			2,414,277	1,729,519	--	--	138,487	114,539	2,552,764	1,733,307

Table 11.1.5: Stratum level camera-based estimates of greater amberjack derived from total count across the Louisiana and Texas shelf in the western Gulf of America. Habitat types include platforms (P), uncharacterized bottom (UCB), artificial reefs (AR), wrecks (W), and natural reefs (NR). For each habitat, depth zones are represented as 20 (20-40m), 40 (40-70m), and 70 (70-150m). Density is given as the mean number of amberjack per structure for artificial reefs, platforms, and wrecks and fish per km2 for uncharacterized bottom and natural reef. Estimates are given for each state-habitat combination and totaled for the entire western Gulf. Estimates of standard error (SE) are given for each habitat stratum.

	Louisiana				Texas				Western Gulf Total	
Stratum Depth	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Estimate**	**Estimate SE**
AR	178.57	508	90,714	25,478	108.19	280	30,292	19,828	121,006	32,285
20	47.44	69	3,273	--	0	60	0	--		
40	232.81	237	55,176	--	59.09	145	8,569	--		
70	159.73	202	32,265	--	289.64	75	21,723	--		
NR	2,722.98	415.4	1,131,126	548,837	1,059.01	306.2	324,268	304,122	1,455,394	627,465
40, 70	2,722.98	415.4	1,131,126	--	1,059.01	306.2	324,268	--		
P	77.86	498	38,774	18,063	62.28	64	3,986	832	42,760	18,082
20	5.44	240	1,306	--	0	5	0	--		
40	52.80	165	8,713	--	54.84	27	1,481			
70	309.19	93	28,755	--	78.29	32	2,505	--		
UCB	151.73	39,938.8	6,059,809	5,983,221	0	47,899.8	0	0	6,059,809	5,983,221
20	0	16,606.0	0	--	0	20,133.1	0	--		
40	0	12,395.2	0	--	0	16,785.4	0	--		
70	554.03	10,937.6	6,059,809	--	0	10,981.3	0	--		
W	0	231	0	0	0.13	104	14	9	14	9
20	0	128	0	--	0	63	0			
40	0	70	0	--	0.33	41	14			
70	0	33	0	--	--	--	--			
TOTAL			7,320,423	6,008,422	--	--	358,560	304,769	7,678,983	6,016,146

Table 11.1.6: Stratum level hydroacoustic-based estimates of greater amberjack across the Louisiana and Texas shelf in the western Gulf of America. Habitat types include platforms (P), uncharacterized bottom (UCB), artificial reefs (AR), wrecks (W), and natural reefs (NR). For each habitat, depth zones are represented as 20 (20-40m), 40 (40-70m), and 70 (70-150m). Density is given as the mean number of amberjack per structure for artificial reefs, platforms, and wrecks and fish per km2 for uncharacterized bottom and natural reef. Estimates are given for each state-habitat combination and totaled for the entire western Gulf. Estimates of standard error (SE) are given for each habitat stratum. Note that the wreck layer in Texas had to be collapsed

	Louisiana				Texas				Western Gulf Total	
Stratum Depth	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Density**	**Area or # of Objects**	**Estimate**	**Estimate SE**	**Estimate**	**Estimate SE**
AR	43.48	508	22,090	6,764	225.47	280	63,133	39,273	85,223	39,851
20	0.35	69	24	--	1.33	60	80	--		
40	37.72	237	8,940	--	55.46	145	8,041	--		
70	64.98	202	13,125	--	733.49	75	55,012	--		
NR	299.56	415.4	124,437	33,219	4.56	306.2	1,397	1,021	125,834	33,234
40, 70	299.56	415.4	124,437	--	4.56	306.2	1,397	--		
P	106.38	498	52,976	29,395	58.07	76	4,413	1,262	57,389	29,422
20	4.93	240	1,182	--	1.81	17	31	--		
40	32.06	165	5,290	--	57.56	27	1,554	--		
70	500.05	93	46,504	--	88.38	32	2,828			
UCB	7.14	39,938.8	285,160	202,407	1.29	47,899.8	61,752	18,553	346,912	203,255
20	0	16,606.0	0	--	0	20,133.1	0	--		
40	0	12,395.2	0	--	0	16,785.4	0	--		
70	26.07	10,937.6	285,160	--	5.62	10,981.3	61,752	--		
W	1.40	231	324	228	0.94	113	107	75	430	240
20	0	128	0	--	--	--	--	--		
40	0.95	70	66	--	--	--	--	--		
70	7.81	33	258	--	--	--	--	--		
20,40,70	--	--	--	--	0.94	113	107	--		
TOTAL			484,987	207,320			130,801	43,465	615,788	211,828

Table 11.1.7: Absolute abundance estimates of greater amberjack in the western Gulf using camera-based (MaxN and Total count) and hydroacoustic-based approaches. Total estimates using camera-based approaches are presented both inclusive and exclusive of uncharacterized bottom habitat (UCB) due to uncertainty in this stratum.

MaxN		
Stratum	**Estimate**	**Estimate SE**
Artificial Reef	50,839	14,100
Natural Reef	751,603	285,899
Platform	18,938	7,436
UCB	1,731,374	1,709,492
Wrecks	10	7
Total	2,552,764	1,733,307
Total without UCB	821,390	286,343
Total count		
Artificial Reef	121,006	32,285
Natural Reef	1,455,394	627,465
Platform	42,760	18,082
UCB	6,059,809	5,983,221
Wrecks	14	9
Total	7,678,983	6,016,146
Total without UCB	1,619,174	628,555
Hydroacoustic		
Artificial Reef	85,223	39,851
Natural Reef	125,834	33,234
Platform	57,389	29,422
UCB	346,912	203,255
Wrecks	430	240
Total	615,788	211,828

Table 11.1.8: The proportion of each size class of greater amberjack within each depth zone in the western Gulf.

Depth Zone	0–400	400–600	600–800	800–1000	1000 +
20–40m	0.111	0.667	0.111	0.000	0.111
40–70m	0.029	0.192	0.192	0.125	0.462
70–150m	0.000	0.031	0.063	0.203	0.703

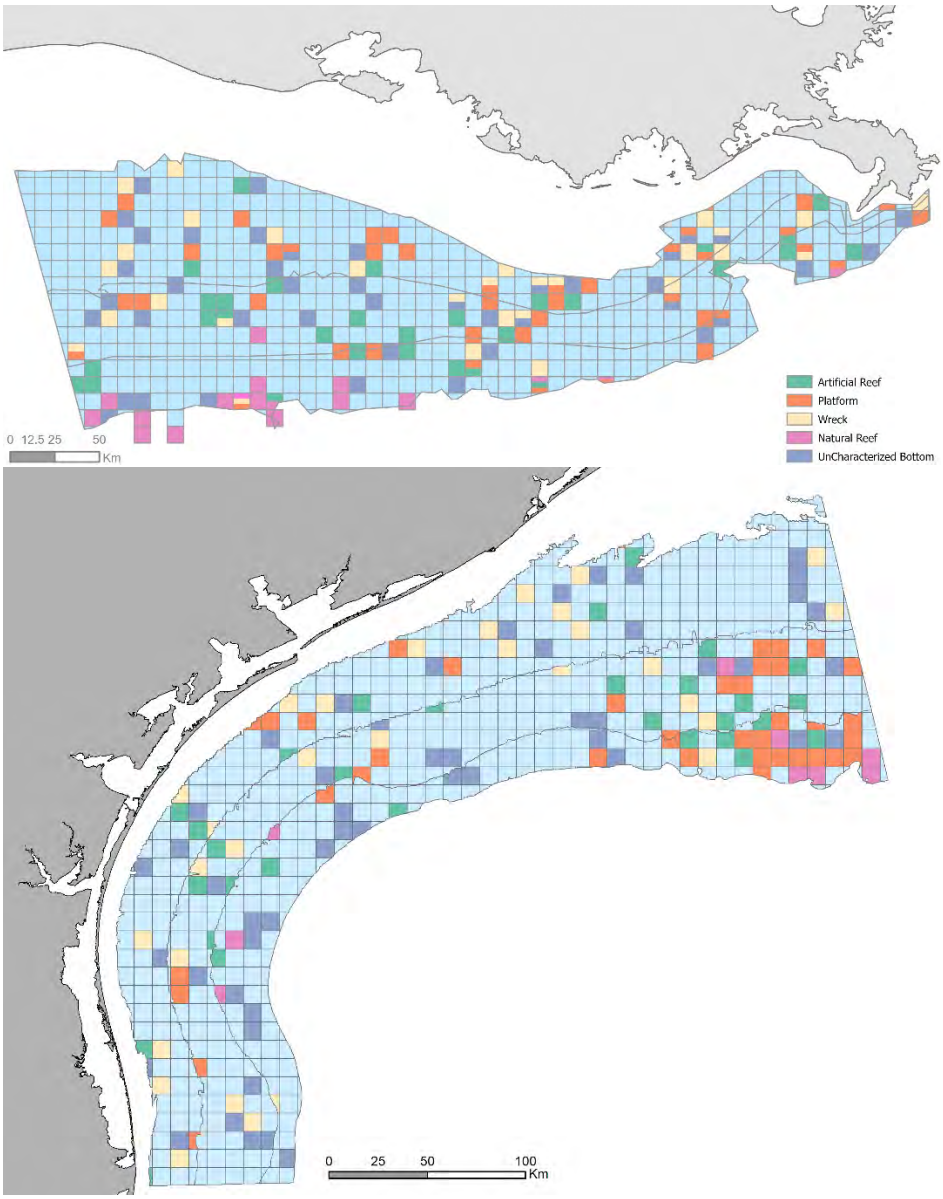


Figure 11.1.1. Distribution of sampling cells randomly selected to estimate greater amberjack abundance in the Louisiana (upper), North Texas (lower left), and South Texas (lower right). Colored cells correspond to habitats sampled using hydroacoustic and videos surveys. Cells with multiple colors indicate cells that were separately selected for more than one stratum. Each grid cell represents 5 x 5 nautical miles and depth zones (20-40m, 40-70m, and 70-150m) are indicated by contours.

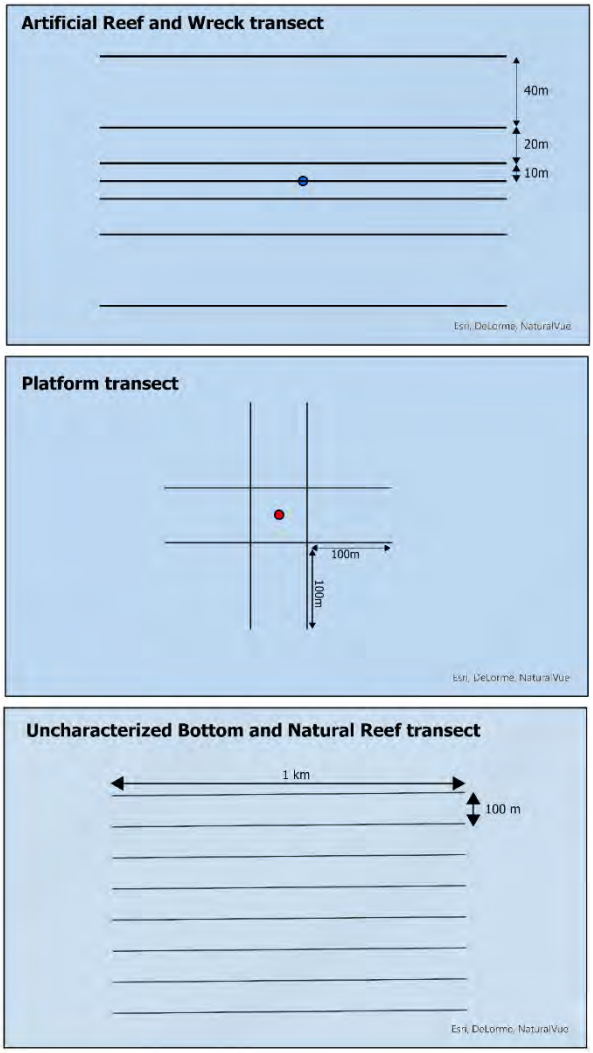


Figure 11.1.2. Hydroacoustic transects performed in the Western Gulf of America for artificial reefs and wrecks, platforms, and uncharacterized bottoms and natural reefs.

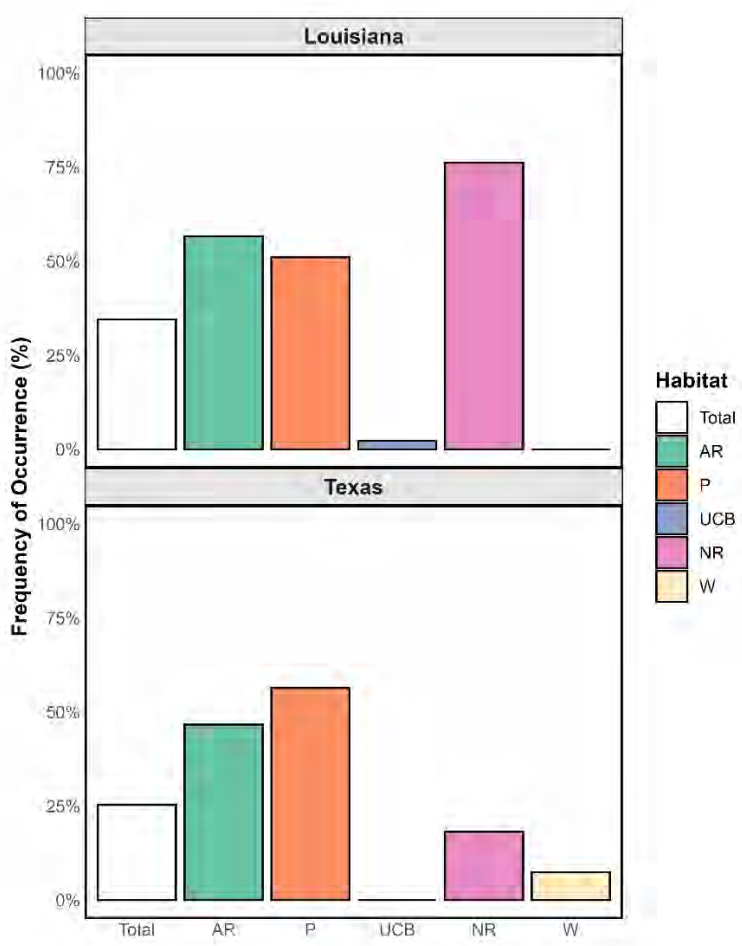


Figure 11.1.3. Frequency of occurrence of greater amberjack (*Seriola dumerili*) observed from stereo video surveys on multiple habitat types in western Gulf of America in Louisiana and Texas.

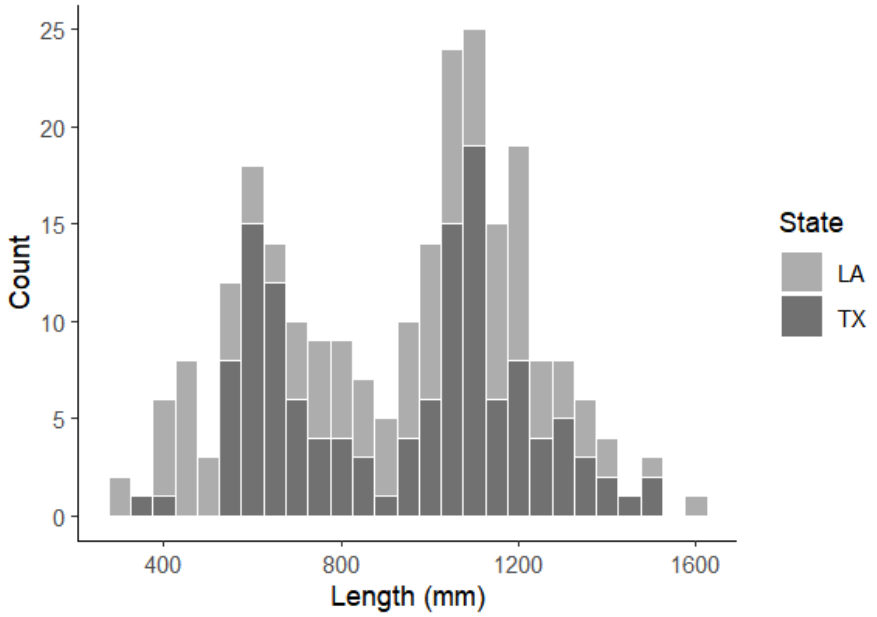


Figure 11.1.4. Length frequency distribution for greater amberjack observed in the western Gulf. LA = Louisiana, TX = Texas.

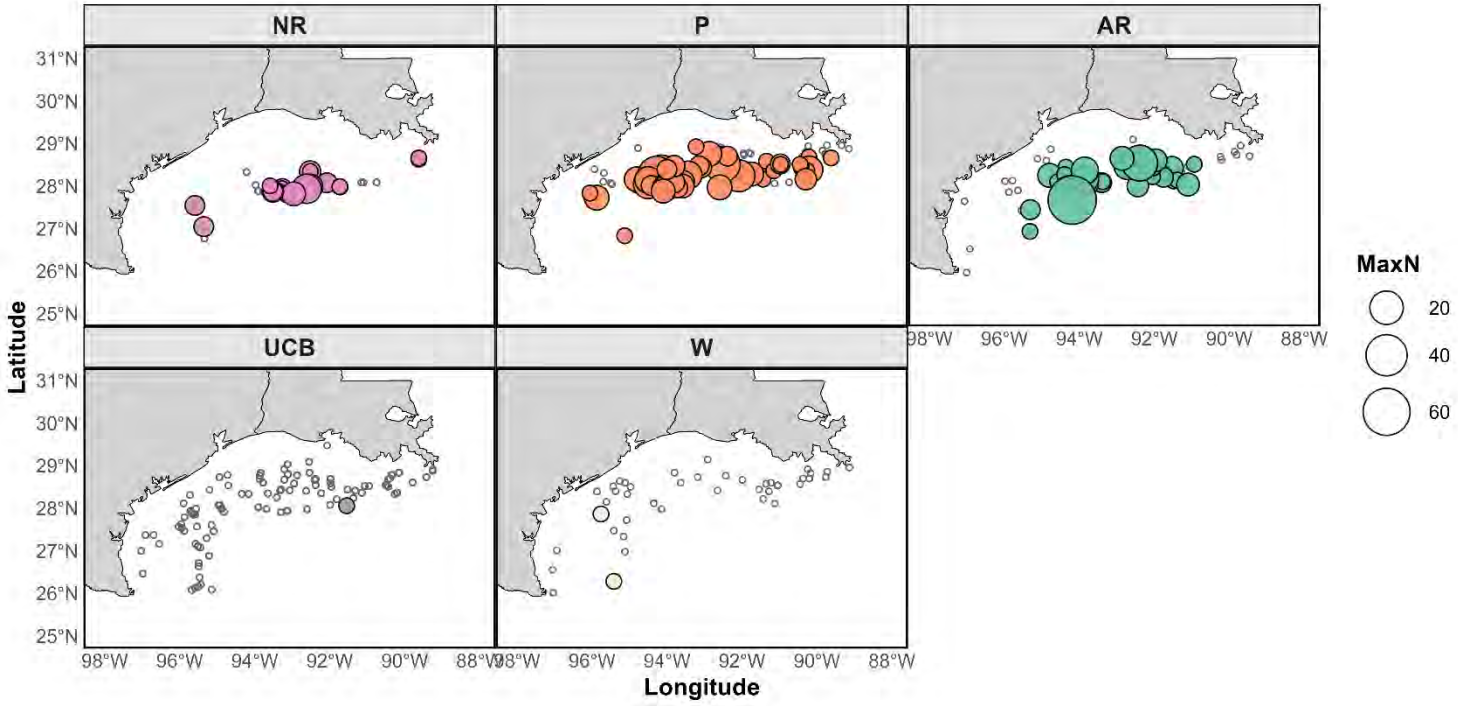


Figure 11.1.5. Spatial distribution of site- and habitat-specific camera-based estimates of relative abundance (MaxN) of greater amberjack (*Seriola dumerili*) in the western Gulf of America from Louisiana to south Texas. The size of the bubble represents the number of amberjack observed and the grey open points represents sites we sampled but did not see amberjack. NR = natural reef, P = platform, AR = artificial reef, UCB = uncharacterized bottom, W = wreck.

Figure 11.1.6. Absolute abundance estimates of greater amberjack in the western Gulf derived from hydroacoustic-based and camera-based approaches. The MaxN approach is shown both inclusive and exclusive of uncharacterized bottom (UCB), while the Total Count approach is shown exclusive of UCB. Shaded orange area shows range for which at least three approaches overlap (~800,00-1,125,000 fish).

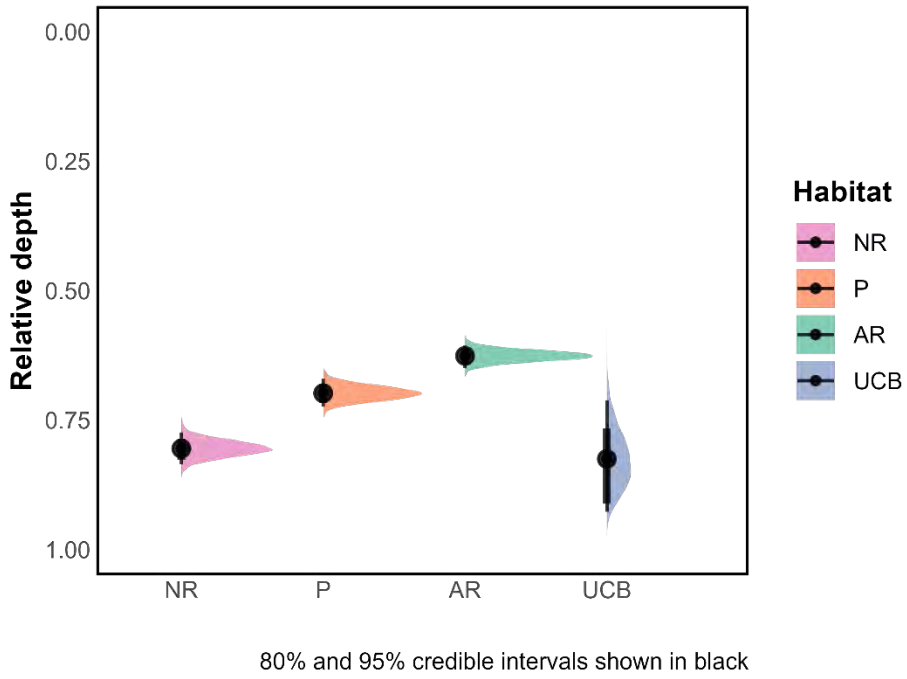


Figure 11.1.7. Relative depth use of greater amberjack (*Seriola dumerili*) on multiple habitat types in the western Gulf of America. A) Density plot of observed relative depths on each habitat type. B) Mean predicted relative depth use from Bayesian linear model and posterior distributions of the means. Associated 80% and 95% credible intervals are show in black lines. NR = natural reef; P = platform; AR = artificial reef; UCB = uncharacterized bottom.

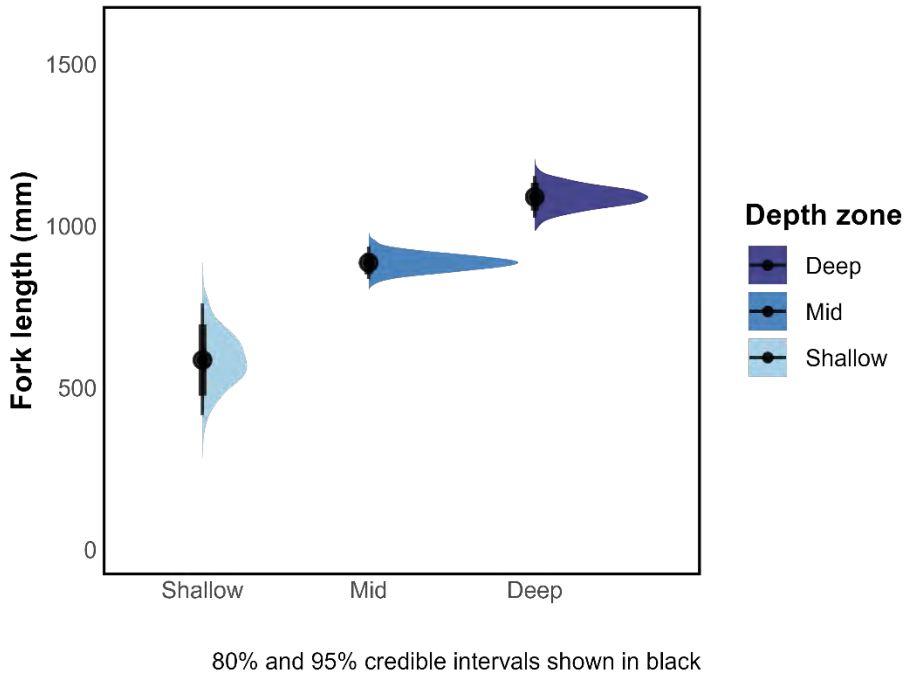


Figure 11.1.8. Length distribution of greater amberjack (*Seriola dumerili*) in multiple depth zones in the western Gulf of America. A) Density plot of observed fork lengths (mm) on each habitat type. B) Mean predicted lengths in each depth zone from Bayesian linear model and posterior distributions of the means. Associated 80% and 95% credible intervals are shown in black lines. Shallow = 20 – 40m; Mid = 40 – 70m; Deep = 70 – 150m.

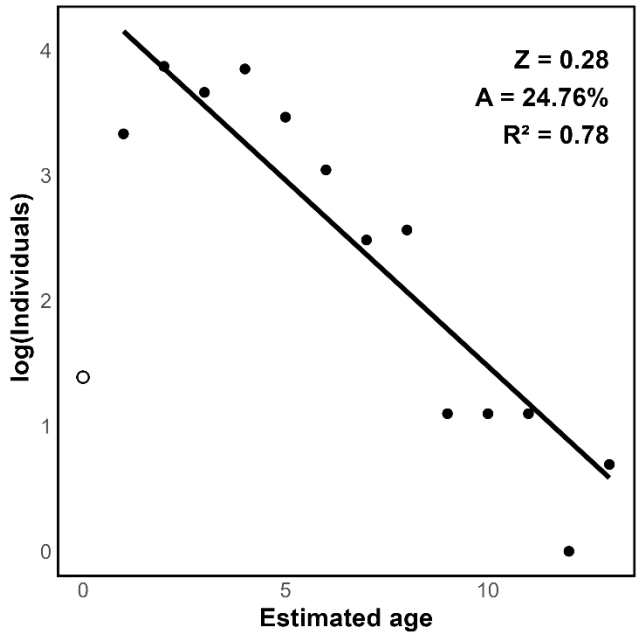


Figure 11.1.9. Catch curve fit to the estimated age distributions (1+) of greater amberjack (*Seriola dumerili*) observed with stereo-video at all study sites from Louisiana to south Texas in the western Gulf of America.



Figure 10. Examples of uncharacterized bottom sites in the western Gulf. The majority of sites were characterized by mud, silt, or sand (top); however, other sites contained hard bottom habitat (bottom).

REFERENCES

Bacheler, N. M., Gregalis, K. C., Gillum, Z. D., Pickett, E. P., Schobernd, C. M., Schobernd, Z. H., & Teer, B. Z. (2022). Using stationary video data to infer relative abundance and distribution of four *Seriola* species along the southeast United States Atlantic coast. Fisheries Research, 249, 106238. https://doi.org/10.1016/j.fishres.2022.106238

Badalamenti, F., Anna, G. D. ', & Mazzola, A. (1995). Feeding habits of young-of-the-year greater amberjack *Seriola dumerili* (Risso, 1810) along the N/W Sicilian Coast. Scientia Marina, 59(3–4), 317–323.

Baker, R., Bilbrey, D., Bland, A., D'alonzo, F., Ehrmann, H., Havard, S., Porter, Z., Ramsden, S., & Rodriguez, A. R. (2022). Underwater Video as a Tool to Quantify Fish Density in Complex Coastal Habitats. Diversity, 14(50). https://doi.org/10.3390/d14010050

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80. https://doi.org/10.18637/jss.v080.i01

Dance, M. A., Patterson, W. F., & Addis, D. T. (2011). Fish community and trophic structure at artificial reef sites in the northeastern Gulf of Mexico. Bulletin of Marine Science, 87(3), 301–324. https://doi.org/10.5343/bms.2010.1040

Denney, C., Fields, R., Gleason, M., & Starr, R. (2017). Development of new methods for quantifying fish density using underwater stereo-video tools. Journal of Visualized Experiments, 129, e56635. https://doi.org/10.3791/56635

Gallaway, B. J., Raborn, S. W., McCain, K. A., Beyea, R. T., Dufault, S., Heyman, W., Putman, N. F., & Egerton, J. (2021). Absolute Abundance Estimates for Red Snapper, Greater Amberjack, and Other Federally Managed Fish on Offshore Petroleum Platforms in the Gulf of Mexico. North American Journal of Fisheries Management, 41(6), 1665-1690. https://doi.org/10.1002/nafm.10678

Gardner, C., Goethel, D. R., Karnauskas, M., Smith, M. W., Perruso, L., & Walter, J. F. (2022). Artificial Attraction: Linking Vessel Monitoring System and Habitat Data to Assess Commercial Exploitation on Artificial Structures in the Gulf of Mexico. Frontiers in Marine Science, 9. https://doi.org/10.3389/fmars.2022.772292

Garner, S. B., Boswell, K. M., Lewis, J. P., Tarnecki, J. H., & Patterson, W. F. (2019). Effect of reef morphology and depth on fish community and trophic structure in the northcentral Gulf of Mexico. Estuarine, Coastal and Shelf Science, 230, 106423. https://doi.org/10.1016/j.ecss.2019.106423

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. PeerJ, 2019(5). https://doi.org/10.7717/peerj.6876

R Core Team (2024). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Wells, R. J. D., & Rooker, J. R. (2004). Distribution, age, and growth of young-of-the-year greater amberjack (*Seriola dumerili*) associated with pelagic Sargassum. Fisheries Bulletin, 102, 545–554.

Appendix 11.1.1. Summary table of sampling effort including camera and hydroacoustic surveys in the western GOM in Louisiana (LA), north Texas (NTX), and south Texas (STX) across all habitat types. Data includes the state, cruise, date, number of days (Days), distance from port (dist km), number of sites (Sites), number of camera surveys (Camera) and number of hydroacoustic surveys (Hydroacoustic). P = platform, AR = artificial reef, W = wreck, NR = natural reef, UCB = uncharacterized bottom.

State	Area sampled	Date	Days	Dist (km)	Sites	P	AR	W	NR	UCB	Camera	Acoustic
Louisiana												
	West Delta	05/22/2023 - 05/23/2023	2	83	17	6	4	3	1	3	33	17
	South Timbalier	06/06/2023 - 06/08/2023	3	87	22	8	2	5	0	7	50	22
	Eugene Island	06/11/2023 - 06/12/2023	2	171	17	4	4	5	0	4	33	17
	Ship Shoal	06/24/2023 - 06/26/2023	3	144	20	10	3	3	1	3	36	20
	Marsh Island	06/29/2023 - 06/30/2023	2	206	19	7	6	1	0	5	39	19
	Vermillon	07/17/2023 - 07/19/2023	3	245	21	5	7	2	1	6	48	20
	West Cameron	07/21/2023 - 07/23/2023	4	285	27	6	4	7	0	10	61	21
	Bright - Elvers - Geyer	8/21/2024	1	240	3	0	0	0	3	0	9	0
	Bouma - Parker	07/02/2024 - 07/03/2024	2	222	3	0	0	0	3	0	10	0
			22		**149**	**46**	**30**	**26**	**9**	**38**	**319**	**136**
North Texas												
	Lease blocks - TX5 and TX6	08/23/2023 - 08/27/2023	3	98	14	3	3	1	0	7	17	14
	Lease blocks - TX6A and TX7B	08/30/2023 - 09/01/2023	1	120	3	1	2	0	0	0	3	3
	Lease blocks - TX7B and TX7C	09/01/2023 - 09/06/2023	4	205	36	18	10	4	1	3	50	36
	Lease blocks - TX5, TX6, TX7A, TX7B, TX7C	09/26/2023 - 09/29/2023	3	188	21	6	1	9	0	5	37	21
			11		**74**	**28**	**16**	**14**	**1**	**15**	**107**	**74**
South Texas												
	Survey1	7/25/2023	1	42	2	0	1	0	0	1	7	2
	Survey2	8/3/2023	1	57	3	0	1	1	0	1	7	3

State	Area sampled	Date	Days	Dist (km)	Sites	P	AR	W	NR	UCB	Camera	Acoustic
	Survey3	8/29/2023	1	56	5	0	1	3	0	1	6	5
	Survey4	9/8/2023	1	68	3	0	1	0	1	1	9	3
	Survey5	9/13/2023	1	61	4	2	1	0	0	1	7	4
	Survey6	9/19/2023	1	85	4	1	1	1	0	1	7	4
	Survey7	9/27/2023	1	117	4	2	1	1	0	0	6	4
	Survey8	9/28/2023	1	41	5	1	1	2	0	1	6	5
	Survey9	10/9/2023	1	69	3	0	2	0	0	1	7	3
	Survey10	10/17/2023	1	70	4	0	1	2	0	1	7	4
	Survey11	12/4/2023	1	13	1	0	1	0	0	0	1	1
	Survey12	12/5/2023	1	66	4	2	1	0	0	1	7	4
	Survey13	1/29/2024	1	22	1	0	0	0	0	1	3	1
	Survey14	1/30/2024	1	81	6	0	0	5	0	1	5	6
	Survey15	1/31/2024	1	28	3	0	1	1	0	1	5	3
			15		**52**	**8**	**14**	**16**	**1**	**13**	**90**	**52**
Collaborative cruise												
Louisiana	Flower Garden	08/11/2023 - 08/15/2023	4	300	18	1	0	0	11	6	86	18
North Texas	Flower Garden and TX5B, TX6A, TX7C	08/16/2023 - 08/24/2023	4	495	16	1	0	0	6	9	73	16
South Texas	Deep UCB and Natural Sites	08/19/2023 - 08/23/2023	4	104	15	0	0	0	2	13	47	15
			12		**49**	**2**	**0**	**0**	**19**	**28**	**206**	**49**

11.2. Alabama-Mississippi

Greater amberjack (*Seriola dumerili*) abundance, distribution, and size-structure on natural and artificial habitats of the north-central Gulf of America

Mark A. Albins, John Hoenig, Julie Gross, Kevin Boswell, Allison White, Crystal Hightower, Sean P. Powers

Introduction

The primary goal of this study was to estimate the abundance of Greater amberjack across the continental shelf (20 to 150m depth) in the Alabama-Mississippi region by habitat (artificial hardbottom, natural hardbottom, and uncharacterized bottom) and age-size class. The seafloor in this region is dominated by unconsolidated sand, mud, and silt, with a large number of artificial reefs in the shallow and mid depth zones (20 to 60m) and a significant amount of shelf-edge associated natural reef in the deeper zone (60 to 150m). An ongoing, long-term (2011 through 2021), systematic, randomized side-scan mapping program, conducted by Dr. Sean Powers' Lab at the University of South Alabama, provided extensive and detailed information on the distribution of both natural and artificial hardbottom habitat across the Alabama portion of this region. However, habitat data for the Mississippi portion of the region consists primarily of lists of artificial reef locations (MDMR, NOAA, BOEM), along with some limited and inconsistent information on reef sizes and relief. Historical mapping work conducted by the USGS (2002) also indicates a substantial amount of natural reef associated with the shelf edge off both Alabama and Mississippi, but does not provide sufficient resolution to delineate reef areas with high enough precision for targeted habitat-specific sampling.

The differences in existing habitat data for the two subregions (Alabama and Mississippi), necessitated separate sampling designs for known hardbottom habitats. Additionally, in order to maximize the data available for our abundance estimate, we designed our Greater Amberjack Hardbottom Survey (GAJ) to compliment and leverage an ongoing, established, reef-fish survey program (Fisheries Independent Survey [FIS], Powers Lab, USA). All three of these designs (Mississippi Survey: MS; Alabama Fisheries Independent Survey: FIS, and Alabama Greater Amberjack Survey: GAJ) targeted hard-bottom habitats and used two-stage cluster sampling, with grid cells randomly selected from a universe of grid cells and reefs within grid cells randomly sampled from reefs present in selected cells (details below). A separate design (UCB) targeted uncharacterized bottom habitat (presumed to be comprised predominantly of unconsolidated sand, silt, and mud seafloor). While three different sampling designs were used to estimate the population over hardbottom habitats in the regional abundance estimation (MS, FIS, and GAJ), these field surveys were all conducted using the same gears, protocols and personnel. In contrast, the UCB survey of uncharacterized bottom habitats was based around use of different gear type and designed to be additive with the hardbottom habitat estimates.

Field Survey Methods

Habitat delineation

The Alabama-Mississippi region extended from the Mouth of the Mississippi river in the west (89°08'W) to the Alabama-Florida Border in the east (87°32'W), and comprised two subregions: Mississippi and Alabama, with the Alabama subregion further divided into six zones based on depth (Shallow: 18-37 m, Mid: 37-55m, and Deep: 55-91m) and inside vs. outside of the Alabama Artificial Reef Zone (AARZ). The zonation scheme used was based on leveraging ongoing habitat mapping and reef-fish monitoring efforts in the Alabama subregion as well as

inherent differences in available habitat data (Figure 12.2.1). Further description and rationale for this spatial partitioning can be found in Gregalis et al. (2012) and Powers et al. (2018, 2023).

There were three primary pre-existing sources of habitat data for this region: (1) lists of artificial reefs, wrecks and energy infrastructure maintained by state (ADCNR, MDCNR) and federal agencies (BOEM, NOAA), and available in online repositories (Figure 12.2.2), (2) mapping data from a 2001 USGS survey of the outer shelf natural reef complex (Figure 12.2.3, Gardner et al., 2002), and (3) an ongoing (2011 to present) comprehensive side-scan habitat mapping effort conducted in the Alabama subregion for the purpose of supporting a comprehensive fisheries independent reef-fish survey (FIS) program (Figure 12.2.4, Gregalis et al. 2012, Powers et al. 2018, 2023). Based on these sources, the majority of bottom habitat in shallow (18 to 36 m) and mid-depth (37 to 55 m) areas within this region can be characterized as unconsolidated mud and sand with an extensive network of artificial reefs (2 to 3 structures per km2) but almost no natural, emergent hardbottom structure. Bottom habitat in the deep zone (56 to 110 m) of the Alabama Artificial Reef Zone (AARZ) includes an extensive area of natural hard bottom, including rocks, boulders and pinnacles centered around the 70 m isobath. The most comprehensive survey of this outer shelf natural reef complex, conducted by the USGS (Figure 12.2.3), mapped approximately 1,625 km2 of rock fields, banks and pinnacles scattered throughout a sand bottom. Approximately 75% of this area is off Alabama and 25% is off Mississippi. The USGS survey did not record data at a suitable resolution to differentiate the spatial extent of hard bottom areas from sand. Fortunately, there is some overlap between this outer shelf reef complex and the more recent quantitative mapping effort of the FIS program.

To estimate the extent of artificial and natural hard bottom in the region, we relied primarily on the extensive, quantitative Fisheries Independent Survey (FIS) side-scan mapping dataset with supplemental information from other sources used to fill gaps in knowledge and complete the overall picture.

FIS Side-scan Mapping Methods and Design

The Alabama offshore area, extending from 88°21’33’’W to 87°32’13’’W and from 30°03’10’’N to the 110m isobath (approx. 29°19’N) was divided into 1423 2-by-2km grid cells. The resulting grid covers most of the area, with some gaps in coverage to avoid operations in major shipping lanes. This grid is further broken down into six strata defined by two factors: depth (shallow = 18 to 37m, mid = 37 to 55, and deep = 55 to 110m) and whether grid cells are inside or outside the Alabama Artificial Reef Permit Zone (AARPZ). Strata A, B, and C are inside the reef permit zone and strata E, F, and G are outside the zone, with A and E in shallow, B and F in mid, and C and G in deep depth categories (Table 12.2.1).

Each year, from 2010 to 2021, we randomly selected some number, depending on funding, of previously unscanned grid cells for side-scan sonar mapping. The allocation of these selections was proportional to the total number of cells in each stratum. Selected cells were surveyed using either an EdgeTech 4200 dual frequency side-scan sonar (300/600 kHz) and a Biosonic echosounder with a 200 kHz single beam transducer, deployed using a side-scan towfish towed at an altitude of approximately 15 m above the seafloor, or a pole-mounted EdgeTech 6205 multiphase echosounder (230/550 kHz). Position, side-scan sonar, and cable-out data were recorded and integrated using Chesapeake Technology Inc. Sonarwiz7 software. This software produced a real-time, fully geo-referenced mosaic of the side-scan sonar data and served as a navigational aid for the vessel during the survey. Bottom targets visualized by the Sonarwiz7 program were captured and displayed on the chart plotter of the program. Based on the side-scan generated map of structures, a contact report was

produced in which the length, width, height, description, latitude, and longitude of each contact (i.e., each structure) within each grid was marked.

The total number of grid cells mapped during this period was 491 (Figure 12.2.4, Table 12.2.1). All contacts in each of these 491 grid cells were identified, digitized, measured, and georeferenced based on expert interpretation of the multiphase side-scan sonar data. We then identified all qualifying reefs as those with height above the seafloor of at least 0.5 m and area of at least 4 m2, following Gregalis et al. (2012). Of the 491 scanned grid cells, 460 contained qualifying reefs by this definition. All qualifying reefs were then categorized based on origin (artificial or natural) and vertical relief (low relief = 0.5 to 3.49 m, high relief ≥ 3.5 m). We were unable to definitively determine the origin of a small percentage of qualifying reefs (1.53 %). The heights and areas of these features (height = 1.08 ± 0.52 m, area = 13.18 ± 11.63 m2 [mean ± SD]) were consistent with artificial reefs in the region and for the purposes of the sample design and abundance estimation, these were treated as artificial reefs. We then calculated the mean and variance across cells for the number and total areal footprint of artificial and natural reefs for each stratum (Tables 12.2.2 and 12.2.3). We provide detailed reef counts and total reef areas for all 491 grid cells mapped between 2010 and 2021 as a data supplement (Appendix E). This supplemental spreadsheet also includes indicators for which cells were included in the sampling universe for the surveys described below.

The stratum-specific reef counts and areal footprints of artificial and natural hardbottom habitats in the Alabama sub-region were estimated from the above-described mapping data by simply calculating the mean number and area of reefs across mapped grid cells in each stratum and then multiplying by the total number of grid cells in each stratum. Estimates along with standard deviations are presented in tables 12.2.4 and 12.2.5.

To estimate the total area of natural hardbottom in the outer shelf reef complex, we applied area estimates from the portion of the reef complex mapped by the FIS program to the entire reef complex area that exists outside the AARZ. This assumes that the density of natural reefs in the FIS mapped portion is similar to that found across the unmapped portion. Based on this calculation, we estimate that there is an additional 4.55 km^2 of high-relief natural reef and 6.48 km^2 of low-relief natural reef in the deep stratum that was not covered by the FIS grid mapping design. Additionally, based on the BOEM database of currently standing rigs and platforms, we estimate the total number of standing oil and natural gas platforms in the Alabama subregion at 18, with one in the shallow zone, four in the mid-depth zone and 13 in the deep zone.

There is not sufficient information about hard bottom habitat off Mississippi to generate total habitat area estimates. However, we make the simplifying assumption that all artificial reefs are known and included in the published MDCNR artificial reef list and rigs-to-reefs list, the BOEM list of currently standing platforms, and the NOAA wrecks and obstructions list. We also assume that the only natural hardbottom reef is included in the USGS mapped outer shelf reef complex. We, again, apply the reef area estimates from that portion of the region that has been mapped quantitatively (FIS) to the entire reef complex. Table 12.2.6 summarizes the number of artificial reefs and area of natural reef in the Mississippi sub-region.

ROV video and hydroacoustic surveys of hardbottom habitats

The survey protocol for hardbottom habitats in the Alabama-Mississippi region included use of both hydroacoustics and ROV mounted video cameras to record fish communities and estimate the abundance of greater amberjack.

Hydroacoustic surveys were performed using two Simrad EK80 transducers: ES 120-7C and ES 70-7C. Transducers were coupled with a motion reference unit and mounted to a towfish which was pulled at a depth of 0.8 m approximately 10 m behind the vessel. Transducers were calibrated during surveys using 32 mm and 22 mm diameter tungsten spheres on all vessels. The hydroacoustic transducer was used to survey an area including and immediately adjacent to the target habitat. We used two different acoustic survey transect patterns in the region. For all natural and artificial reefs except standing oil and gas platforms, we ran seven 200 m long transects with 10 m spacing between the center transect and the first concentric pair of neighboring transects, 20 m between the first concentric pair and the second concentric pair of transects, and 40 m between the second pair of concentric transects and the outermost pair of transects. The center transect was positioned over the coordinates of the camera deployment or target structure. For surveys at oil and gas platforms, we used an open star pattern which consisted of three to eight transects driven alongside the platforms as close to the structure as was deemed safe by the vessel captain. Transects were driven along each side of the platform complex (or between individual platforms in the case of a cluster of structures not connected above the water), and each transect extended 100 m beyond the structure on either side (Figure 2).

Video surveys were conducted using a four thruster Outland 2000 ROV equipped with a high-definition, 1080-line, color camera. The ROV was also equipped with a sonar with a 75 m detection range and 360◦ viewing capabilities allowing the operator to approach large structures. The ROV maneuvers at approximately 0.25 m·s−1 and 3–4 m from the bottom. The ROV umbilical (250 m) was attached to a 10 kg depression weight, used to reduce the umbilical's catenary. The terminus of the depression weight was maintained 5–10 m from the seafloor, followed by 50 m of unweighted umbilical cord suspended with low buoyancy floats. For each site, the ROV was positioned 5–7 m away from the structure and the cameras pointed at the structure. Five minutes of video was recorded. The process was repeated on the opposite side of the structure. Video imagery from the ROV was recorded in HD and analyzed in the laboratory.

ROV video processing for hardbottom habitats

Fish recorded by the ROV were identified to the lowest possible taxa, enumerated, and measured (see below). For highly mobile fish (e.g., amberjack, red snapper, etc.), abundance was estimated using the MaxN method, wherein the still frame with the most fish visible represents the minimum number of fish present in the sampled area (Schobernd et al. 2013). The still frame used for the maxN count is the frame that has the highest number of that species in view. This method produced an index of relative abundance for each species on each ROV dive. For species that were slow moving (e.g., gray triggerfish) or near sedentary (red lionfish) total counts of fish on the reef could be made without concern of double counting fish. For these species, this method produced an estimate of absolute abundance per reef. Measurements (total length, TL) of some fish were possible because the ROVs were equipped with a pair of Digi-Key 2.5 milliWatt red lasers aligned in parallel, separated by 3-cm as a frame of reference (Caimi and Tusting 1987). Fish must be nearly perpendicular to the camera and have both lasers illuminate their body to be measured.

Camera-based estimates of site-specific density for hardbottom habitats

In order to estimate site-specific density of GAJ we used the MaxN count divided by the portion of the effective sampling area (ESA) of the camera gear that contained the target habitat type. However, determination of the ROV video camera ESA is not straightforward. The area directly viewed by the ROV video camera during a single survey can be roughly approximated by a circle with a radius equal to the distance of the ROV from the target reef feature (approx. 5m). However, because the ROV remains on the bottom for an extended period of time,

moving animals can approach, enter, and exit the detection range of the gear during the survey. This means that the effective sampling area (ESA) of the camera is likely to be many times larger than the circular detection area. The ESA of any particular gear, for a given target species is determined by the combined probability of the gear detecting an animal when present and the probability of an animal encountering the gear while moving within its home range. This probability can be estimated by calculating the volume under the detection probability surface, which typically declines with distance from the sampling gear. We approximated this volume by the 2-dimensional area of a unit-high cylinder, with the cylinder area estimated as the area encompassed by the 95% kernel density of amberjack positions during a 20-minute time interval (approximate duration of ROV video survey). In order to estimate this area, position data were collected from a sample of 40 acoustically tagged Greater Amberjack across four different virtual positioning system (VPS) receiver arrays in the Alabama artificial reef zone. VPS receiver arrays allow repeat triangulations of the positions of acoustically tagged fish within the array over time. Position data were collected over a range of time periods from several days to two months. We parsed the VPS position data into 20-minute periods for each fish, eliminated intervals with fewer than six positions, and then calculated the area of the 95% kernel density (KDE) of these positions within and across tagged fish (mean: 1.93 ha, SE: 3.31). KDE were fit using the kde2d function in the MASS package in R. Bandwidths were chosen automatically based on a Gaussian reference distribution rule-of-thumb within the function and ranged from 0.25 to 722.5 in the x-axis and 0.12 to 649.3 in the y-axis. We then used the mean 95% KDE to approximate the area used by an individual Greater Amberjack during a 20-minute time interval. This area was translated into a circle with radius 78.3 m and was used to approximate the ROV-video-camera effective sampling area during a single deployment.

For each ROV survey, we then calculated the area of target-reef habitat (from side-scan mapping data, described above) within the ROV-video-camera ESA. This area was then used as the denominator in the site-specific density estimates.

Acoustic-based estimates of site-specific density on hardbottom habitats

Hydroacoustic surveys and the analysis of associated data are described in more detail in chapter 8, but in brief, surveys at artificial reef and natural reef sites consisted of 7 parallel transects of 200 m from the center of the reef, extending out to 70 m from the center point while surveys at oil and gas platforms consisted of 4 perpendicular transects on all four sides of the structure, extending 100m beyond the platform. Paired ROV video surveys were conducted at the center of these areas. This design was implemented with the intent for camera surveys to obtain species composition information and hydroacoustic surveys to obtain a total count of individual large bodied (>20cmTL), disaggregated, water column (>0.5m off seabed) fish targets in the acoustic transect area covered. Therefore, the numerator for the site-specific density estimates was the acoustic count of observed during the acoustic survey of the site, multiplied by a deflation factor determined using the camera counts of Greater amberjack and all other observed large-bodied, water column fishes. The denominator of the site-specific density estimate was the area of the circle used to clip the acoustic transects multiplied by the proportion of target reef in the covered area.

Towed camera survey of uncharacterized bottom habitat

Uncharacterized (not known hardbottom) seafloor habitat was sampled using the Camera-Based Assessment Survey System (C-BASS; Grasty 2014; Lembke et al. 2017) which consisted of a series of six high-definition video cameras used in stereo and individually. The system also has a range of environmental sensors (temperature, depth, salinity, CDOM) and navigation-related sensors (e.g., pitch, roll, yaw, altitude). The system is “flown” in

real time using imagery supplied from the tow sled to the vessel deck. Depth alterations are done via adjusting the towing cable length to affect the height over bottom. The system is generally flown at an altitude of 2-5 m above the sea floor, depending on water visibility and bottom topography, and is towed at speeds of 1.5 – 2 m s-1. All efforts were made to sample the selected transects as planned; however, unforeseen constraints to sampling (e.g., presence of oil rigs or obstacles in the transect path, unsafe proximity to high traffic areas, presence of actively fishing vessels) resulted in transects that were either cut short of the 6.5 nm target, moved to an adjacent grid cell, or towed in an alternate randomly selected compass direction.

Imagery from the C-BASS video transects were processed in 15 s intervals, although scrolling a few sections in each direction was allowed to provide context and ensure that classification adequately characterized the area. Observed habitats were classified according to the substrate and biotic components of the Coastal and Marine Ecological Classification Standard (CMECS) scheme. Characterized metrics included seafloor induration, cover and composition of primary attached biota, and primary and secondary components of hardbottom and softbottom habitat, respectively. In addition, all observed fish were enumerated and identified to the lowest possible taxon; if no taxonomic distinction was possible, observed individuals were classified as unidentified small and large fishes, depending on size. Positional data, including roll, pitch, yaw, altitude, and tow speed were analyzed to estimate area viewed by each transect. Transect-specific density was simply calculated as the number of GAJ observed divided by the area viewed.

Sampling Design and Estimators

Study design for hardbottom habitat

We are leveraging a pre-existing project in the Alabama-Mississippi (AL-MS) region. The University of South Alabama Fisheries Ecology Lab's Fishery Independent Survey (FIS), to increase the number of samples available to estimate GAJ abundance in the region. The FI Survey has a complicated sample design, as it is a long-term general reef-fish survey that has been modified to meet changing data needs over the years. Additionally, this survey likely under-samples certain uncommon habitat types that are important for Greater Amberjack, including high-relief artificial reefs and both high- and low-relief natural reefs. A second survey, referred to as the GAJ Count, was implemented in AL-MS to supplement the FIS survey and to mitigate issues of under-sampled habitat via increased sampling of these three uncommon habitat types.

The sampling universe for both Alabama hardbottom survey designs (FIS and GAJ Count) is based on a long-term side-scan sonar mapping effort which uses random sampling from a grid to characterize reef-fish habitat in the selected grid cells. This existing mapping effort allows us to estimate the numbers and sizes of artificial and natural reefs in the region (see previous section) and provides the sampling universe for fish community surveys. In summary, there are a total of 1423 2km x 2km grid cells in the habitat-mapping universe. This is broken down into six strata defined by two factors: depth (shallow = 18 to 37m, mid = 37 to 55, and deep = 55 to 110m) and whether grid cells are inside or outside the Alabama Artificial Reef Zone (AARZ). Strata A, B, and C are inside the reef permit zone and strata E, F, and G are outside the zone, with A and E in shallow, B and F in mid, and C and G in deep depth categories (Table 12.2.1). Of the 1423 total grid cells in the region, 491 cells in total have been habitat-mapped between 2010 and 2021. These cells are used to estimate the number and areas of reefs of each habitat type and serve as the pool of grid cells that could be selected for sampling under the GAJ Count survey. For FIS, the selection pool of grids is restricted to those 289 cells that were mapped in recent years (2017-2021); a previous study revealed that early mapping data (2010 – 2016) is unreliable for locating small

artificial reef structures (which FIS predominantly samples) due to degradation, movement and deployment of new structures.

Fishery Independent Survey Design

As stated, early grid mapping data (2010 – 2016) has proven unreliable in terms of predicting the numbers of and locations of one of the most common habitat types in the region (small, i.e., low-relief, artificial reefs). Therefore, the FIS sampling universe was limited to the 289 grid cells that were mapped in recent years (2017 – 2021). Of these 289 grid cells, 31 had no qualifying reefs (neither artificial nor natural) and were eliminated from the sampling frame. A further 41 grid cells (16% of 258) were eliminated because they did not contain at least one qualifying reef (either natural or artificial) that had not been sampled in previous years. This secondary elimination is not a true random sample of a random sample of grid cells; we assume that the 41 cells eliminated are not different in terms of GAJ abundance per structure. Thus, the sampling frame of cells for the FIS reef fish survey consisted of 217 grid cells that contained at least one reef (by the above definition) that had never been sampled for fish in the past. The FIS pool of 217 grid cells available for sampling are divided amongst the strata as follows: {A: 84, B: 64, C: 31, E: 5, F: 18, G: 15} (Table 2).

Of the 217 qualifying grid cells, a total of 36 grids were randomly selected for sampling (A: 12, B: 8, C: 4, E: 2, F: 6, G: 4). Selection was allocated among the strata proportional to area of a stratum, but even allocation was somewhat limited by logistical constraints and the small numbers of qualifying cells in some strata. In previous years of the FIS survey, we selected two reefs from each selected grid in the A, B and C strata (inside the permit zone) and one reef from each selected grid in the E, F, and G strata (outside the permit zone). But due to a reduced budget, the total number of reefs sampled had to be reduced. Therefore, we randomly designated a subset of A and B strata's grid cells, along with all of the E, F, and G cells, as "single-reef" cells, in which only one reef would be sampled (A: 4, B: 2, C: 0, E: 2, F: 6, G: 4). In A, B, and C stratum, "double-reef" cells, in which two reefs would be sampled, were randomly selected from those grid cells that contained at least two qualifying reefs that were at least 400m apart (A: 8, B: 6, C: 4).

The above procedures for selecting cells and reefs imply three separate sampling designs, one for strata A & B, a second for stratum C, and a third for strata E & F & G. Below are explanations of each sampling design, followed by the estimators for those strata.

Sampling Design for Strata A and B:

The sampling was a two-step process (see figure 3 in Supplement). For single cell reefs in a stratum s : if one-or-more natural reefs were present, a natural reef was randomly for sampling. If no natural reefs were present, one artificial reef was randomly selected for sampling.

First, t_s cells are randomly selected for single reef sampling from the T_s total cells available for sampling in the stratum. The T_s cells originally available for sampling is divided into M_j cells, with only one qualifying reef, and N_s cells, with two or more reefs, such that $T_s = M_s + N_s$ The number of cells in stratum s available for sampling after the first selection of single reefs is the number of cells with two or more qualifying reefs that haven't been sampled in previous years, N_s, minus the number of cells in the collection of N_s cells that were selected in the first step that have two or more reefs. This is $N_s - (t_s - x_s)$ where x_s is the number of cells from the sample of t_s cells selected in the first step that are members of the M_s cells with only one qualifying reef. x_s is a random

variable with a hypergeometric distribution, $x_s \sim Hg(M_s, N_s, t_s)$. A sample of S_s cells was selected in the second step.

The cells sampled in strata A and B were selected giving more weight to cells with two or more reefs relative to cells with just a single reef. To deal with this, we can post-stratify the results into two strata, one with cells having a single reef and the other with cells having two or more reefs. We compute abundance estimates for each stratum separately and then combine the estimates by weighting by estimates of the fraction of the cells that are in each stratum (as estimated from the mapping of cells).

Estimators for the FIS Design

Estimator for natural reefs for Strata A and B.

The cells selected only had artificial reefs so no estimates can be made for GAJ on natural reefs in strata A and B.

Estimator for artificial reefs for Strata A and B.

Mean per cell estimator. We post-stratified the data for each stratum, A and B, into two post-strata, one with cells having a single reef and the other with cells having two or more reefs. We compute abundance estimates for each post-stratum separately and then combine the estimates by weighting by estimates of the fraction of the cells that are in each stratum (as estimated from the mapping of cells).

The total abundance in a stratum (A or B) is given by

$$\hat{Y}_{tot} = N_{tot} \frac{N_{single}\underline{Y}_1 + N_{multiple}\underline{Y}_2}{N_{zero} + N_{single} + N_{multiple}}$$

where N_{tot} is the total number of cells in the stratum, N_{zero} is the number of mapped cells with no reefs, N_{single} is the number of mapped cells with just one reef, $N_{multiple}$ is the number of mapped cells with multiple reefs (so $N_{zero} + N_{single} + N_{multiple} = N_{mapped}$ (the number of cells that were mapped), $\underline{Y}_1$ and $\underline{Y}_2$ are the mean estimated cell total abundances for sample cells that have one and multiple reefs, respectively (i.e., the mean total abundances/cell for the two post-strata). It is assumed that cells with no reefs (of any kind) have no GAJ.

The total abundance in a cell with one reef is simply the MaxN count. For cells with two or more reefs present, the total abundance in a cell is the mean MaxN count for the cell times the number of reefs in the cell.

The variance of $\hat{Y}_{tot}$ can be estimated using the delta (Taylor's series) method. Let $\hat{P}_1 = \frac{N_{single}}{N_{mapped}}$ and $\hat{P}_2 = \frac{N_{multiple}}{N_{mapped}}$. Then $\hat{Y}_{tot}$ can be written as

$\hat{Y}_{tot} = N_{tot}\hat{P}_1\underline{Y}_1 + N_{tot}\hat{P}_2\underline{Y}_2$.

There are four variables: $\underline{Y}_1$, $\underline{Y}_2$, $\hat{P}_1$, and $\hat{P}_2$. The proportions $\hat{P}_1$and $\hat{P}_2$have a trinomial distribution; hence the estimated variances and covariance are

$$\hat{Var}(\hat{P}_1) = N_{mapped}(\hat{P}_1)(1 - \hat{P}_1)$$

$$\hat{Var}(\hat{P}_2) = N_{mapped}(\hat{P}_2)(1 - \hat{P}_2)$$

$$\hat{Cov}(\hat{P}_1, \hat{P}_2) = -N_{mapped}\hat{P}_1\hat{P}_2$$

(Recall that N_{mapped} is the total count of cells mapped, including cells found to have no reefs.) The covariance of every other pair of variables is zero because the parameters were estimated from independent surveys.

The variance of $\hat{Y}_{tot}$ can be estimated by:

$$\hat{Var}(\hat{Y}_{tot}) = (N_{tot}\hat{P}_1)^2\hat{Var}(\underline{Y}_1) + (N_{tot}\underline{Y}_1)^2\hat{Var}(\hat{P}_1) + (N_{tot}\hat{P}_2)^2\hat{Var}(\underline{Y}_2) + (N_{tot}\underline{Y}_2)^2\hat{Var}(\hat{P}_2) + 2(N_{tot})^2\underline{Y}_1\underline{Y}_2\hat{Cov}(\hat{P}_1, \hat{P}_2)$$

Ratio estimator: An alternative to the mean-per-cell estimator is a ratio estimator applied to each post stratum. For post stratum 1, which comprises the cells with a single reef, the ratio estimate of the stratum total is

$$\hat{Y}_{str1-R} = N_{tot}\hat{P}_1\frac{\sum_{i=1}^{n_1} c_i}{n_1}$$

where $\hat{Y}_{str1-R}$ is the ratio estimator of abundance in post stratum 1 and n_1 is the number of cells sampled in the stratum. Note that for this stratum, the ratio estimate is the same as the mean per cell estimate because $M_i = 1$ for all i, and $\sum M_i = n_i$ since each cell has exactly 1 reef.

For post stratum 2, which comprises cells with 2 or more reefs, the ratio estimate of the stratum total is

$$\hat{Y}_{str2-R} = \frac{N_{tot}}{N_{mapped}}\left(\sum_{i=1}^{N_{multiple}} M_i\right)\frac{\sum_{i=1}^{n_2} M_i\bar{y}_i}{\sum_{i=1}^{n_2} M_i}$$

where the sum of the M_i values is over the $N_{multiple}$ mapped cells for which $M_i \geq 2$, n_2 is the number of cells sampled in the post stratum 2, and $\bar{y}_i$ is the mean abundance over the reefs sampled in cell i. (The sum of M_i is the number of reefs in those cells of the N_{mapped} cells that have two or more reefs, the fraction N_{tot}/N_{mapped} expands the count of reefs to the total number of cells. Therefore, the product of N_{tot}/N_{mapped} and the sum of the $\sum M_i$ gives an estimate of M_0, the total number of reefs in cells that contain two or more reefs, and the quotient at the far right is the ratio of abundance per reef. An estimate of the total abundance in the N_{tot} cells is given by $\hat{Y}_{str1-R}$ + $\hat{Y}_{str2-R}$. An estimate of the variance of the estimated total abundance is obtained by adding the variances of the two stratum estimates.

Sampling Design for Stratum C:

Four cells were randomly selected and two reefs were selected per selected cell. Priority was given to sampling natural reefs so up to 2 natural reefs were selected (depending on how many natural reefs occurred in the cell). After selecting natural reefs, selection was made of up to 2 artificial reefs so that a total of 2 reefs (of any type) were selected if possible. This resulted in a sample of 3 cells where two natural reefs were selected and one cell where 1 natural and 1 artificial reef was selected.

Estimator for Natural reefs in stratum C.

Four cells were sampled and in three of them, sampling took place at two natural reefs. In the fourth cell, one natural reef and one artificial reef were sampled. Since priority was given to sampling natural reefs, the presence or absence of artificial reefs had no influence on the selection of natural reefs.

Of the cells available for survey, let n^* = number of cells selected for sampling and n = the number of cells selected that have one or more natural reefs to be sampled ($n \leq n^*$). Within the n cells that contain natural reefs, m_i denotes the number of natural reefs to be sampled in cell i ($m_i = 0, 1$, or 2 depending on how many reefs are in the cells). The abundance of amberjack on natural reef j, in cell i is:

$y_{ij} = c_{ij}A_{ij}/a_{ij}$

where c_{ij} is the MaxN of amberjack, A_{ij} is the total area of the jth reef in cell i, and A_{ij} is the area viewed by the camera on reef j, in cell i.

The mean abundance of amberjack on natural reefs in cell i can be calculated as:

$$\overline{y_i} = \frac{\sum_{j=1}^{m_i} y_{ij}}{m_i}$$

Mean per cell estimator: The mean per cell estimator (MPC) for the total abundance of greater amberjack in the stratum can calculated as:

$$\hat{\hat{C}}_{MPC} = \left(N_{tot}\frac{N^{**}}{N^*}\right)\frac{\sum_{i=1}^{n} \mathrm{M}_i\overline{y_i}}{n}$$

where N_{tot} is the total number of cells in the stratum, N^* is the number of mapped cells in the stratum, N^{**} is the number of cells mapped and found to have natural reefs, M_i is the total number of natural reefs in cell i, and n is the number of cells sampled with natural reef.

Ratio estimator: Alternatively, a ratio estimator for the stratum can be defined as:

$$\hat{\hat{C}}_{Ratio} = \left(N_{tot}\frac{{M_0}^{**}}{N^*}\right)\frac{\sum_{i=1}^{n} \mathrm{M}_i\overline{y_i}}{\sum_{i=1}^{n} \mathrm{M}_i}$$

Where quantities are defined as for the mean per cell estimator, and ${M_0}^{**}$ is the total number of reefs in the N^* mapped cells in the stratum.

The variance of both the mean per cell and ratio estimators can be derived using can be found using Goodman's (1960) exact formula. For random variables U and V, the exact estimated variance is

$\hat{Var}(UV) = U^2\hat{Var}(V) + V^2\hat{Var}(U) - \hat{Var}(U)\hat{Var}(V)$.

Estimator for artificial reefs in stratum C.

The data for artificial reefs in stratum C consists of observed abundance at one reef in one cell. That count was 0 so the estimate of abundance on artificial reefs in stratum C is zero. The variance cannot be estimated since there is only one observation.

Sampling Design for Strata E, F, and G:

A simple random sample of cells is selected from those cells that have been mapped and found to have qualifying reefs not all of which were sampled previously. This is not a random sample of mapped cells with qualifying reefs because cells excluded from consideration are not random but, rather, reflect the specific

attribute that there are few (i.e., one) reef present. Nonetheless, we make the assumption that cells that were excluded (because of previous complete sampling) are not different from cells with multiple reefs. Additionally, in strata E, F and G, only one reef was randomly selected per cell and the selected reef was always a natural reef if natural reef habitat was present. Thus, for these strata, the design for natural reefs is a two-stage cluster sample with one observation per cluster. It is not possible to estimate the within and between cluster components of variance, but a conservative estimate of the variance of the estimated abundance can be obtained (Cochran 1977). For these strata, an artificial reef in a cell was sampled only if there was no natural reef present. Thus, there is no guarantee that any artificial reefs sampled are representative of the population of artificial reefs in the stratum.

The sampling design for artificial reefs in strata E, F and G differs from that for natural reefs in an important respect. The sample of cells for sampling natural reefs was drawn from the n^* cells selected for the sampling; a natural reef was sampled in all of those cells that had natural reefs. In contrast, for artificial reefs, only those cells among the n^* selected that did not have any natural reef habitat had an artificial reef sampled (if present). Thus, n cells were selected for sampling natural reefs and n_a cells were sampled for artificial reefs, with n_a being $\leq n^* - n$ (it would be equal to $n^* - n$ if all cells not selected for sampling a natural reef had an artificial reef present). Although n^* and n_a are random variables, we treat them as known constants. The cells sampled for artificial reefs are not a random sample but we treat them as such on the assumption that artificial reefs in cells excluded (because they have a natural reef) are not different from cells that have no natural reef habitat.

Estimator for natural and artificial reefs in strata E, F, and G.

We'll use the following notation (see diagram below). Let:

N_{tot} = number of grid cells in the stratum

N^* = number of cells in the stratum that have been mapped ($N^* < N_{tot}$)

N = number of cells that have been mapped and found to have reefs of any kind ($N \leq N^*$)

$\widetilde{N}$ = number of cells that have been mapped in recent years, found to have reefs (of any kind), and not all the reefs have been previously sampled ($\widetilde{N} \leq N$)

$N^{**}{}_{hab}$= number of cells that have been mapped and found to have reef habitat ($N^{**} \leq N^*$), for a given habitat type (artificial or natural)

n^* = number of cells selected for surveying from the $\widetilde{N}$ cells available for sampling

n_{hab} = number of cells selected for surveying that have one or more natural reefs ($n \leq n^*$)

The total abundance in cell i for each of the cells selected is estimated by:

$$\hat{C}_i = \frac{c_i}{a_i} A_i$$

where c_i is the MaxN count, a_i is the area viewed, and A_i is the total area of reef in cell i. Note c_i = 0 if there is no reef habitat present in the cell.

Mean per cell estimator: The average abundance per cell is estimated as

$$\hat{\bar{C}} = \frac{1}{n}\sum_{i=1}^{n} \hat{C}_i$$

where the sum is over the n cells with reef habitat not sampled previously.

Using the mean per unit estimator, the total abundance in the stratum, $\hat{\hat{C}}$, is the product of the total number of cells times the fraction of cells with reef habitat times the average abundance in cells with reef habitat (estimated from n cells with reef habitat not previously sampled). Thus,

$$\hat{\hat{C}} = N_{tot} \frac{{N^{**}}_{hab}}{N^*} \hat{\bar{C}}$$

$\hat{\hat{C}}$ is the product of a constant, N_{tot}, a binomial random variable, ${N^{**}}_{hab}/N^*$, and an estimate from two-stage cluster sampling with one observation per cluster.

The estimated variance of the binomial random variable is

$$\hat{Var}\left(\frac{{N^{**}}_{hab}}{N^*}\right) = = \frac{{N^{**}}_{hab}}{N^*}\left(1 - \frac{{N^{**}}_{hab}}{N^*}\right)/N^*$$

The variance of $\hat{\hat{C}}$ is somewhat problematic because the first stage sampling of cells is not a random sample of all mapped cells because cells where all reefs were previously sampled are excluded. However, the percentage of cells that were excluded is small and thus the sample is similar to a random sample of all mapped cells and is treated as such. When there is one observation per cell, as is the case here, the variance can be approximated as

$$\hat{Var}\left(\hat{\bar{C}}\right) = \frac{\sum_{j=1}^{n}(y_i - \bar{y})^2}{n_{hab}\,(n_{hab} - 1)}$$

where $y_i = c_i A_i / a_i$ and $\bar{y}$ is the mean of the n_{hab} observations (for a given habitat) of y.

The variance of a product of two independent random variables can be found using Goodman's (1960) exact formula. For random variables U and V, the exact estimated variance is

$$\hat{Var}(UV) = U^2\hat{Var}(V) + V^2\hat{Var}(U) - \hat{Var}(U)\hat{Var}(V)$$

Substituting equations, the relevant values yields:

$$\hat{Var}\left(N_{tot}\frac{{N^{**}}_{hab}}{N^*}\hat{\bar{C}}\right) = N_{tot}^2\left\{\left\{\frac{{N^{**}}_{hab}}{N^*}\right\}^2 \hat{Var}\left(\hat{\bar{C}}\right) + \hat{\bar{C}}^2\,\hat{Var}\left(\frac{{N^{**}}_{hab}}{N^*}\right) - \hat{Var}\left(\frac{{N^{**}}_{hab}}{N^*}\right)\hat{Var}\left(\hat{\bar{C}}\right)\right\}.$$

Ratio Estimator: An alternative estimator would be a ratio estimator given by

$$\hat{Y}_R = \frac{M_o^{**}}{N^*} N_{tot} \frac{\sum_{i=1}^{n} M_i \frac{c_{ij}A_i}{a_{ij}}}{\sum_{i=1}^{n} M_i}$$

where M_o^{**} is the number of reefs in the N^* cells that were mapped, N^* is the number of cells that were mapped (over all years), N_{tot} is the number of grid cells in the stratum, n is the number of cells sampled, M_i is the number of reefs in cell i, c_{ij} is the count at the reef j sampled in cell i, a_{ij} is the area observed at the reef j in cell i, and A_i is the total area of the reef observed in cell i. The variance for this estimator can be derived in the same manner as the mean per cell estimator.

Supplemental uncharacterized habitat sampling:

As a supplement to the C-BASS towed camera sampling, which is limited to sampling outside of areas with high densities of artificial and natural reefs, we incorporated uncharacterized bottom, or "no-known-structure" (NS), sites into our regular reef survey. For strata inside the reef zone (A, B, and C) we randomly selected two of the grid cells selected for reef surveys in each stratum for sampling unconsolidated habitat. For strata outside the AARPZ zone (E, F, and G), we sampled unconsolidated habitat in all of the grid cells selected for reef surveys. In both cases a single random location, at least 400m from any known vertical relief, was selected in each cell for unconsolidated bottom survey. This means that our estimate of Greater Amberjack abundance from these surveys is conditional on the grid selection criteria used for the FIS reef fish survey, namely grids had to have at least one qualifying reef that had not been previously sampled in order to be a candidate for selection. We make the simplifying assumption that this is an unbiased estimate of GAJ abundance across the entire gridded region. This constitutes a two-stage cluster sampling plan with one or two observations per cluster (grid cell).

Greater Amberjack Survey Design

Since the GAJ Count survey was focused on providing an estimate of GAJ abundance on high-relief artificial reefs and natural reefs, and because these reefs are much less likely to move, degrade, or become buried over time, we were more confident relying on full list of 491 grid cells that had been mapped between 2010 and 2021. We then filtered out cells that had no qualifying structures (at least 0.5m height and at least 4 m^2 in area) and cells that did not contain at least one of target habitat (high-relief artificial, low-relief natural, or high-relief natural). This left us with a total of 94 qualifying grid cells in the following strata: {A: 24, B: 16, C: 40, E: 0, F: 3, and G:11} (Table 3). But, it is important to note that for any sampled low relief artificial and unconsolidated bottom habitats, we can only derive estimates that are conditional on their having at least one reef present that is high relief artificial or high-or-low relief natural. This is because any cell that did not have one of these three reef types present were excluded from sample selection.

We then used stratified-random sampling to select 43 of these cells for surveys (A: 18, B: 12, C: 6, E: 0, F: 3, G: 4). We did our best to allocate effort proportionally by area, but there were not enough qualifying grid cells in some strata to make this possible across all strata.

For each of the selected grid cells, we randomly selected one reef in each of four habitat categories (if present) in the grid cell (low-relief artificial, high-relief artificial, low-relief natural, and high-relief natural). Selected reefs within a grid had to be at least 400 m from one-another. This could theoretically result in as many as four reefs selected per grid (if all four habitats were present) and as few as one (if only one present), but most grid cells had two categories of reef present. In practice, when a cell lacked some of the priority habitat types, the field crew tried to meet the target of four sites sampled in the cell by sampling at multiple low relief artificial reefs.

Note for natural reefs, the sampling scheme is three-stage: pick cells at the first stage, reefs within cells at the second stage, and patch within a reef the third stage. Because one reef and one patch within a reef is selected, we don't have replicate observations at the second and third stages. We can still get unbiased estimates of the total but we can't estimate the components of variance. A conservative (biased high) estimate of the variance can be obtained from the variance at the first stage.

Note: there was overlap between the GAJ grid-cell pool and the pool used for FIS grid cell selections (50 cells were common between the two pools of potential selections). Essentially, at the grid-cell selection phase, both FIS and GAJ procedures were run independently, with their own set of criteria, on two overlapping groups of

grid cells. Two grid cells ended up being selected for both projects, but reef selection was treated differently for these for the GAJ project. Where the GAJ grid-cell selection procedure selected a cell that was already selected for FIS (two grid cells), we randomly selected one additional reef (if present) in each of the four categories if that category had not already been selected in the FIS site selection process.

Active oil and gas platforms: this category of artificial reef was treated separately from other artificial reefs. There are 18 standing, active platforms in the Alabama sub-region. We randomly selected three platforms in Alabama (two in F-stratum and 1 in G-stratum) for GAJ abundance sampling. Note that sampled platforms were randomly selected from a list of platforms and did not use the grid system to select sample locations.

No-structure sites: For strata outside the AARPZ (E, F, and G), we also selected one location per selected grid cell (i.e., one location in each of the 3 cells selected in stratum F and one in each of the 4 cells selected in stratum G) to conduct a survey of unconsolidated/unknown bottom. For strata inside the reef zone (A, B, and C) we randomly selected one of the previously selected grid cells in each stratum for sampling unconsolidated/unknown habitat, i.e., one location in stratum A, one in stratum B, and one in stratum C). In either case, the specific location sampled was a random point at least 400m from any known vertical relief.

Estimators for GAJ sampling:

Separate, independent estimates are made for each of the habitat types. The following notation is used for a particular habitat type. Quantities in parentheses were used in the description of the design but are not used in the estimation of abundance.

N_{tot} = number of cells in the stratum

N_{mapped}= number of cells mapped (over all years)

N_{pool} = number of mapped cells that have at least one of the 3 habitats of interest

N_{hab} = number of mapped cells having the habitat under consideration ($N_{hab} \leq N_{pool}$)

N_{samp} = number of cells selected for sampling, not all of which will have the habitat of interest

$N_{samphab}$ = number of cells selected for sampling the habitat under consideration. $N_{samphab} \leq N_{samp}$. Although $N_{samphab}$ is a random variable it is treated as fixed.

Then, an estimator of the abundance in the stratum, using the mean per cell, is as follows:

$$\hat{Y} = \frac{\sum_{i=1}^{N_{samphab}} y_i}{N_{samphab}}$$

where y_i is abundance of greater amberjack on the habitat under consideration in cell i, and $\hat{Y}$ represents the mean abundance in the sampled cells that contain a given habitat.

Then, $N_{hab}\ \hat{Y}$ is an estimate of the abundance on the habitat in the N_{hab}cells with the habitat, which is also an estimate of the abundance on the habitat in the N_{mapped} cells that have been mapped. Consequently, an estimate of abundance on the habitat in all of the cells in the stratum can be found as:

$\hat{Y}_{strat} = \frac{N_{tot}}{N_{mapped}} N_{hab}\hat{Y} = N_{tot}\frac{N_{hab}}{N_{mapped}}\hat{Y}$.

The estimated variance of this estimator is:

$$\hat{V}\left(\hat{Y}_{stratum}\right) = (N_{tot})^2 \hat{V}\left(\frac{N_{hab}}{N_{mapped}}\hat{Y}\right).$$

The quantity $\hat{P} = \frac{N_{hab}}{N_{map}}$ is a binomial random variable with estimated variance $\hat{P}(1-\hat{P})/N_{mapped}$. The variance of the quantity $\hat{Y}$ is estimated from the first stage variance (of a multi-stage sampling design), as before. The variance of the product of $\frac{N_{hab}}{N_{map}}$ and $\hat{Y}$ is estimated using Goodman's exact variance of products formula.

An alternative, ratio estimator can be developed as follows. Let $N_{feature}$ be the number of occurrences (reefs) of the habitat being considered in the N_{mapped} cells, and $\hat{Y}$ be the mean abundance per feature in the n cells sampled. Then, the total abundance in the stratum can be estimated by

$$\hat{Y}_{stratum-ratio} = N_{tot}\frac{N_{feature}}{N_{mapped}}\hat{Y}\ .$$

Then, let M_i be the number of occurrences of the habitat in cell i, c_{ij} be the MaxN count at the one feature (the jth reef) sampled in cell i, a_{ij} be the area viewed to get the MaxN count, and A_{ij} be the area of the reef j. The ratio estimator for number of features can be defined as:

$$\hat{Y}_{stratum-ratio} = N_{tot}\frac{N_{feature}}{N_{mapped}}\hat{Y} = N_{tot}\frac{N_{feature}}{N_{mapped}}\frac{\sum_{i=1}^{n} M_i \frac{c_{ij}A_{ij}}{a_{ij}}}{\sum_{i=1}^{n} M_i}\ .$$

The ratio estimator can also be formulated in terms of area of natural reef habitat instead of number of natural reef features. Then,

$$\hat{Y}_{strat-ratio} = N_{tot}\frac{A_{feature}}{N_{mapped}}\hat{Y} = N_{tot}\frac{A_{feature}}{N_{mapped}}\frac{\sum_{i=1}^{n} A_i \frac{c_{ij}}{a_{ij}}}{\sum_{i=1}^{n} A_i}$$

where $A_{feature}$is the total area of natural reef in the mapped cells, and A_i is the area of natural reef in i^{th} cell. This equation applies to the special case where there is one observation per cell. If multiple sites with a particular habitat are sampled, then an additional summation is needed to combine the observations from the multiple sites (see equation 11, Chapter 11).

Of the three standing platforms selected for sampling in Alabama, only one was successfully sampled, and no Greater Amberjack were observed. No formal estimators are defined for this reef category in the region. However, we provide an ad-hoc approach to estimation of GAJ abundance by applying mean reef-level density estimates from adjacent Louisiana depth strata to platforms in Alabama.

Additionally, we apply a similar approach to estimating the abundance of Amberjack on the portion of the shelf-edge natural reef complex that is not covered by the FIS mapping survey and was not included in our sample design. In this case we apply the reef-level areal density estimated in adjacent mapped and sampled strata to our estimate of natural reef are in the un-mapped portion of the shelf-edge region.

Mississippi survey design

In the case of Mississippi, there was no leverageable project and no available comprehensive habitat mapping data. Therefore, we designed the sampling around the existing lists of artificial reefs available from various

databases (MDMR, LDFW, NOAA, BOEM). Previous mapping efforts by USGS also indicated a substantial amount of natural reef along the shelf-edge, but did not provide high-enough resolution to either systematically sample these reef areas or estimate their extent.

We began by extending the existing AL grid system to include the area between the existing grid and the Mississippi river. This new Mississippi grid contains 1201 cells. We then plotted all known MS and LA State artificial reefs, all high- and medium-confidence AWOIS wrecks, and all active BOEM oil and gas platforms, onto this grid. We reduced the pool of grid cells to those that contained at least one low-relief or high-relief artificial reef. This resulted in a pool of 41 cells, which were categorized into three depth strata (E, F, and G: corresponding to the three depth strata used for the AL sample design). We have included a spreadsheet detailing the number of cells in each depth stratum and the number of reefs of each relief category (high and low) in each of these cells. We then randomly selected 17 of these grid cells (E: 7 of 18 available, F: 4 of 6 available, and G: 6 of 17 available). Sampling effort was not proportional to area due to logistical constraints. The distance of the G stratum from our home port necessitated expensive overnight trips, so this stratum was undersampled compared to the total area of the stratum and the number of cells containing artificial reefs. Conversely, the F stratum, which contained only 6 cells with artificial reefs, was undersampled compared to the total area of the stratum, but oversampled compared to the number of cells containing reefs. For each of the selected grid cells, we randomly selected one reef in each of the two relief categories (high and low) if present for sampling. However, many grid cells did not contain reefs in both relief categories. Most of the shallow (E stratum) cells only contained low-relief reefs and most of the deep (G stratum) cells only contained high-relief reefs.

Active oil and gas platforms were treated separately from other artificial reefs. There is a total of 83 active platforms in the Mississippi region. Therefore, we randomly selected three platforms (E: 1 of 22, F: 1 of 17, G: 1 of 45) for GAJ abundance sampling. This design undersamples Mississippi platforms, but again, we were up against logistical constraints and did the best that we could to allocate effort efficiently across the regions. Note that sampled platforms were randomly selected from a list of platforms and did not use the grid system to select sample locations.

Estimators for MS sampling

Of the cells selected for sampling for artificial reefs E: 7 of 18 available, F: 4 of 6 available, and G: 6 of 17 available), only 2 cells were sampled in each E and F strata (4 total) due to logistical limitations and poor visibility conditions. In each cell, one artificial, low relief reef was sampled; no amberjack were observed on the sampled reefs. Due to the high rate of non-responses and no amberjack observed on sampled reefs, no formal estimators are defined for this region.

Of the three Mississippi platforms selected for sampling, poor visibility resulted in unusable ROV video footage. Therefore, we do not have a formal sample-based estimate of Amberjack abundance on platforms in the Mississippi region.

Due to insufficient sampling in the Mississippi sub-region, we provide an ad-hoc approach to estimation of GAJ abundance. We applied mean reef-level density (fish per reef) estimates from adjacent Alabama strata to Mississippi strata, including low-relief artificial reefs and high-relief artificial reefs in each depth category (shallow, mid, and deep strata). We also apply mean platform-level density (fish per structure) estimates from adjacent Louisiana depth strata to Mississippi strata. These estimates are not based on sample design theory, but we believe that they represent the best available estimates given the limited data.

Results

Size distribution of Greater Amberjack from video survey

We estimated fork length and total length for 97 individual Greater Amberjack observed during ROV video surveys, using a parallel laser system. Total length measured by the ROV camera gear during the Alabama portions of the survey ranged from 231 to 977mm, with a mean length of 537.4 (SD: 166.5). The length distribution was bimodal with peaks at 400 and 600mm (Figure 12.2.8). The proportion of fish in five size bins across each of three depth zones is presented in table 12.2.7.

Abundance estimates by stratum are presented in Tables 12.2.8 and 12.2.9.

TABLES

Table 12.2.1: Distribution of habitat-mapping grid cells among the six strata in the Alabama region. The sampling universe is categorized into six strata based on three depth categories (Shallow = 18.3 - 36.6m, Mid-depth = 36.6 - 54.9, and Deep = 54.9 – 109.7m) and whether the cells are inside or outside the Alabama Artificial Reef Permit Zone (AARPZ). A randomly selected subset of these cells was mapped using side-scan sonar between 2010 and 2021.

Stratum name	Depth category	AARPZ	Total number of cells	Number of cells mapped (2010-2021)
A	Shallow	Inside	398	173
B	Mid-depth	Inside	297	119
C	Deep	Inside	111	61
E	Shallow	Outside	131	28
F	Mid-depth	Outside	338	77
G	Deep	Outside	148	33
Total			*1423*	*491*

Table 12.2.2 Mean and standard deviation (SD) of artificial and natural reef counts per mapped cell for each stratum from randomized side-scan mapping in the Alabama region. Note: the number of mapped cells in each stratum are given above in Table 12.2.1.

				Low-relief		High-relief	
Habitat Type	Stratum name	*Mean count*	*SD*	*Mean count*	*SD*	*Mean count*	*SD*
Artificial	A	7.05	4.51	6.82	4.39	0.23	0.76
	B	5.87	3.63	5.72	3.61	0.15	0.62
	C	2.44	2.85	2.43	2.86	0.02	0.13
	E	1.68	2.94	1.68	2.94	0	0
	F	1.86	1.81	1.83	1.82	0.03	0.16
	G	2.91	3.63	2.76	3.44	0.15	0.87
Natural	A	0.08	0.47	0.08	0.47	0	0
	B	0.10	0.45	0.11	0.45	0	0
	C	85.70	206.64	79.00	196.98	6.67	19.03
	E	0	0	0	0	0	0
	F	0.01	0.11	0.01	0.11	0	0
	G	12.10	34.93	10.90	33.76	1.21	4.17

Table 12.2.3: Mean and standard deviation (SD) of artificial and natural reef area (m^2) per mapped cell for each stratum from randomized side-scan mapping effort in the Alabama region. Note: the number of mapped cells in each stratum are given above in Table 12.2.1.

				Low-relief		High-relief	
Habitat Type	Stratum name	*Mean area (m^2)*	*SD*	*Mean area (m^2)*	*SD*	*Mean area (m^2)*	*SD*
Artificial	A	146.00	310.81	102.00	204.45	44.40	227.82
	B	67.10	68.70	64.50	68.63	2.59	10.91
	C	73.00	211.90	57.30	177.48	15.80	123.29
	E	17.20	28.00	17.20	28.00	0	0
	F	51.80	133.42	37.90	76.75	13.90	90.83
	G	3270.00	5458.94	2920.00	5244.04	355.00	2037.15
Natural	A	57.30	509.90	57.30	509.90	0	0
	B	20.90	108.17	20.90	108.17	0	0
	C	95200.00	224276.61	55200.00	174928.56	40000.00	130766.97
	E	0	0	0	0	0	0
	F	0.66	5.80	0.66	5.80	0	0
	G	3120.00	8485.28	2250.00	7341.66	871.00	4159.33

Table 12.2.4 Estimated count and standard deviation (SD) of qualifying (>0.5 m high, >4m^2 area) artificial and natural reefs for each stratum and totaled across strata in the Alabama region. Note: the number of mapped cells in each stratum are given above in Table 12.2.1.

				Low-relief		High-relief	
Habitat Type	Stratum name	*Estimated count*	*SD*	*Estimated count*	*SD*	*Estimated count*	*SD*
Artificial	A	2805.90	89.89	2714.36	87.64	91.94	15.11
	B	1743.39	62.61	1698.84	62.14	44.85	10.68
	C	270.84	30.02	269.73	30.13	1.82	1.35
	E	220.08	33.70	220.08	33.70	0	0
	F	628.68	33.19	618.54	33.40	8.79	2.94
	G	430.68	44.20	408.48	41.79	22.50	10.59
	Total	***6099.57***	***130.74***	***5930.03***	***128.25***	***169.89***	***21.57***
Natural	A	29.89	9.38	29.89	9.38	0	0
	B	32.37	7.71	32.37	7.71	0	0
	C	9512.70	2177.09	8769.00	2075.28	740.37	200.45
	E	0	0	0	0	0	0
	F	4.39	2.10	4.39	2.10	0.00	0
	G	1790.80	424.92	1613.20	410.76	179.08	50.75
	Total	***11370.16***	***2218.20***	***10448.86***	***2115.58***	***919.45***	***206.78***

Table 12.2.5: Estimated area (ha) and standard deviation (SD) of artificial and natural reef for each stratum and totaled across strata in the Alabama region. Note: the number of mapped cells in each stratum are given above in Table 12.2.1.

				Low-relief		High-relief	
Habitat Type	Stratum name	*Est. area (ha)*	*SD*	*Est. area (ha)*	*SD*	*Est. area (m^2)*	*SD*
Artificial	A	5.81	62.01	4.06	40.79	1.77	45.45
	B	1.99	11.84	1.92	11.83	0.08	1.88
	C	0.81	22.32	0.64	18.70	0.18	12.99
	E	0.23	3.20	0.23	3.20	0	0
	F	1.75	24.53	1.28	14.11	0.47	16.70
	G	48.40	664.11	43.22	637.97	5.25	247.83
	Total	***58.99***	***667.93***	***51.33***	***639.81***	***7.74***	***252.86***
Natural	A	2.28	101.73	2.28	101.73	0	0
	B	0.62	18.64	0.62	18.64	0	0
	C	1056.72	23629.01	612.72	18429.87	444.00	13777.16
	E	0	0	0	0	0	0
	F	0.02	1.07	0.02	1.07	0	0
	G	46.18	1032.28	33.30	893.15	12.89	506.00
	Total	***1105.82***	***23651.77***	***648.94***	***18451.79***	***456.89***	***13786.44***

Table 12.2.6 Summary of artificial and natural habitats in the Mississippi sub-region.

Stratum name	Depth	Artificial reefs (count)	Oil and Gas Platforms (count)	High-relief Natural reefs (area, ha)	Low-relief Natural reefs (area, ha)
E	Shallow	107	17	0	0
F	Mid	14	45	0	0
G	Deep	6	21	3.00	7.76

Table 12.2.7. Estimated proportions of each size-class (mm TL) of GAJ within each depth zone in Alabama. Estimates derived from the total length of a subset of individuals observed (measured using parallel laser system) during the ROV video survey.

Depth zone	**0-200**	**200-400**	**400-600**	**600-800**	**800-1000**	**1000 +**
Shallow	0.257	0.641	0.067	0.032	0	0.002
Mid-depth	0.152	0.645	0.118	0.067	0.018	0
Deep	0.421	0.289	0.289	0	0	0

Table 12.2.8. Stratum level camera-based estimates of Greater Amberjack on artificial and natural reefs derived from MaxN counts across the Alabama sub-region in the north-central Gulf of America. Habitat types include high-relief artificial reefs (ART_HIGH) and low-relief artificial reefs (ART_LOW). Spatial zones include shallow (A), mid (B), and deep (C) zones inside the Alabama Reef Permit Zone and shallow (E), mid (F), and deep (G) zones outside the Reef Permit Zone. Alabama estimates are based on sampling-design approach.

PROJECT	STRATUM_ALL	STRATUM	HAB_RLF	CELLS SAMPLED (ROV)	REEFS SAMPLED (ROV)	ESTIMATE_CAMERA	SE_CAMERA	SCALED_CV_%
GAJ COUNT	A_ART_HIGH	A	ART_HIGH	10	10	905.7	537.7	59.3684443
GAJ COUNT	A_ART_LOW	A	ART_LOW	11	11	141837.3	120503.1	84.95868153
GAJ COUNT	A_NAT_HIGH	A	NAT_HIGH	0	0	0	0	0
GAJ COUNT	A_NAT_LOW	A	NAT_LOW	3	3	159.9	53.7	33.58348968
STRATUM TOTAL				24	24	142,902.90	120504.2622	84.32597391
GAJ COUNT	B_ART_HIGH	B	ART_HIGH	5	5	580.7	289.4	49.83640434
GAJ COUNT	B_ART_LOW	B	ART_LOW	12	19	1379.1	521.4	37.80726561
GAJ COUNT	B_NAT_HIGH	B	NAT_HIGH	0	0	0	0	0
GAJ COUNT	B_NAT_LOW	B	NAT_LOW	2	2	10.2	12.9	126.4705882
STRATUM TOTAL				19	26	1,970.00	596.5031433	30.27934738
GAJ COUNT	C_ART_HIGH	C	ART_HIGH	0	0	NA	NA	NA
GAJ COUNT	C_ART_LOW	C	ART_LOW	2	4	500.5	603.7	120.6193806

GAJ COUNT	C_NAT_HIGH	C	NAT_HIGH	4	4	698695.4	1007858.2	144.248581
GAJ COUNT	C_NAT_LOW	C	NAT_LOW	3	3	146901.5	67144.5	45.70715752
STRATUM TOTAL				9	11	846,097.40	1010092.482	119.38253
GAJ COUNT	E_ART_HIGH	E	ART_HIGH	0	0	0	0	0
FI SURVEY	E_ART_LOW	E	ART_LOW	1	1	46738.9	0	0
GAJ COUNT	E_NAT_HIGH	E	NAT_HIGH	0	0	0	0	0
GAJ COUNT	E_NAT_LOW	E	NAT_LOW	0	0	0	0	0
STRATUM TOTAL				1	1	46,738.90	0	0
GAJ COUNT	F_ART_HIGH	F	ART_HIGH	1	1	0	0	0
FI SURVEY	F_ART_LOW	F	ART_LOW	6	6	9969.1	6854.7	68.75946675
GAJ COUNT	F_NAT_HIGH	F	NAT_HIGH	0	0	0	0	0
GAJ COUNT	F_NAT_LOW	F	NAT_LOW	0	0	0	0	0
STRATUM TOTAL				7	7	9,969.10	6854.702298	68.76
GAJ COUNT	G_ART_HIGH	G	ART_HIGH	0	0	NA	NA	NA
FI SURVEY	G_ART_LOW	G	ART_LOW	1	1	0	0	0
GAJ COUNT	G_NAT_HIGH	G	NAT_HIGH	0	0	NA	NA	NA

GAJ COUNT	G_NAT_LOW	G	NAT_LOW	1	1		24412	0	0
STRATUM TOTAL				2	2	24,412.00		0	0
GRAND TOTAL				**62**	**71**	**1,072,090.3**	**1,017,278.4**	**94.9**	

Table 12.2.9. Stratum level acoustic-based estimates of Greater Amberjack on artificial and natural reefs derived from MaxN counts across the Alabama sub-region in the north-central Gulf of America. Habitat types include high-relief artificial reefs (ART_HIGH) and low-relief artificial reefs (ART_LOW). Spatial zones include shallow (A), mid (B), and deep (C) zones inside the Alabama Reef Permit Zone and shallow (E), mid (F), and deep (G) zones outside the Reef Permit Zone. Alabama estimates are based on sampling-design approach.

PROJECT	STRATUM_ALL	STRATUM	HAB_RLF	CELLS SAMPLED (ACOUSTIC)	REEFS SAMPLED (ACOUSTIC)	ESTIMATE_EK Method 1	SE_EK Method 1	SCALED_CV_%
GAJ COUNT	A_ART_HIGH	A	ART_HIGH	9	9	409.2	217.4	53.12805474
GAJ COUNT	A_ART_LOW	A	ART_LOW	11	11	6233	4337.7	69.59249158
GAJ COUNT	A_NAT_HIGH	A	NAT_HIGH	0	0	0	0	0
GAJ COUNT	A_NAT_LOW	A	NAT_LOW	3	3	11.4	3.6	31.57894737
STRATUM TOTAL				23	23	6653.6	4343.103614	65.27449221
GAJ COUNT	B_ART_HIGH	B	ART_HIGH	5	5	259.1	87.5	33.77074489
GAJ COUNT	B_ART_LOW	B	ART_LOW	19	10	1148	411.9	35.87979094
GAJ COUNT	B_NAT_HIGH	B	NAT_HIGH	0	0	0	0	0
GAJ COUNT	B_NAT_LOW	B	NAT_LOW	0	0	NA	NA	NA
STRATUM TOTAL				24	15	1407.1	421.1119804	29.92765122
GAJ COUNT	C_ART_HIGH	C	ART_HIGH	0	0	NA	NA	NA
GAJ COUNT	C_ART_LOW	C	ART_LOW	1	2	446.8	345.1	77.23813787
GAJ COUNT	C_NAT_HIGH	C	NAT_HIGH	3	3	123542.4	11636	9.418628746

GAJ COUNT	C_NAT_LOW	C	NAT_LOW	0	0	0	0	0
STRATUM TOTAL				4	5	123989.2	11641.13881	9.388832903
GAJ COUNT	E_ART_HIGH	E	ART_HIGH	0	0	0	0	0
FI SURVEY	E_ART_LOW	E	ART_LOW	1	1	9471.4	0	0
GAJ COUNT	E_NAT_HIGH	E	NAT_HIGH	0	0	0	0	0
GAJ COUNT	E_NAT_LOW	E	NAT_LOW	0	0	0	0	0
STRATUM TOTAL				1	1	9471.4	0	0
GAJ COUNT	F_ART_HIGH	F	ART_HIGH	1	1	0	0	0
FI SURVEY	F_ART_LOW	F	ART_LOW	6	6	7652.6	5757.5	75.23586755
GAJ COUNT	F_NAT_HIGH	F	NAT_HIGH	0	0	0	0	0
GAJ COUNT	F_NAT_LOW	F	NAT_LOW	0	0	0	0	0
STRATUM TOTAL				7	7	7652.6	5757.510808	75.23600878
GAJ COUNT	G_ART_HIGH	G	ART_HIGH	0	0	NA	NA	NA
FI SURVEY	G_ART_LOW	G	ART_LOW	1	1	0	0	0

GAJ COUNT	G_NAT_HIGH	G	NAT_HIGH	0	0	NA	NA	NA	
GAJ COUNT	G_NAT_LOW	G	NAT_LOW	1	1	136587.1	0		0
STRATUM TOTAL				2	2	136587.1	0		0
GRAND TOTAL				61	53	285,761.0	13,700.5	4.8	

FIGURES

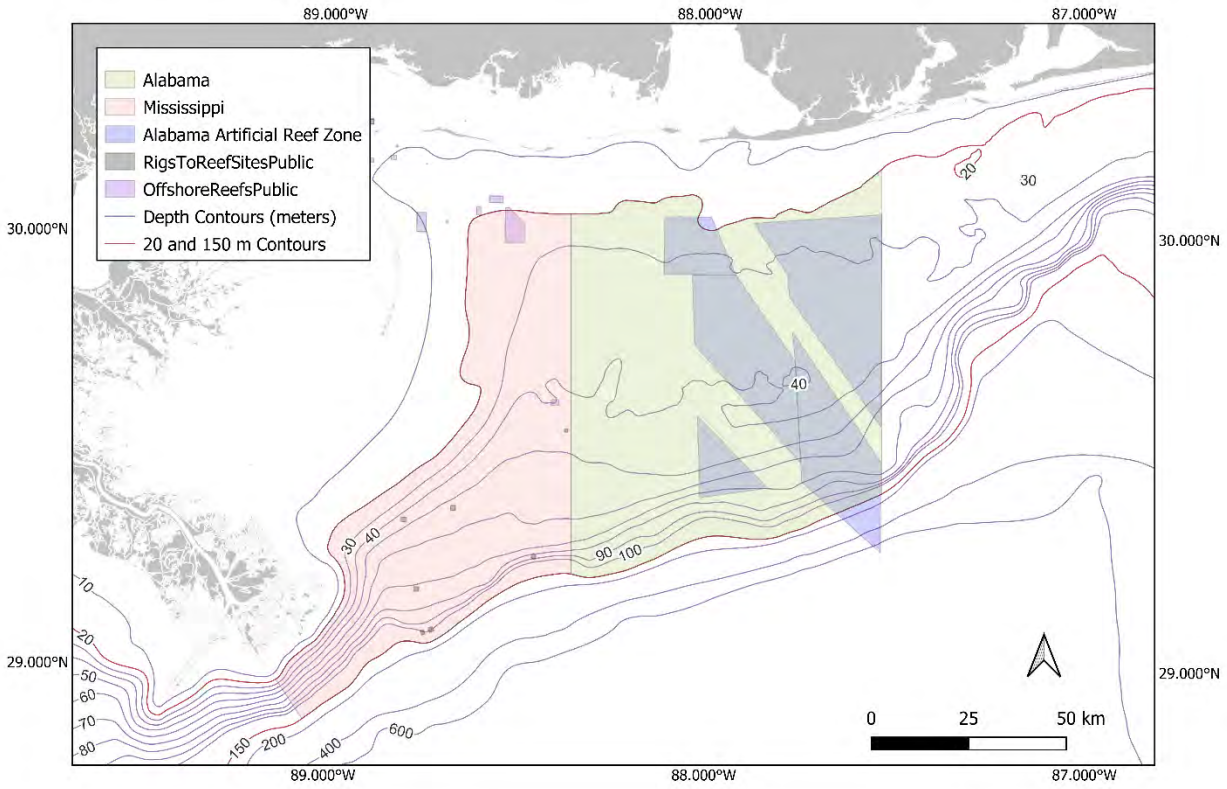


Figure 12.2.1. Sampling frame for Alabama Mississippi region. The region was further subdivided into Mississippi, Alabama (AARZ), and Alabama (non AARZ), with each of these subregions divided into three depth categories (shallow: 20-40m, mid: 40-60m, and deep: 60-150m).

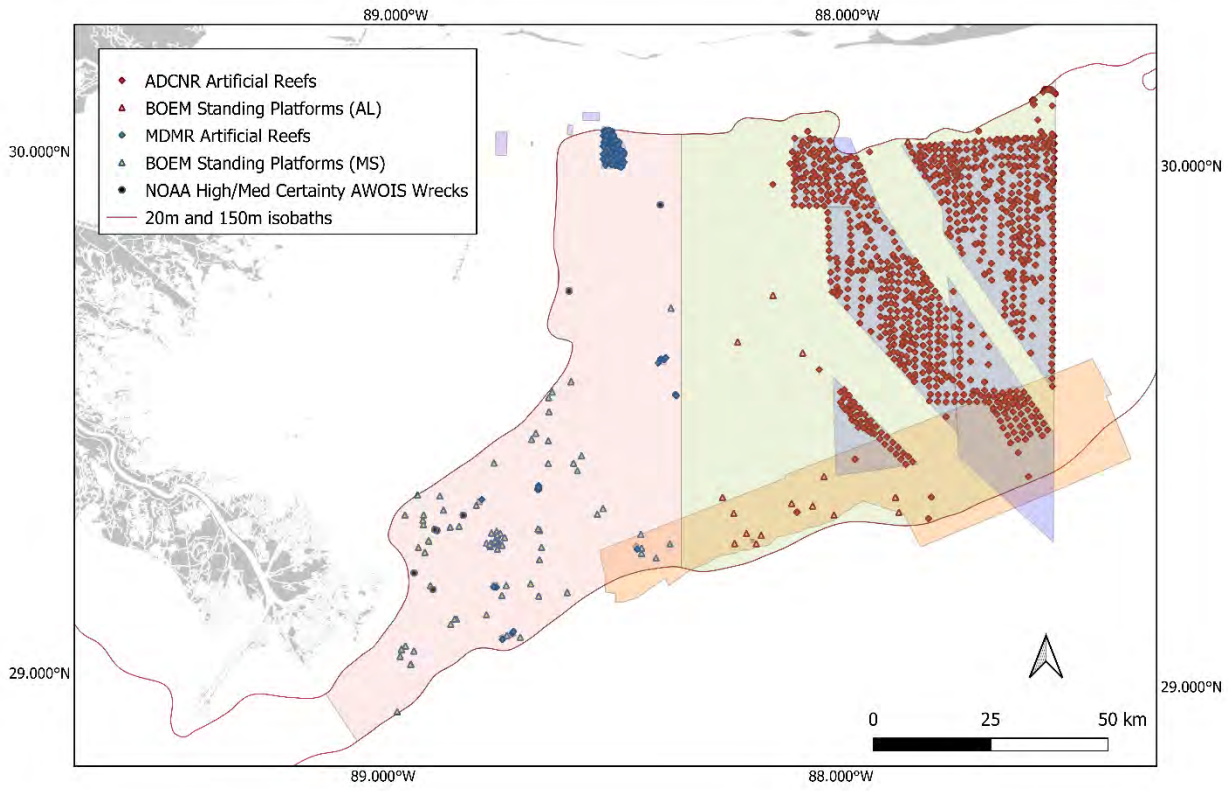


Figure 12.2.2 Map of Alabama-Mississippi region with known artificial structures from state and federal databases.

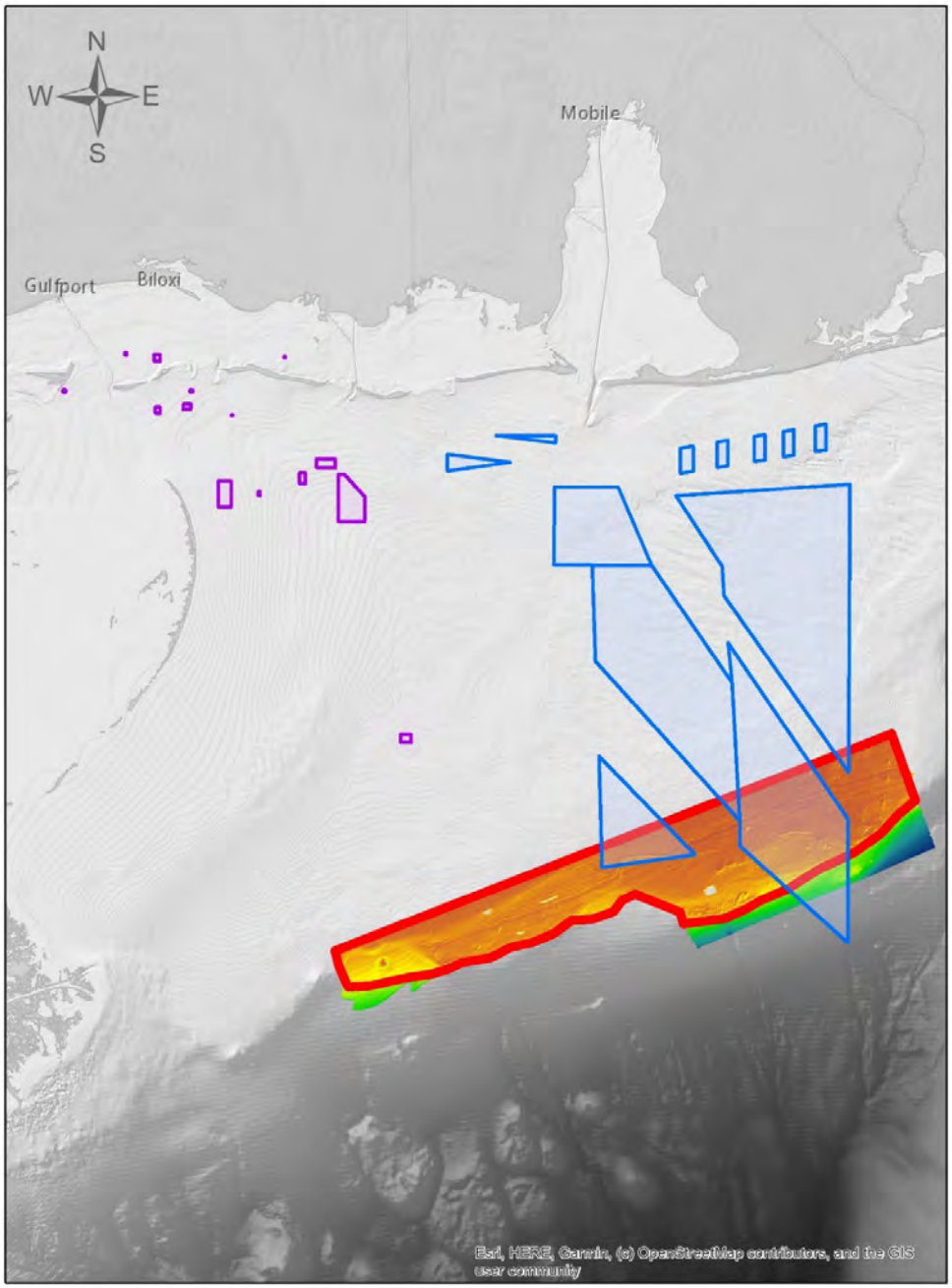


Figure 12.2.3. Mississippi and Alabama coastal waters showing the extent of the natural hard bottom banks (Alabama pinnacles and Alps; red polygon) centered around the 70 m isobath surveyed by the 2002 USGS multi-beam study. The blue and purple polygons represent the Alabama and Mississippi artificial reef zones, respectively.

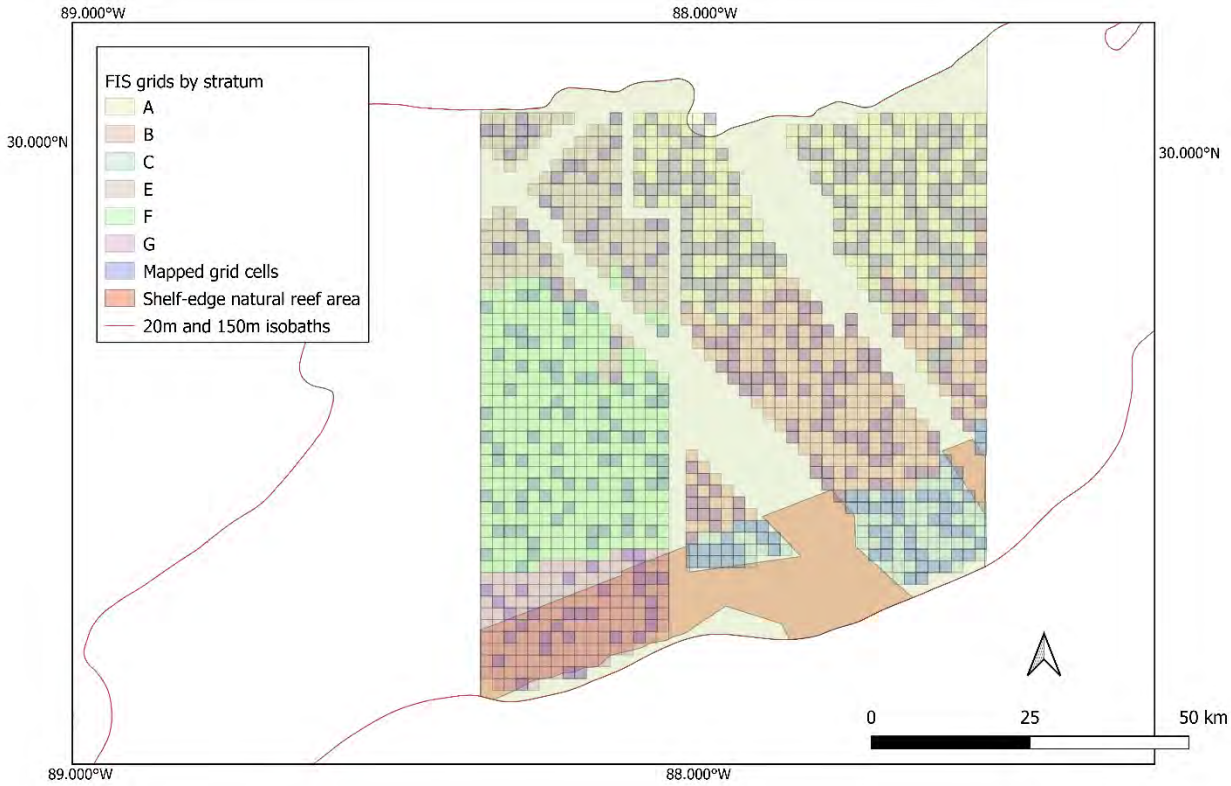


Figure 12.2.4. Alabama Fisheries Independent Survey (FIS) grid. Shaded cells were mapped using side-scan sonar between 2010 and 2021. The area in orange represents the shelf-edge natural reef complex mapped by the USGS in 2002.

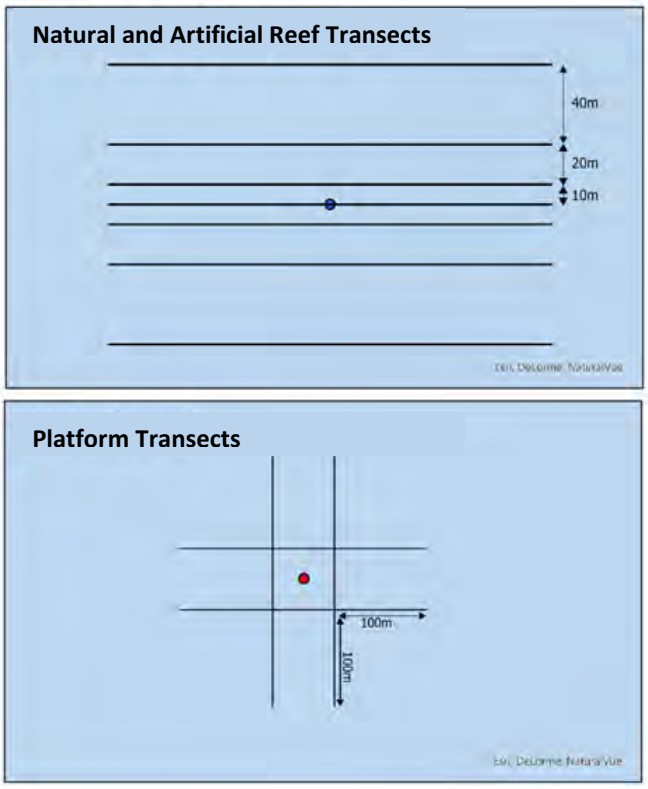


Figure 12.2.5. Hydroacoustic transects performed in the Alabama-Mississippi region for natural and artificial reefs and platforms.

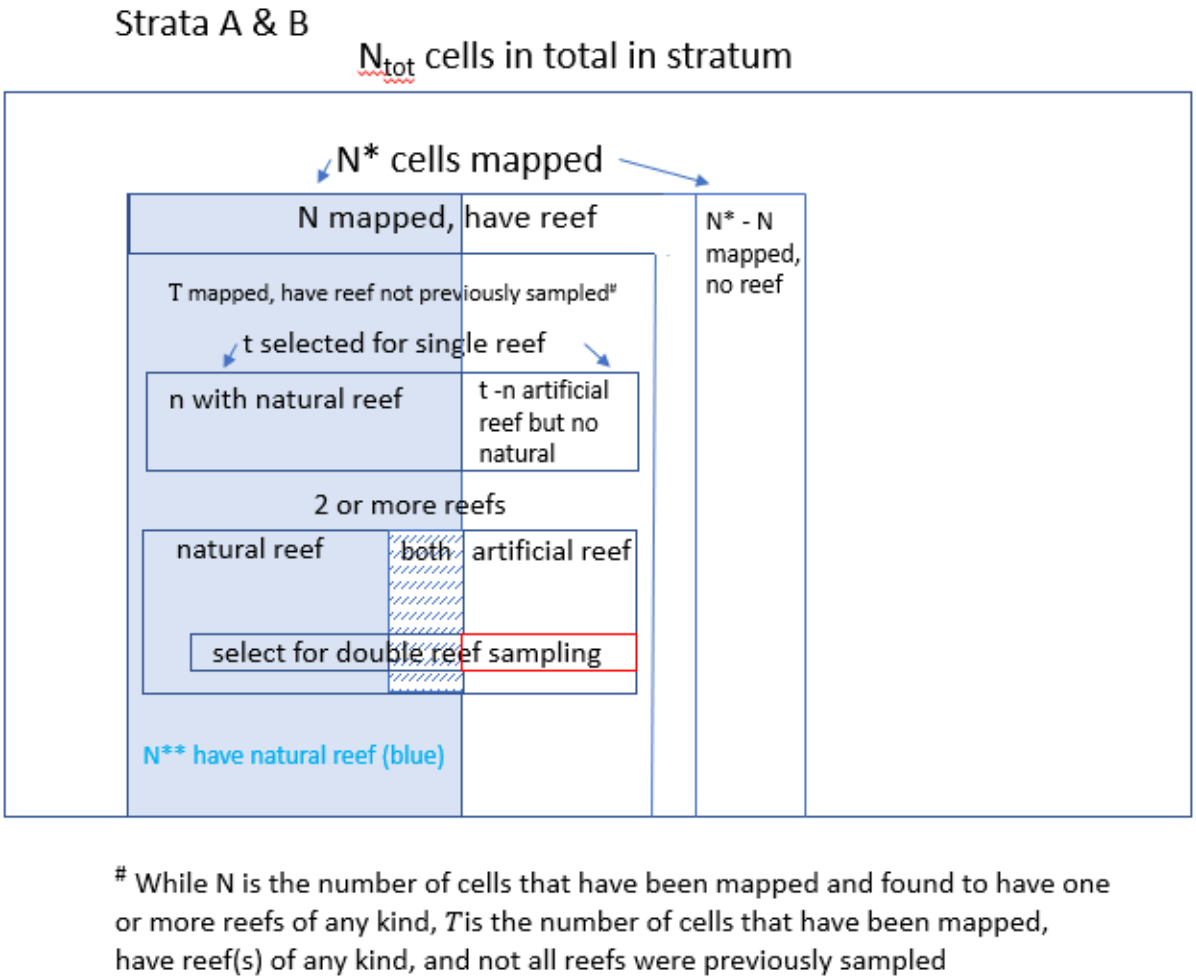
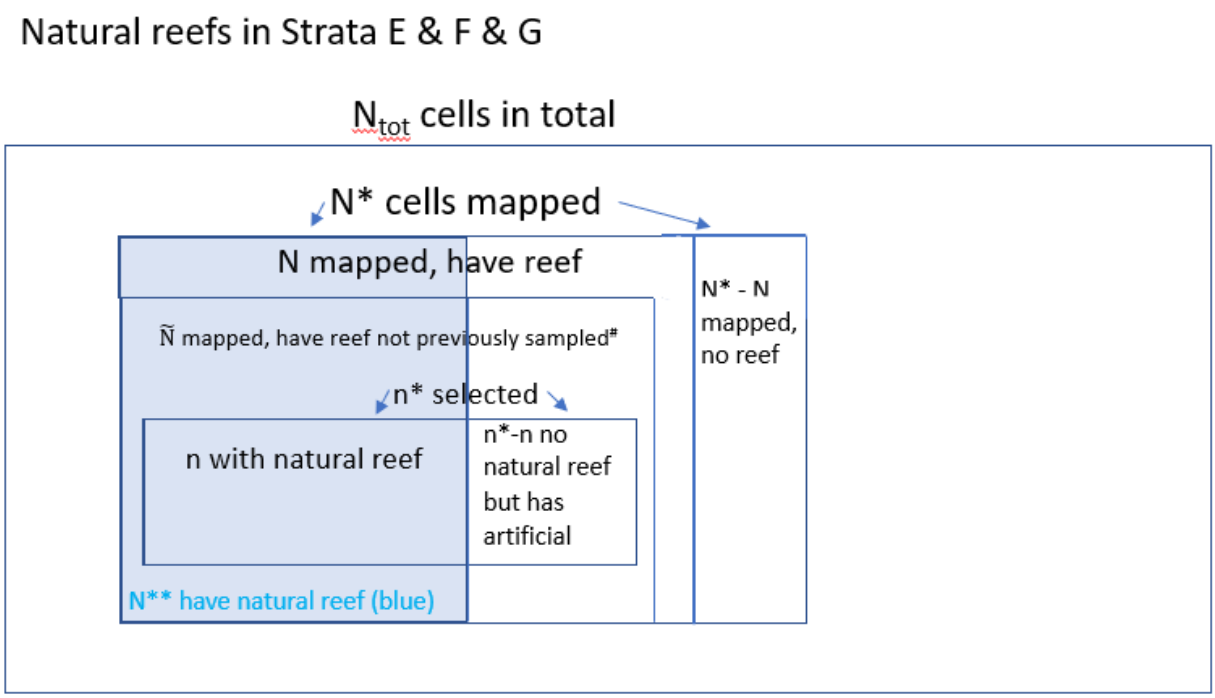


Figure 12.2.6. Notation for sampling strata A and B. Blue cross-hatching indicates cells with two or more reefs and both natural and artificial reefs; blue highlight indicates cells with natural reefs. Red rectangle indicates that the cells chosen to have two reefs sampled all had only artificial reefs (no natural reefs).



While N is the number of cells that have been mapped and found to have one or more reefs of any kind, $\tilde{N}$ is the number of cells that have been mapped, have reef(s) of any kind, and not all reefs were previously sampled

Figure 12.2.7. Notation for sampling strata E, F and G.

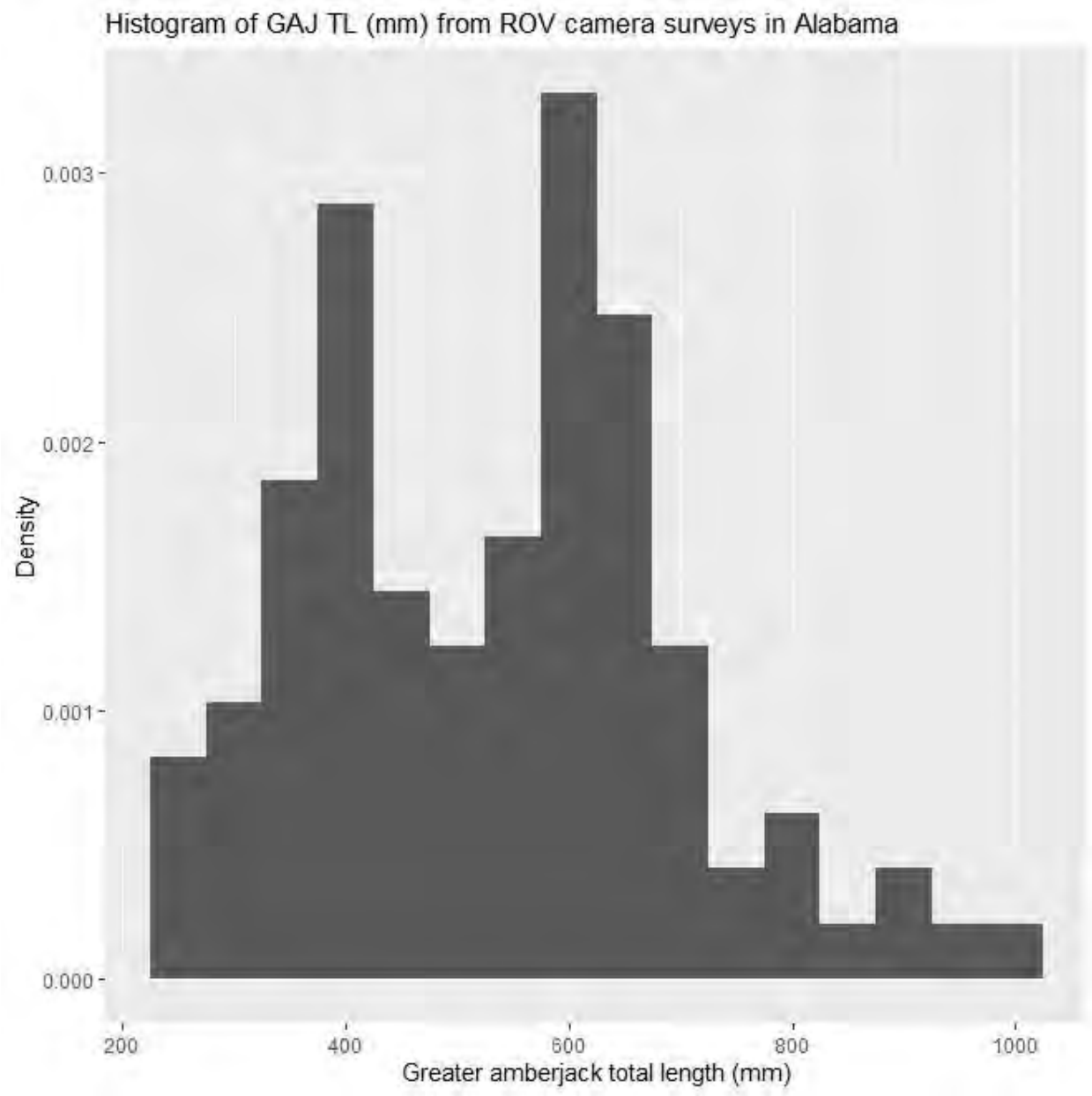


Figure 12.2.8. Histogram of total length measurements of Greater amberjack observed during the ROV survey. Only a subset (n = 97) of observed fish was measured using the laser system.

References

Bacheler, N.M., and Conn, P.B. 2013. Examining the utility of alternative video monitoring metrics for indexing reef fish abundance. Can. J. Fish. Aquat. Sci. 71: 464–471. doi:10.1139/cjfas-2013-0086.

Caimi, F.M., and Tusting, R.F. 1987. Application of lasers to ocean research and image recording systems. In Proceedings of the International Conference on LASERs 1986, pp. 518–524.

Cochran, W.G. 1977. Sampling Techniques, 3rd edition. John Wiley, New York.

Goodman, L. A. 1960. On the exact variance of products. Journal of the American Statistical Association 55:708-713.

Grasty, S. E. 2014. Use of a towed camera system for estimating reef fish population densities on the West Florida shelf. M.S. thesis, x p. University of South Florida, St. Petersburg, Florida.

Gregalis, K. C., L. S. Schlenker, J.M. Drymon, J.F. Mareska, and S.P. Powers. 2012. Evaluating the performance of vertical longlines to survey reef fish populations in the northern Gulf of Mexico. Transactions of the American Fisheries Society 141:1453–1464.

Isermann, D. A., & Knight, C. T. 2005. A Computer Program for Age–Length Keys Incorporating Age Assignment to Individual Fish. North American Journal of Fisheries Management, 25(3), 1153–1160. https://doi.org/10.1577/M04-130.1Schobernd, Z.H.,

Lembke, C., S. Grasty, et al. 2017. The Camera-Based Assessment Survey System (C-BASS): A towed camera platform for reef fish abundance surveys and benthic habitat characterization in the Gulf of Mexico. Continental Shelf Research 151: 62-71.

Powers, S. P., J. M. Drymon, C. L. Hightower, T. Spearman, G. S. Bosarge, and A. Jefferson. 2018. Distribution and Age Composition of Red Snapper across the Inner Continental Shelf of the North-Central Gulf of Mexico. Transactions of the American Fisheries Society, Vol. 147(5): 791-805. https://doi.org/10.1002/tafs.10081

Powers, S. P., J. M. Drymon, C. L. Hightower, L. M. Carleton, and J. M. Hoenig. 2023. Direct estimates of reef fish abundance across an artificial reef network. Canadian Journal of Fisheries and Aquatic Sciences. 80(4): 676-689. https://doi.org/10.1139/cjfas-2022-0094

11.3. Western Florida

Regional Estimate III (West Florida/Eastern Gulf of America)

Ted Switzer, Sean Keenan, Lynne Stokes, Ellen Taylor, Kevin Boswell, Allison White, Steve Murawski, Josh Kilborn

Introduction

In the eastern Gulf of America / West Florida region, the estimation of absolute abundance of Greater Amberjack was accomplished to a large degree by leveraging efforts from the Gulf Fishery Independent Survey of Habitat and Ecosystem Resources (G-FISHER) program. G-FISHER is an annual survey conducted collaboratively by the National Marine Fisheries Service (NMFS) and the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute (FWC-FWRI). Historically involving three separate video surveys, (NMFS – Pascagoula, NMFS – Panama City, and FWC-FWRI), reef fish survey efforts were integrated under G-FISHER beginning in 2020 (Switzer et al. 2023). The G-FISHER survey involves a combination of underwater video surveys using stereo-baited remote underwater video (S-BRUV) arrays and standardized habitat mapping surveys (Keenan et al. 2022) and is designed to estimate relative abundance and conduct habitat characterization of both natural and artificial hardbottom habitats throughout the Gulf of America. To facilitate efforts to utilize data from G-FISHER to estimate absolute abundance of Greater Amberjack, the G-FISHER survey was augmented in several ways. First, at a subset of G-FISHER sampling stations, a concurrent hydroacoustic survey was conducted during S-BRUV deployment to produce an estimate of the distribution and abundance of large-bodied pelagic fishes throughout the water column, including waters above those typically viewed by the stationary S-BRUV camera system. Second, a comparison of abundance estimates between single-camera and spherical-camera systems was conducted to quantify potential biases in single-camera abundance estimates for Greater Amberjack and develop calibration coefficients to estimate abundance where only single-camera reads were conducted. And third, a series of paired baited / unbaited S-BRUV trials were conducted to evaluate the potential influence of bait on abundance estimation of Greater Amberjack.

To complement the G-FISHER survey of natural and artificial hardbottom habitats, a survey of unconsolidated / uncharacterized bottom habitats was conducted using the Camera-Based Survey Assessment System (C-BASS) by the University of South Florida College of Marine Science. Using a paired combination of a towed underwater video and concurrent hydroacoustic survey, this survey was designed to estimate the distribution and abundance of Greater Amberjack throughout the vast areas of presumably non-reef habitat throughout the eastern Gulf of America / West Florida Region.

Combined, these research and survey activities were designed to address the following objectives in the eastern Gulf of America / West Florida Region:

a. Estimate the quantity of natural hardbottom habitat, artificial hardbottom habitat, and unconsolidated habitat from the Florida Keys to the Florida-Alabama border.
b. Provide data on site-specific abundances of Greater Amberjack and other *Seriola* spp. on natural and artificial hardbottom habitats from S-BRUV surveys and, where available, concurrent hydroacoustic surveys.
c. Provide data on site-specific abundances of Greater Amberjack and other *Seriola* spp. on unconsolidated/uncharacterized bottom habitats from C-BASS and concurrent hydroacoustic surveys.

d. Quantify potential biases in abundance estimation of single-camera S-BRUV surveys through the analysis of coincided spherical camera S-BRUV data and develop calibration equations to estimate abundance of Greater Amberjack and other *Seriola* spp. where only single-camera data were available.
e. Evaluate the influence of bait on site-specific abundance estimation of Greater Amberjack and other *Seriola* spp. surveyed using S-BRUV arrays.
f. Produce region-wide estimates of the absolute abundance of Greater Amberjack by habitat, region, and size/age class where possible.

Method

Survey design: Survey of natural hardbottom habitats (G-FISHER)

The G-FISHER reef fish survey of natural hardbottom habitats utilized a stratified-random survey design whereby annual sampling effort is allocated among both spatial and habitat strata. Details of the rationale and development of the G-FISHER survey design can be found in Switzer et al. (2023), but in brief, historical reef fish survey data were analyzed to delineate biologically meaningful spatial and habitat strata based on assemblage-level analyses of data for managed reef fishes using classification and regression tree analyses. In the eastern Gulf of America, a 3 x 3 factorial spatial stratification scheme was implemented, where sampling effort was partitioned into three regional strata (NC: North-Central Gulf, BB: Big Bend, and SF: Southwest Florida) and three depth strata (Nearshore: 10 – 25 m, Offshore: 25 – 55 m, and Deep: 50 – 180 m; Figure 11.3.1). In terms of habitat stratification for sampling of natural hardbottom habitats, a 3 x 3 factorial stratification scheme was also implemented, where sampling effort was partitioned into three levels of relative relief (Low-relief, Moderate-relief, and High-relief) and three levels of reef scale (Small-scale: < 100 m2, Medium-scale: 100 – 1,000 m2, and Large-scale: > 1,000 m2; Figure 11.3.2). Multiple habitat types, or geoforms, identified through standardized habitat mapping surveys were aggregated into the three levels of relative relief (Table 11.3.1).

Allocation of annual sampling effort among identified spatial and habitat strata was based on a combination of both habitat availability and managed species richness. To estimate availability of natural hardbottom habitats for each combination of spatial and habitat strata and inform the G-FISHER design, data from standardized, random habitat mapping surveys conducted through 2019 were analyzed (Keenan et al. 2022). First, composition of hardbottom habitats (percentage of randomly mapped area comprised of each hardbottom stratum) were compiled for each spatial stratum. Next, these percentages were extrapolated to the total area within each spatial stratum to produce region-wide estimates of habitat availability. Estimates of habitat availability were then 4th-root transformed to downweight the contribution of several highly prevalent habitats to effort allocation. In addition, average managed species richness was calculated for each combination of spatial and habitat strata. The final allocation scheme was based on the proportional product of transformed habitat availability data and managed species richness, with adjustments to ensure that at least two sites were selected for sampling annually (Figure 11.3.3; Table 11.3.2).

Prior to sampling each year, sites are randomly selected from a universe of previously identified habitat features (Table 11.3.2). First, a series of primary sampling units (microgrids: 0.1 nm latitude x 0.1 nm longitude) are selected without replacement that meet criteria for each sampling stratum. Next, the specific sampling location within each selected microgrid is randomly assigned to a point within the corresponding habitat stratum polygon. Finally, all selected points are reviewed to ensure a minimum of 100 m spacing between points in adjacent microgrids; if initial spacing is less than 100 m (typically less than 25 sites per year), specific sampling locations are adjusted within the originally selected habitat stratum. During each sampling year, all efforts are expended to ensure that all sites are sampled; however, due to inclement weather, mechanical issues with

vessels, and camera failures, useful data are lacking from a small subset of sites each sampling year. There is no inherent bias in sites dropped each year, although best efforts are applied to try and ensure that sampling effort occurs within each stratum.

In 2023, hydroacoustic surveys were conducted at a subset of randomly selected natural hardbottom sites being sampled as part of the G-FISHER survey. Due to limited availability of hydroacoustic equipment and logistical challenges of multiple concurrent research cruises, we were unable to select sites following the fully stratified G-FISHER design. However, hydroacoustic sampling sites were allocated and randomly selected within all spatial strata.

In 2022, a subset of natural hardbottom G-FISHER survey stations was selected for a paired baited-unbaited deployment study to quantify the impacts of bait on estimation of abundance and size composition of managed reef fishes. These data were augmented with additional sites selected in 2023; sites for the paired baited-unbaited deployment study were selected from sites that were not selected for the hydroacoustic survey.

Survey design: Survey of artificial hardbottom habitats (G-FISHER)

Like the survey of natural hardbottom habitats, the G-FISHER reef fish survey of artificial hardbottom habitats utilized a stratified-random survey design whereby annual sampling effort is allocated among both spatial and habitat strata (Switzer et al. 2023). Spatial stratification of the survey of artificial hardbottom habitats is identical to that of the natural hardbottom habitat survey (Figure 11.3.1). In terms of habitat stratification, a 3 x 3 factorial stratification scheme was also implemented for the survey of artificial hardbottom habitats, where sampling effort was partitioned into three levels of relative relief (Low-relief, Moderate-relief, and High-relief) and three levels of reef scale (Small-scale: < 25 m2, Medium-scale: 25 – 100 m2, and Large-scale: > 100 m2; Figure 11.3.4). Multiple habitat types, or geoforms, identified through standardized habitat mapping surveys were aggregated into the three levels of relative relief (Table 11.3.3).

Allocation of annual sampling effort among identified spatial and habitat strata was based on a combination of both habitat availability and managed species richness. Due to the assumption that the universe of artificial hardbottom habitats was fully described, there was no need to extrapolate estimates of total habitat availability. However, most known artificial reef sites have not been mapped. Therefore, to estimate availability of artificial hardbottom habitats for each combination of spatial and habitat stratum, locations of all known artificial hardbottom habitats were intersected with microgrids (0.1 nm latitude x 0.1 nm longitude), and the total area of all microgrids was summed for each sampling stratum. Estimates of habitat availability were then 4th-root transformed to downweight the contribution of several highly prevalent habitats to effort allocation. In addition, average managed species richness was calculated for each combination of spatial and habitat strata. The final allocation scheme was based on the proportional product of transformed habitat availability data and managed species richness, with adjustments to ensure that at least two sampling sites were selected annually (Figure 11.3.5; Table 11.3.4). Because artificial hardbottom habitats are much less prevalent than natural hardbottom habitats in the eastern Gulf of America, fewer total sites were selected each year.

Prior to sampling each year, sites are randomly selected from a universe of previously identified artificial hardbottom habitat features (Table 11.3.4). First, a series of primary sampling units (microgrids: 0.1 nm latitude x 0.1 nm longitude) are selected that meet criteria for each sampling stratum; microgrids are selected without replacement for each sampling stratum but would be eligible for selection for other sampling strata should multiple habitat strata be present, although this is rare. Next, the specific sampling location within each selected microgrid is randomly assigned to a point within the appropriate habitat polygon. Finally, all selected points are

reviewed to ensure there are no spatial conflicts (sites within 100 m of each other), and all conflicts are resolved to ensure full implementation of the intended survey design. During each sampling year, all efforts are expended to ensure that all sites are sampled; however, due to inclement weather, mechanical issues with the vessel, and camera failures, useful data are lacking from a small subset of sites each sampling year. There is no inherent bias in sites dropped each year, although best efforts are applied to try and ensure that sampling effort occurs within each stratum.

In 2023, hydroacoustic surveys were conducted at a subset of randomly selected artificial hardbottom sites being sampled as part of the G-FISHER survey. Due to limited availability of hydroacoustic equipment and logistical challenges of multiple concurrent research cruises, we were unable to select sites following the fully stratified G-FISHER design. However, hydroacoustic sampling sites were allocated and randomly selected within all spatial strata.

In 2022, a subset of artificial hardbottom G-FISHER survey stations was selected for a paired baited-unbaited deployment study to quantify the impacts of bait on estimation of abundance and size composition of managed reef fishes. These data were augmented with additional sites selected in 2023; sites for the paired baited-unbaited deployment study were selected from sites that were not selected for the hydroacoustic survey.

Survey design: Survey of unconsolidated / uncharacterized habitats (C-BASS)

The survey domain for the survey of unconsolidated or uncharacterized habitats extended from the Louisiana-Mississippi Border to the Florida Keys, including waters 20 – 150 m deep (Figure 11.3.6). The approximate total area of the sampling frame was 119,000 km2. The sampling frame was gridded into 5’ by 5’ blocks, and subregions were defined (NC: North-Central Gulf, BB: Big Bend, and SF: Southwest Florida) based on G-FISHER regions (Figure 11.3.1).

To inform the sampling design for the C-BASS survey, data from FWRI and NOAA S-BRUV surveys from 2010 – 2020 (N = 9,426) were compiled and georeferenced, and MaxN abundance data for *Seriola dumerili* were intersected with the gridded sampling frame (Figure 11.3.7). For each grid cell, the total number of sampling events was calculated (Figure 11.3.8), as was the mean observed count of S. *dumerili* (Figure 11.3.9). Within each region, grids were assigned to one of four strata based on sampling effort and average S. *dumerili* abundance (varies by subregion) as: Never Sampled (no survey effort reported within a grid cell); Zero Catch (sampling effort occurred but no S. *dumerili* observed; Low Catch (relatively low CPUE for S. *dumerili*); High Catch (relatively high CPUE for S. *dumerili*). It was estimated that approximately 90 stations could be sampled given available ship time; sampled grids were randomly selected and allocated among strata with proportionately higher sampling effort applied to both Low Catch and High Catch grids (Table 11.3.5). For each randomly selected grid, a randomly selected compass direction (0 – 359°) was assigned, and an approximately 6.5 nm transect line was generated using the midpoint of the sampling grid as the midpoint of the transect, oriented in the randomly selected compass direction.

Survey design: Habitat mapping

Although several specific regions have been well mapped (see overview by Hine and Locker 2011), a comprehensive habitat map does not exist for most of the West Florida region. The West Florida region includes a diversity of habitats ranging from the biogenic reefs in the Florida Keys to the extensive areas of flat hard bottom in the Big Bend region and finally includes isolated hard bottom areas in the Florida Panhandle. Included within the matrix of natural hard bottom and unconsolidated sediments are concentrations of man-made

artificial reef habitats including construction material, vessels and prefabricated modules among others (Paxton et al. 2024). Available literature on the benthic habitats of the West Florida region provide valuable context on the diversity and geology of habitat throughout the region. However, our understanding of benthic habitats is insufficient to provide reliable estimates of overall habitat coverage and composition suitable for scaling density data to habitat-specific absolute abundance estimates and lacks the fine-scale resolution to direct sampling efforts.

As part of what would evolve into the G-FISHER reef fish survey, the Florida Fish and Wildlife Research Institute – Fisheries Independent Monitoring program (FWRI-FIM) implemented a habitat mapping survey in 2010 to identify and characterize hardbottom habitats. Spatial extent of the FWRI-FIM reef fish and habitat mapping survey expanded through time as the funding supporting these efforts increased. From 2010 – 2013, these efforts were restricted to waters 10–110 m between 26º and 28º N latitude corresponding to the NMFS Statistical Zones (statistical zones) 4 and 5 (Keenan et al. 2022; Figure 11.3.10). In 2014, the reef fish survey area expanded to include the western Florida Panhandle (from 86º to approximately 87º 30’ W; encompassing statistical zones 9 and 10) and waters out to 180 m depth. Beginning in 2016, sampling was further expanded into all statistical zones along the Florida coastline (statistical zones 2–10), from the Florida Keys and Dry Tortugas (terminates at 24º 35.0’ N when east of 83º W) to the Florida-Alabama border 87º 30’ W. Thus, the total area that could be surveyed (i.e. potential survey area) was approximately 146,534 km2 (Figure 11.3.10).

Survey locations from 2010 to 2023 were randomly selected from a gridded universe within the potential survey area using ArcGIS. Standardized and randomized habitat-mapping surveys (each covering an area of approximately 2.0 km2 for surveys between 2010 and 2019 and approximately 4 km2 from 2020 to 2023) were conducted to identify and delineate hard bottom habitats using a side-scan sonar (L3-Klein 3900) operating at 445 kHz. Annual survey efforts were allocated among predetermined depth strata (10–25 m; 25.1–50 m; and 50.1–180 m) within each statistical zone with a goal of 100 surveys per statistical zone per year. Survey lines generally ran perpendicular to the shore, covering survey areas approximately 0.5 km N–S and 4 km E–W (2010-2018) or 8 km E-W (2019-2023) along the peninsula (statistical zones 2–6) or 0.5 km E–W and 4 km N–S (2010-2018) or 8 km N-S (2019-2023) in the panhandle (statistical zones 7–10). Each survey plan allowed for 10% overlap between mapping swaths to correct for potential positional errors. Habitat mapping data were then post-processed to delineate hardbottom habitat features; features were classified into a variety of geologic, biogenic, and anthropogenic reef features that were then aggregated into G-FISHER habitat strata. See Keenan et al. (2022) for additional details on data collection and processing methods.

Field methods: S-BRUV surveys

All natural and artificial hardbottom sites were sampled using an S-BRUV array consisting of five independent stereo-video recorders mounted orthogonally to each other inside a 0.8m wide × 1.0m tall aluminum frame (Figure 11.3.11). Cameras were positioned to provide a near-360-degree view of the surrounding seafloor. Two of the five cameras were paired with a second satellite camera (one at a baseline of 231 mm and the other at a baseline of 300 mm) mounted in parallel to provide stereo video measurements; each S-BRUV was calibrated before and immediately after the sampling season to ensure accuracy of stereo-video measurements. All cameras were synched and recorded with an 86o horizontal field of view and 61o vertical field of view at a pixel resolution of 1,920 H x 1,200 V. Data were recorded to an onboard computer system which also provided power via an external battery. An In-Situ Aqua Troll 500 water quality datasonde was mounted to each S-BRUV to record temperature (°C), salinity (psu) and dissolved oxygen (mg L-1). Each S-BRUV was baited with four

mackerel (Scomber spp.) halves (approx. 450 g) and approximately 225 g of squid (Illex sp.) and deployed for 35 minutes prior to retrieval.

At a subset of natural and artificial hardbottom sites, a paired hydroacoustic survey was conducted during the S-BRUV deployment. Full details of hydroacoustic sampling methods are provided in Chapter 9, but in brief, hydroacoustic equipment was deployed using either a towed, pole-mounted, or hull-mounted system depending on the vessel. Regardless of deployment method, a combination of three (38 kHz, 70 kHz, 120 kHz) Kongsberg EK 80 wideband autonomous transceivers was used to map the three-dimensional distribution of fish surrounding the S-BRUV array. At each site, the hydroacoustic survey consisted of a series of seven 200-m parallel transects; the center transect passed over the site of the S-BRUV deployment, while each successive parallel transect was increasingly further away from the center line (two transects at a targeted 7-m distance, two transects at a targeted 18-m distance, and two transects at a targeted 35-m distance).

Field methods: C-BASS surveys

Unconsolidated sediment habitats were sampled using the Camera-Based Assessment Survey System (C-BASS; Grasty 2014; Lembke et al. 2017) which consisted of a series of six high-definition video cameras used in stereo and individually (Figure 11.3.12). The system also has a range of environmental sensors (temperature, depth, salinity, CDOM) and navigation-related sensors (e.g., pitch, roll, yaw, altitude). The system is "flown" in real time using imagery supplied from the tow sled to the vessel deck. Depth alterations are done via adjusting the towing cable length to affect the height over bottom. The system is generally flown at an altitude of 2-5 m above the sea floor, depending on water visibility and bottom topography, and is towed at speeds of 1.5 – 2 m s-1. All efforts were made to sample the selected transects as planned; however, unforeseen constraints to sampling (e.g., presence of oil rigs or obstacles in the transect path, unsafe proximity to high traffic areas, presence of actively fishing vessels) resulted in transects that were either cut short of the 6.5 nm target, moved to an adjacent grid cell, or towed in an alternate randomly selected compass direction.

Field methods: Paired baited / unbaited S-BRUV surveys

At a subset of natural and artificial hardbottom G-FISHER sites, paired baited-unbaited S-BRUV deployments were conducted in 2022 and 2023. At each site, deployment order (baited-unbaited or unbaited-baited) was randomly assigned prior to sampling to account for the potential confounding effect of deployment order. Once each sampling site was reached the appropriate S-BRUV gear (if baited, the standard G-FISHER bait was used; if unbaited, no bait was used) was deployed. After the standard deployment time, the S-BRUV was retrieved, and the complementary S-BRUV deployment was conducted at the same sampling site.

Data processing methods: S-BRUV surveys

For all S-BRUV deployments, the relative abundance of all fishes was estimated as MaxN, or the maximum number of individuals of each taxon (note that counts from different taxonomic levels are not mutually exclusive, for example *Seriola dumerili* and *Seriola* spp., so are not additive) observed in a single screen frame to avoid the possibility of double-counting. MaxN was limited to 299 or less, as beyond 299, it is difficult to estimate the number of individuals observed accurately. Each video was viewed for a total of 20 minutes, beginning approximately 10 minutes after the gear reached the bottom; in addition, MaxN of *Seriola* spp. was determined prior to the standard G-FISHER read time (e.g., during the camera drop). All fishes were identified to the lowest possible taxonomic level, most commonly species. For all S-BRUV sets, MaxN was determined by processing data from one of the five cameras on the array. The camera chosen for estimation of MaxN was the

one with the best view of targeted hardbottom habitat; when multiple cameras had similar views of hardbottom habitat, the camera chosen for estimation of MaxN was randomly chosen, unless one was paired with the satellite stereo-video camera. A variety of metrics of habitat quality were characterized for each of the five cameras, including various measures of relief, substrate coverage and composition, biota coverage and composition, and habitat complexity among others. Physical and habitat data were synthesized to provide the suite of covariates included in routine methods to develop indices of relative abundance (Thompson et al. 2020). These include G-FISHER habitat and spatial strata, month, latitude and longitude (in decimal degrees), depth (m), relative visibility (a qualitative measure of water clarity ranging from 0, or clear, to 10, or unreadable), substrate maximum relief (m), and the presence/absence of several classes of attached biota (sponges, hard coral, soft coral, algae, seagrass, or other unidentified sessile organisms).

For instances where the camera read for estimation of MaxN was paired with a satellite camera, stereo-video measurements were obtained for a subset of taxa, including all managed species as well as any species that attains a maximum size of greater than 300 mm. Measurements were taken at a single point during the video in which the most number of individuals were measurable (both the snout and tail were visible on both cameras and the individual was generally oriented perpendicular to the video camera). Stereo-video measurements represent fork length (FL, in mm). Often only a subset of individuals observed on video were measurable, and there are instances where no individuals were measurable due either to stereo camera malfunction or fish orientation.

For all S-BRUV sites paired with hydroacoustic surveys, full spherical cameras reads were completed to avoid potential underestimation of site-specific abundance from single-camera reads. In addition, for sites with hydroacoustic surveys, the abundance of all *Seriola* taxa (*Seriola dumerili*, S. *fasciata*, S. *rivoliana*, S. *zonata*, and *Seriola* spp.) was determined from the period of the initial drop through the start time of the standard read to account for any *Seriola* spp. that may have been present prior to normal video reading times.

Data processing methods: C-BASS surveys

Imagery from the C-BASS video transects were processed in 15 s intervals, although scrolling a few sections in each direction was allowed to provide context and ensure that classification adequately characterized the area. Observed habitats were classified according to the substrate and biotic components of the Coastal and Marine Ecological Classification Standard (CMECS) scheme. Characterized metrics included seafloor induration, cover and composition of primary attached biota, and primary and secondary components of hardbottom and softbottom habitat, respectively. In addition, all observed fish were enumerated and identified to the lowest possible taxon; if no taxonomic distinction was possible, observed individuals were classified as unidentified small and large fishes, depending on size. Positional data, including roll, pitch, yaw, altitude, and tow speed were analyzed to estimate area viewed by each transect. All instances where Greater Amberjack were observed on C-BASS in association with natural or artificial hardbottom habitats were excluded, as these habitats were surveyed by S-BRUVs.

Analytical methods: Developing calibration coefficients between single-camera and spherical-camera S-BRUV reads

Abundance estimation from interpretation of data from a single video camera have the potential of underestimating true site-specific abundance, particularly for schooling species or at high site abundances where more individuals are present than can fit on a single video frame. To address this issue, Schobernd et al. (2014) developed an alternative to MaxN, termed MeanCount (or SumCount), that they suggested more linearly

tracked true abundance. While this may be the case, Campbell et al. (2015) demonstrated that alternative modeling relationships (e.g., power regression) could be applied to the interpretation of MaxN data. To further evaluate these issues, a fully spherical camera system was developed, and data on paired single-camera and spherical-camera reads have been collected over the past several years to quantify whether statistically robust calibration equations could be developed to predict total site abundance from single-camera MaxN data. Accordingly, estimates of spherical-camera counts were generated for sites where only single-camera reads were calculated using Poisson regression in a Bayesian framework, details of which are provided in (Chapter 10).

Analytical methods: Comparison between baited and unbaited S-BRUV deployments

Comparisons of fish abundance between baited and unbaited S-BRUV deployments were conducted using a multispecies approach and permutational analysis of variance (PERMANOVA). First, data were filtered to include taxa that were (1) identified to species, and (2) occurred in a minimum of 5% of all baited and unbaited camera reads. Data were fourth root transformed to reduce the influence of highly abundant taxa, and a Bray-Curtis dissimilarity matrix was constructed by adding in a dummy value to facilitate the calculation of paired similarity values. A mixed-effects PERMANOVA was then fit to determine whether assemblage structure differed between the fixed-effects factors of Bait (baited or unbaited) or Habitat (natural or artificial), and whether the Bait x Habitat interaction was significant. Sampling Site, nested within habitat, was included as a random effect. For factors found to be significant, similarity percentages (SIMPER) analyses were conducted to identify taxa that cumulatively contributed to 85% of observed differences in assemblage structure.

Analytical methods: Estimation of absolute abundance of Greater Amberjack – camera-based estimate

Data from stereo camera and historical habitat mapping surveys were integrated to produce estimates of Greater Amberjack for West Florida. Several critical data sources were compiled in support of these efforts (Figure 11.3.13), details of which follow below.

Site-specific abundance

The abundance of Greater Amberjack at each sampling site was estimated in spherical camera abundance currency. For sites where spherical camera reads were conducted, site-specific abundance represented MaxN from the spherical camera reads. For sites where only single camera reads were conducted, single camera counts were converted to spherical camera counts using the Bayesian calibration conversion model (Chapter 10).

Effective sampling area of reef habitat

Analysis of S-BRUV videos yields a qualitative measure of relative visibility which can be used to estimate the area of the seafloor viewed for each camera site (on the order of meters; Figure 11.3.14). However, it is likely that some proportion of individuals observed originate from areas beyond that which can be directly observed on video. As the S-BRUVs used in this survey are baited, there is likely some degree of directional bait attraction which would attract fish from an unknown distance. It is also likely that there is a measure of reactiveness to gear moving through the water. Accordingly, using the area observed by the camera as a denominator for density estimation would likely result in a significant overestimation of true Greater Amberjack density at each sampling site.

We examined two data sources to guide our coarse estimation of effective sampling area. First, we examined several years of historical data to characterize the timing of MaxN at both natural and artificial hardbottom

habitats (Figure 11.3.15). Enumeration of MaxN begins approximately 10 minutes after the S-BRUV reaches the seafloor, and 20 minutes of video are read in total. Therefore, timing of MaxN ranged from 10 – 30 minutes from the time the S-BRUV reached the seafloor. For natural reef surveys, MaxN occurred at 18.2 minutes on average (median of 17.8 minutes), while for artificial reef surveys, MaxN occurred at 18.0 minutes on average (median of 18.5 minutes). In the eastern Gulf of America, the average current velocity is ~ 0.1 m/s (Liu and Weisberg 2012), so in that time we estimated that on average the bait plume could travel a distance of ~ 108 m. The second data source we examined involved available VPS telemetry data from both calibration studies conducted off Alabama (Chapter 10) as well as an unrelated study conducted off North Carolina. For the North Carolina study, data for Greater Amberjack were queried for 30-minute time periods (identical to the total bottom deployment time of the S-BRUV arrays) with multiple positions recorded to produce an estimate of home range using a weighted autocorrelated kernel density estimator (Figure 11.3.16). In total, three estimates were produced for five individual Greater Amberjack; estimated area of use ranged from 2,782 – 254,666 m2, with an average area of 67,357 m2 and a median area of 44,613 m2. In comparison, the area of a circle with a radius of 108 m (the estimated bait plume distance) is 36,644 m2. Based on this information, we chose to apply an effective sampling area of 36,644 m2 as the primary measure of effective sampling area for all subsequent analyses. However, understanding that the choice of effective sampling area is highly uncertain, we also conducted sensitivity analyses using a radius of 10 m (the estimated distance of visibility under good conditions; 314 m2) and 50 m (in between the two; 7,854 m2). Full details are provided for analyses using the 108 m buffer, whereas only resultant final abundance estimates are presented for analyses using the 10 m and 50 m buffers.

Within the effective sampling area for each site, both reef and non-reef habitats are often present. As the objective of the S-BRUV survey was to estimate Greater Amberjack abundance on natural and artificial hardbottom habitats, we calculated the total area of reef habitat within the effective sampling area (assuming that the contribution of non-reef habitats to observed Greater Amberjack abundances was minimal based on data from the C-BASS survey). As all sites have previously been mapped, we first clipped available habitat mapping data within a 108 m radius for each sampling site (Figure 11.3.14). For sites with complete habitat mapping coverage, we then calculated the total area coverage of natural and/or artificial hardbottom habitat within the clipped buffer by summing up the area of all habitat polygons. For sites with incomplete habitat mapping coverage, estimates of natural and artificial hardbottom composition were extrapolated to the unmapped portion. Results from these analyses produced estimates of natural, artificial, and total hardbottom habitat within the effective sampling area of each camera site, and these estimates of reef habitat coverage, not the total effective sampling area of 36,644 m2, were used in calculations to convert site-specific Greater Amberjack counts to densities.

Estimating site selection probabilities

Site selection probabilities included two components: probability a microgrid was mapped and the probability a mapped microgrid was selected for sampling. The probability that a microgrid was mapped was the number of mapped microgrids within each spatial stratum divided by the total number of microgrids within the spatial stratum. The probability that a microgrid is selected for sampling for a particular natural or artificial hardbottom stratum each year was calculated by dividing the annual sampling allocation by the total number of microgrids available for random selection for each sampling stratum (Tables 11.3.2 and 11.3.4 for natural and artificial hardbottom habitats, respectively).

Estimating natural, artificial, and non-reef habitat coverage

Data from standardized, random habitat mapping surveys (n = 4,137) conducted from 2010 – 2023 were analyzed to estimate coverage of natural and artificial hardbottom habitats and unconsolidated sediment habitats throughout the eastern Gulf (Figures 11.3.10; 11.3.17). For each survey, the average proportion of the mapped areas containing natural hardbottom habitat, artificial hardbottom habitat, and unconsolidated sediment habitats was first calculated, then these proportions were averaged over all surveys (± SE) then used to estimate habitat coverage throughout the unmapped areas within each spatial stratum. This provided a spatially explicit estimate of benthic habitats throughout the West Florida survey frame. Because the survey footprint expanded through time as funding allowed, total mapping effort varied among NMFS statistical zones. Accordingly, summaries of habitat proportionality were first calculated for each zone and depth bin, then summed across the entire survey domain to provide domain-wide estimates of natural and artificial hardbottom habitats as well as unconsolidated sediment habitats.

Estimating Greater Amberjack size composition

To partition the final absolute abundance estimates into size/age classes, we first generated overall size composition of Greater Amberjack for the S-BRUV surveys by spatial stratum for natural and artificial reef habitats, respectively (collapsing over habitat strata). Due to sparseness of size composition data at the spatial stratum level, size composition data were generated in 150 mm standard length (SL) bins. No size data were available for the Greater Amberjack observed on the C-BASS survey of unconsolidated sediment habitat.

Estimating Greater Amberjack abundance

A design-based estimate of Greater Amberjack absolute abundance was generated from an analysis of S-BRUV (natural and artificial reef habitats) and C-BASS data in the eastern Gulf of America. An overview of general analytical methods was provided in Chapter 10.1.

Analytical methods: Estimation of absolute abundance of Greater Amberjack – acoustics-based estimate

Data from paired camera and hydroacoustic surveys were integrated to develop an acoustics-based estimate of Greater Amberjack. Due to limitations in funding and the availability of acoustic gear, there were multiple natural (Table 11.3.2) and artificial reef sampling strata (Table 11.3.4) for which no acoustic data were collected. To compensate for this and enable region-wide estimation of Greater Amberjack abundance using acoustic data, the decision was made to collapse over the area coverage and relative relief component of habitat strata, while retaining spatial stratification (three regions by three depth strata) for both natural and artificial hardbottom habitats, respectively. The number of sites where acoustic data were available varied, however at least one site was completed within each spatial stratum (Table 11.3.6). Details of estimation methods follow below.

Acoustic-based estimates of potential Greater Amberjack abundance

Hydroacoustic data were processed according to protocols described in full detail in Chapter 9. These analyses produced estimates of potential *Seriola*, or all targets thought to be greater than approximately 200 mm FL, observed more than 0.5 meters above the sea floor, and not tightly aggregated. These estimates likely contain targets from large-bodied fishes other than Greater Amberjack; therefore, estimates of potential Greater Amberjack then need to be corrected based on species composition from coincident video surveys.

Estimating Greater Amberjack partitioning ratios

As acoustic estimates of potential *Seriola* include all large-bodied fishes, we generated site-specific estimates of the percentage of large-bodied fishes observed that were confirmed to be Greater Amberjack. Estimates of total large-bodied fishes per site include all species that routinely reach sizes of 200 mm or greater. Notable exceptions to this included two species (Blue Runner and Vermilion Snapper) who often occur in large numbers, a significant percentage of which are small individuals. For these two species, we calculated the percentage of individuals measured that were greater than 200 mm, then adjusted counts by multiplying by this percentage. Stereo measurements are often not possible for a particular sampling site, therefore these corrective percentages were calculated and applied globally.

Estimating Greater Amberjack abundance

Due to the decision to collapse across habitat strata (reef area and relative relief), acoustic-based estimates were generated for natural and artificial habitat for each spatial stratum. For all sites where the number of large-bodied fishes observed was > 30 individuals, we used site-specific correction ratios. For all other sites, camera data were aggregated by relief strata (combining all reef area components of habitat strata) within each spatial stratum.

For each acoustics survey, estimates of the total area swept were calculated. As surveys cover both reef and non-reef habitats, we estimated the total reef area within the acoustics survey by multiplying the total area swept by the estimated percentage of sampling area comprised of reef habitat as determined from the 108 m buffer analyses. These data were then used to generate design-based estimates of Greater Amberjack.

Results

Survey accomplishments: Survey of natural hardbottom habitats (G-FISHER)

In 2023, a total of 934 successful S-BRUV deployments were completed on natural hardbottom habitats in Florida Gulf waters (Table 11.3.2). Of these, concurrent hydroacoustic surveys were conducted at 203 sites (Table 11.3.2).

From single-camera natural hardbottom video reads, *Seriola dumerili* were observed on 128, or 14%, of all videos read (Figure 11.3.18). Unidentified *Seriola* spp. were present on 62 total natural hardbottom sites (Figure 11.3.19), including 49 where *Seriola dumerili* were not observed. In total, 439 *Seriola dumerili* were observed on natural hardbottom habitats on single-camera reads, with average relative abundances highest in the North Central Gulf region (Table 11.3.7; Figure 11.3.20) and in both high-relief habitats as well as small, low-relief habitats (Table 11.3.8; Figure 11.3.21). A total of 271 *Seriola dumerili* were measured with stereo video; individuals ranged from 180 – 1000 mm FL, although individuals ranging from 380 – 440 mm FL were most commonly observed (Figure 11.3.22).

From spherical-camera natural hardbottom video reads, *Seriola dumerili* were observed on 15, or 7%, of all videos read (Figure 11.3.23). When examining video footage prior to standard video read time, an additional 8 instances of *Seriola dumerili* were identified. Unidentified *Seriola* spp. were present on 11 spherical camera reads (Figure 11.3.234, with an additional 5 instances identified when examining video footage prior to standard video read times.

Survey accomplishments: Survey of artificial hardbottom habitats (G-FISHER)

In 2023, a total of 195 successful S-BRUV deployments were completed on artificial hardbottom habitats in Florida Gulf waters (Table 11.3.4). Of these, concurrent hydroacoustic surveys were conducted at 41 sites (Table 11.3.4).

From single-camera artificial hardbottom video reads, *Seriola dumerili* were observed on 58, or 30%, of all videos read (Figure 11.3.25). Unidentified *Seriola* spp. were present on 40 total artificial hardbottom sites (Figure 11.3.26), including 16 where *Seriola dumerili* were not observed. In total, 297 *Seriola dumerili* were observed on artificial hardbottom habitats on single camera reads, with average relative abundances highest in the North Central Offshore strata as well as Offshore and Deep strata in South Florida (Table 11.3.9; Figure 11.3.27) and in association with large habitat features of both medium and high relief (Table 11.3.10; Figure 11.3.28). A total of 130 *Seriola dumerili* were measured with stereo video; individuals ranged from 220 – 1000 mm FL, although individuals ranging from 420 – 480 mm FL were most commonly observed (Figure 11.3.29).

From spherical-camera artificial hardbottom video reads, *Seriola dumerili* were observed on 16, or 39%, of all videos read (Figure 11.3.30). When examining video footage prior to standard video read time, an additional 2 instances of *Seriola dumerili* were identified. Unidentified *Seriola* spp. were present on 13 spherical camera reads (Figure 11.3.31), with an additional 4 instances identified when examining video footage prior to standard video read times.

Survey accomplishments: Survey of unconsolidated / uncharacterized habitats (C-BASS)

A total of 42 towed C-BASS transects were conducted on unconsolidated sediment habitats in Florida Gulf waters (Table 11.3.5; Figure 11.3.32). These transects covered a total of 510,000 m (average distance per transect of 12,148 ± 1,272 m) and sampled a total of 6,749,148 m2 (average area per transect of 160,694 ± 53,313 m2). Among all C-BASS transects conducted in West Florida waters, *Seriola dumerili* were observed only once in association with unconsolidated sediment habitats; four individuals were observed at one Moderate Catch site in the North Central region (Figure 11.3.32).

Results: Developing calibration coefficients between single-camera and spherical-camera S-BRUV reads

From 2020 – 2023, a total of 648 paired single-camera and spherical-camera reads were completed for sites sampled in Florida Gulf waters, of which *Seriola dumerili* were observed on at least one read for 90 total sites (Figure 11.3.33). For most sites, spherical-camera reads observed more *Seriola dumerili* than did single-camera reads, although this is not universal (Figure 11.3.34). Of the 90 sites in which *Seriola dumerili* were observed, they were detected on single-camera reads on 67 sites, for a single-camera detection probability of 74%. Spherical camera counts, along with 68% prediction intervals, were generated for sites where only single-camera reads were calculated (Figure 11.3.35) based on results from Bayesian Poisson regression described in Chapter 10.2.

Results: Comparison between baited and unbaited S-BRUV deployments

A total of 104 paired baited – unbaited sites were sampled, including 46 artificial reef sites and 58 natural reef sites (Figure 11.3.36). PERMANOVA results indicated that, after accounting for the random site effect, overall assemblage structure differed significantly between habitats and bait treatments, but the interaction between habitat and bait was not significant (Table 11.3.11). In terms of *Seriola dumerili*, abundances were much higher on artificial than natural hardbottom habitats (Figure 11.3.37); for natural habitats, abundances were higher on baited sets, while for artificial habitats, abundances were higher on unbaited sets.

Results: Estimation of absolute abundance of Greater Amberjack – camera-based estimate

Site-specific abundance

We initially compared Greater Amberjack abundances between standard S-BRUV reads and those obtained prior to standard S-BRUV reads (including the camera drop through the water column), to determine whether S-BRUV reads on the seafloor may have been biased by excluding individuals in the water column (Figure 11.3.38). For sites where both reads were completed (N = 246), Greater Amberjack were only observed on 41 sites. Of these, there was no clear relationship between the two measures of Greater Amberjack abundance; Greater Amberjack were only observed on standard S-BRUV reads at 14 sites, while at 10 sites they were only observed during the read conducted prior to the standard S-BRUV read. Where Greater Amberjack were observed on both reads (N = 17), higher abundances were observed during the standard S-BRUV read 10 times. As there was no clear bias of the standard S-BRUV counts, and neither count correlated well with acoustic estimates of Greater Amberjack (Figure 11.3.39), we decided to move forward with standard S-BRUV counts only for consistency.

Frequency distribution of Greater Amberjack site-specific abundances on natural and artificial hardbottom habitats that are fed into statistical analyses are presented in Figure 11.3.40. All site-specific abundances represent spherical camera currency, where single-camera counts were converted to spherical-camera counts (rounding down to the nearest whole integer) using results from Bayesian Poisson regression analyses to define camera calibration factors. For natural hardbottom habitats, counts ranged from 0 – 74 Greater Amberjack per site, with 805 sites (86%) where Greater Amberjack were not observed. For artificial hardbottom habitats, counts ranged from 0 – 42 Greater Amberjack per site, with 135 sites (69%) where Greater Amberjack were not observed.

Effective area of reef habitat sampled

The frequency distribution of the effective area of reef habitat that fell within the 108 m buffer around natural and artificial hardbottom sampling sites is presented in Figure 11.3.41. For natural reef sampling sites, the effective area of natural reef habitat in proximity to the S-BRUV array ranged from 0 to 36,643 m2, with an average of 9,239 m2. For artificial reef sampling sites, the effective area of artificial reef habitat in proximity to the S-BRUV array ranged from 0 to 19,504 m2, with an average of 1,429 m2.

Estimating site selection probabilities

Site selection probabilities for natural and artificial hardbottom sampling strata are provided in Tables 11.3.2 and 11.3.4, respectively.

Estimating natural and artificial habitat coverage

Proportion of mapped area comprised of natural hard bottom habitat varied by statistical zone with the lowest proportion in Zone 3 (west of Florida Bay) with 0.0036 and greatest in Zone 7 (Big Bend) with 0.0779 (Table 11.3.12). Variation in the proportion of the mapped area containing hard bottom habitat as well as the overall size of the statistical zone generated large differences in extrapolated habitat area. In general, the central west Florida shelf had the greatest amount of hard bottom including zone 5 (879.9 km2), zone 6 (912.4 km2) and zone 7 (880.3 km2). Zones in the Florida Panhandle (9 and 10) had less natural hard bottom habitat overall (Table 3.2). Extrapolated area of natural hard bottom for the entire region was 3,979.025 km2 which is similar to prior approximations (Keenan et al. 2022). Corresponding estimates of artificial hardbottom (15.68 km2) and unconsolidated sediment habitats (142,466.27 km2) were also generated.

Estimating Greater Amberjack abundance

Camera-based estimates of abundance of *Seriola dumerili* were made using the stratified ratio estimator with area-based sampling units. The count values used for abundance were observed MaxN spherical camera counts on sampled sites where available, and the mean of the model-based spherical camera counts (as described in Chapter 9.2 and displayed in Figure 11.3.35) on the sites where only single camera MaxN values were observed.

The sample base weights used in the estimator were the reciprocal of the product of two sampling probabilities. The first was the mapping probability; i.e. the proportion of the grid cells mapped in each region and depth area by the G-FISHER program. The second was the microgrid site selection probability for each sampled grid cell, shown in Tables 11.3.2 and 11.3.4 for natural and artificial habitats. These weights were adjusted for non-response using weighting class adjustment, where the weighting classes were the cells of the 3 x 3 habitat stratification to which the missing observation belonged.

Estimates and their standard errors were computed using the R package survey (10.32614/CRAN.package.survey), which appropriately handles variance estimation for complex sample designs.

Estimates of abundance for natural and artificial habitat from the camera count data from the use of buffers of 108 m, 50 m, and 10 m are shown in Tables 11.3.13 – 11.3.15. Each table displays the estimates aggregated to the spatial stratum level, as well as the overall estimate of *Seriola dumerili* for natural and artificial habitats, respectively. As the buffer area (i.e., effective sample area) decreased, this site specific density increased and therefore increased total abundance. The estimate using the smallest buffer, which is the typical visual distance of the S-BRUV cameras, generated unrealistically high estimates (Table 11.3.15). The estimate using the 50m buffer generated most realistic results and was therefore used for final camera-based estimates for natural and artificial habitats (Table 11.3.14).

Results: Estimation of absolute abundance of Greater Amberjack – C-BASS estimate

Densities of Greater Amberjack observed in the C-BASS towed camera survey were calculated when Amberjack were observed over unconsolidated sediment. Positive observations were only observed during one 15 sec video segment in the North Central region (Figure 11.3.32). Utilizing density estimate from this observation and extrapolating to unconsolidated sediment areas for the region resulted in 37,482 ± 42,402.5 estimated Greater Amberjack.

Results: Estimation of absolute abundance of Greater Amberjack – acoustics-based estimate

Acoustic-based estimates of potential Greater Amberjack abundance

Frequency distribution of site-specific acoustic counts of large-bodied, disaggregated, water column fish (LBWCF, i.e. potential Greater Amberjack) on natural and artificial hardbottom habitats are presented in Figure 11.3.42. For natural hardbottom habitats, acoustic counts ranged from 0 – 442 LBWCF per site, with an average of 8.5 per site. For artificial hardbottom habitats, counts ranged from 0 – 44 LBWCF per site, with an average of 1.6 per site. Agreement between paired S-BRUV and acoustic surveys, in terms of presence/absence of Greater Amberjack (S-BRUV) compared to presence/absence of LBWCF (acoustics), was generally weak (Figure 11.3.39; Table 11.3.16). For both natural and artificial reef habitats, only seventeen instances occurred where Greater Amberjack were observed on video and LBWCF were also detected on acoustics. While it may not be particularly surprising that LBWCF were observed by acoustics when no Greater Amberjack were observed on the cameras, the high incidence of cases when Greater Amberjack were seen by the cameras when no LBWCF were counted

by the acoustic survey is problematic. In particular, the low number of sites were both camera counts of Greater Amberjack were greater than zero and hydroacoustic counts of LBWCF were greater than zero makes it impossible to use a site-specific camera-based species ratio correction for the acoustic counts.

Acoustics-based estimates of Greater Amberjack associated with natural and artificial reef habitats are provided in Table 11.3.17. Abundance estimates are provided for each spatial stratum. For natural habitats, abundance estimates were greatest for BBO strata (n=281,524) with lower estimates from BBN, BBO, NCO, SFO and NCN. Estimates of zero resulted in NCD, SFN and SFD; these strata contained no Greater Amberjack observed on any S-BRUV/acoustic pair, therefore ratios applied to acoustic data resulted in zero estimates. For artificial reef habitats, estimates were greatest for SFN stratum (n=83,391) followed by BBN (n=63,964) with positive values also from NCN, NCO, NCD, BBO and SFD (Table 11.3.17). Estimates of zero came from BBD and SFO as no Greater Amberjack were observed from S-BRUVs paired with acoustics. Final acoustic-based estimates were 328,165 Greater Amberjack from Natural habitats and 188,753 from artificial reef habitats.

Discussion

We produced a design-based estimate of age-1 and older Greater Amberjack in the eastern Gulf of America (Gulf) by integrating camera and habitat mapping data from two complementary surveys. Our abundance estimation process was facilitated greatly by the availability of data from the G-FISHER program, an ongoing broad-scale, fishery-independent survey whose data was used to provide abundance estimates of Greater Amberjack on both natural and artificial hardbottom habitats as well as estimate habitat coverage throughout the West Florida region. Data from G-FISHER were valuable for several key regions. First, the G-FISHER survey utilizes a statistically rigorous survey design to partition and allocate a robust amount of annual sampling effort among biologically meaningful spatial and habitat strata. Second, G-FISHER provide abundance and size composition data using time-tested stereo-baited remote underwater video (S-BRUV) arrays that typically exhibit less species and size selectivity than traditional capture gears. And third, the G-FISHER sampling frame was developed using more than a decade of randomized habitat mapping data that were used to provide (1) fine-scale estimates of proximate reef habitat surrounding each gear deployment, and (2) broad-scale estimates of regional habitat coverage. This survey was complemented by a towed camera survey designed to provide broad coverage on areas of presumed unconsolidated sediment, or non-reef habitat. Combined, data from these two surveys were used to provide Greater Amberjack abundance estimates on natural hardbottom habitats, artificial hardbottom habitats, and non-reef habitats, respectively.

Our approach to camera-based abundance estimation relied on some key assumptions, several of which were based on experience and intuition rather than empirical data. However, we generally aimed to err conservatively with our decision-making process, so that our estimates represent a worst-case scenario (i.e., Greater Amberjack abundance is likely higher than our estimates would indicate). First, we acknowledge that there is some uncertainty in the ability of video readers to accurately identify *Seriola* spp. on video, especially at small sizes (Figure 11.3.43). To compensate for this, our video reading protocols require that video readers only identify individuals to species if they are certain of their identification, and all videos are reviewed by a minimum of two expert readers. Should there be any uncertainty, observed individuals would be identified to genus only (*Seriola* spp.), and these counts were excluded from analyses. Second, the choice of our abundance metric, MaxN, is in itself conservative. Some evidence suggests that MaxN may underestimate true abundance, especially among schooling species. Nevertheless, the use of MaxN where all individuals were observed on a single screen shot ensures that double counting is not possible. Even when MaxN was converted to spherical camera currency through calibration equations, these conversions were based on hundreds of paired

observations. Third, Greater Amberjack are often found great distances off the seafloor that would be well above the field of view of stationary S-BRUV arrays. We attempted to assess whether we could identify or define potential biases in standard S-BRUV reads through the use of both early video reads (as the camera passed through the water column) as well as paired hydroacoustic surveys at a subset of sampling sites. Data from these comparisons were highly variable and did not indicate the presence of significant or persistent bias in the standard S-BRUV counts. Therefore, we chose to use standard S-BRUV counts for consistency, understanding that site-specific detection probability of Greater Amberjack is imperfect. And fourth, instead of using the area viewed by the camera in our density estimates (~ 10 m radius under ideal conditions), we initially estimated a conservative and much greater effective area of sampling (~ 108 m radius). Greater Amberjack are known to be reactive to sampling vehicles moving through the water, and our pilot study indicated at least a modest influence of bait attraction, especially on natural hardbottom sites. Therefore, our initial estimates represent the likely the furthest distance we would expect Greater Amberjack to be drawn from. Results from sensitivity analyses of buffers of both 10 m and 50 m indicate that choice of effective sampling area has significant implications to final abundance estimates. Admittedly these assumptions come with a certain degree of uncertainty that cannot be directly quantified by the current study, and the highly variable nature of Greater Amberjack abundance and site-specific occupancy certainly complicates these efforts. However, we feel the abundance estimation framework we developed is robust and reproducible, and with a series of focused studies designed to test or quantify these assumptions, could have broad applicability for other managed reef fishes in the region.

Acoustic data were not collected with every S-BRUV dataset, therefore to generate region-wide abundance estimate, we were required to aggregate camera data across broader habitat strata. To provide acoustic-based estimates for both natural and artificial habitats within each spatial strata, we collapsed data over S-BRUV survey's habitat strata of reef area (large, medium, small) and relief categories (high, mid, low). Overall abundance estimates were of a similar order of magnitude to those generates from camera-based process, although there were some spatial and habitat-specific differences observed (Tables 11.3.14 and 11.3.17). Artificial reef habitats received a greater proportion of the total estimate from the acoustic-based approach than the camera based, which could be attributed to smaller scale habitat features typical of artificial reefs. Acoustic results indicate that the acoustics-based estimation protocols explored here show promise. However, there are several sources of uncertainty that require further consideration. These predominantly relate to the use of camera data to partition acoustics-based estimates of large-bodied fishes to species. Our approach was to include all species that routinely reach a size greater than 200 mm in our partitioning estimates. This also includes several species that spend most of their time on the bottom and likely within the acoustics shadow, therefore might not have been observed in the acoustics survey. We do not have positional data for individuals observed on video, so we cannot determine which individuals would not be observed on acoustics, but including these species may have biased our proportions low. In addition, although we excluded small Blue Runner and Vermilion Snapper from our estimates of large-bodied reef fishes, there are other species that might also be observed at small sizes. Further efforts to better partition the camera-based data would improve our ability to subsequently partition the acoustics data. Another source of uncertainty includes the estimates of effective area sampled by the acoustics survey. For our approach, rather than the total area scanned by the acoustic equipment, we subset this area by the proportion of mapped hard-bottom habitat within the 108m buffer associated with the S-BRUV camera set, paired with that specific acoustic survey. The actual amount of hard bottom surveyed by the acoustic survey may have been larger or smaller; which ultimately could skew the density estimate from the specific site.

Tables and Figures

Table 11.3.1. Complete list of natural hard bottom types (geoforms) identifiable via interpretation of side-scan sonar backscatter data and included in the G-FISHER survey. Relief strata indicates whether each geoform was included in the low relief, moderate relief, or high relief in the G-FISHER habitat stratification scheme.

Natural hardbottom geoform	Relief strata
Seagrass	Low relief
Flat hard bottom	Low relief
Pavement	Low relief
Mixed hard bottom	Low relief
Potholes	Moderate relief
Rubble field	Moderate relief
Fracture	Moderate relief
Reef rubble	High relief
Fragmented hard bottom	High relief
Individual patch reef	High relief
Aggregate patch reef	High relief
Aggregate coral reef	High relief
Ledge	High relief
Boulder	High relief
Spur and groove	High relief
Spring or sink	High relief
Pinnacle	High relief
Escarpment	High relief

Table 11.3.2. Summary of the G-FISHER survey design for natural hardbottom habitats. Spatial strata are a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep). Habitat strata are a combination of hardbottom origin (Natural), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large). The total number of available reef features and annual effort allocation are reported for each stratum. Selection probability represents the probability of selection of primary sampling units (microgrids). The last two columns summarize successful sampling effort (sites sampled with usable video data), including those sites read for single-camera abundance and spherical-camera abundance, in 2023. The G-FISHER survey design does not differentiate between the MS/AL and Florida portion of the North Central Gulf region, but only sites off the coast of Florida were included in effort summaries and subsequent analyses.

Spatial stratum	Habitat stratum	Total mapped microgrids	Annual allocation	Selection Probability	Single-camera sites	Spherical-camera sites (acoustics)
NCN	NLS	105	7	0.067	7	0
NCN	NLM	430	22	0.051	21	3
NCN	NLL	671	30	0.045	31	1
NCN	NMS	23	3	0.130	3	0
NCN	NMM	3	2	0.667	2	0
NCN	NML	0	0		0	0
NCN	NHS	43	3	0.070	3	0
NCN	NHM	112	19	0.170	18	2
NCN	NHL	75	14	0.187	12	0
NCO	NLS	526	3	0.006	3	1
NCO	NLM	1927	16	0.008	12	0
NCO	NLL	2615	29	0.011	28	4
NCO	NMS	295	5	0.017	3	0
NCO	NMM	81	4	0.049	3	0
NCO	NML	0	0		0	0
NCO	NHS	158	4	0.025	4	0
NCO	NHM	351	15	0.043	12	2
NCO	NHL	394	31	0.079	30	2
NCD	NLS	247	5	0.020	5	0
NCD	NLM	684	12	0.018	11	0
NCD	NLL	1460	29	0.020	26	0
NCD	NMS	589	6	0.010	5	0
NCD	NMM	200	6	0.030	5	0
NCD	NML	26	8	0.308	5	0
NCD	NHS	260	10	0.038	10	0
NCD	NHM	516	18	0.035	18	0
NCD	NHL	976	30	0.031	28	1
BBN	NLS	1814	5	0.003	3	2
BBN	NLM	5711	14	0.002	13	6
BBN	NLL	8945	30	0.003	27	11
BBN	NMS	15	3	0.200	3	1

Spatial stratum	Habitat stratum	Total mapped microgrids	Annual allocation	Selection Probability	Single-camera sites	Spherical-camera sites (acoustics)
BBN	NMM	6	3	0.500	2	1
BBN	NML	2	1	0.500	1	1
BBN	NHS	526	20	0.038	16	3
BBN	NHM	1794	17	0.009	15	4
BBN	NHL	1992	30	0.015	28	14
BBO	NLS	1813	13	0.007	13	0
BBO	NLM	3872	25	0.006	24	8
BBO	NLL	4181	53	0.013	49	14
BBO	NMS	264	12	0.045	12	4
BBO	NMM	184	7	0.038	6	2
BBO	NML	2	1	0.500	1	0
BBO	NHS	86	2	0.023	2	0
BBO	NHM	369	21	0.057	18	4
BBO	NHL	522	33	0.063	30	11
BBD	NLS	905	11	0.012	9	5
BBD	NLM	1217	16	0.013	9	4
BBD	NLL	2264	36	0.016	30	11
BBD	NMS	2170	12	0.006	3	1
BBD	NMM	1575	9	0.006	3	1
BBD	NML	47	5	0.106	5	1
BBD	NHS	679	8	0.012	2	1
BBD	NHM	999	11	0.011	2	2
BBD	NHL	497	18	0.036	10	2
SFN	NLS	269	7	0.026	7	0
SFN	NLM	812	15	0.018	12	5
SFN	NLL	1193	32	0.027	29	5
SFN	NMS	9	3	0.333	2	0
SFN	NMM	0	0		0	0
SFN	NML	0	0		0	0
SFN	NHS	118	3	0.025	3	1
SFN	NHM	335	17	0.051	10	4
SFN	NHL	1379	35	0.025	27	14
SFO	NLS	1492	12	0.008	12	4
SFO	NLM	2227	23	0.010	23	7
SFO	NLL	1374	31	0.023	27	5
SFO	NMS	363	7	0.019	7	2
SFO	NMM	241	9	0.037	9	1
SFO	NML	1	1	1.000	1	0
SFO	NHS	99	9	0.091	7	2
SFO	NHM	227	16	0.070	16	6
SFO	NHL	561	34	0.061	32	9

Spatial stratum	Habitat stratum	Total mapped microgrids	Annual allocation	Selection Probability	Single-camera sites	Spherical-camera sites (acoustics)
SFD	NLS	1713	7	0.004	6	0
SFD	NLM	1539	11	0.007	10	1
SFD	NLL	970	15	0.015	17	4
SFD	NMS	4444	7	0.002	7	0
SFD	NMM	3059	8	0.003	6	2
SFD	NML	60	5	0.083	3	0
SFD	NHS	563	6	0.011	5	0
SFD	NHM	494	11	0.022	8	0
SFD	NHL	272	18	0.066	7	0

Table 11.3.3. Complete list of artificial hardbottom types (geoforms) identifiable via interpretation of side-scan sonar backscatter data and included in the G-FISHER survey. Relief strata indicates whether each geoform was included in the low relief, moderate relief, or high relief in the G-FISHER habitat stratification scheme.

Artificial hardbottom geoform	Relief strata
Cable	Low relief
Dredge deposit	Low relief
Tires	Low relief
Dredged channel	Low relief
Rock piles	Low relief
Pipelines	Low relief
Construction materials	Moderate relief
Reef modules	Moderate relief
Chicken coops	High relief
Marine wreckage	High relief
Other vehicles	High relief
Small vessels	High relief
Military tanks	High relief
Oil-platform materials	High relief
Aircraft	High relief
Large vessels	High relief

Table 11.3.4. Summary of the G-FISHER survey design for artificial hardbottom habitats. Spatial strata are a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep). Habitat strata are a combination of hardbottom origin (Artificial), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large). The total number of available reef features and annual effort allocation are reported for each stratum. Selection probability represents the probability of selection of primary sampling units (microgrids). The last two columns summarize successful sampling effort (sites sampled with usable video data), including those sites read for single-camera abundance and spherical-camera abundance, in 2023. The G-FISHER survey design does not differentiate between the MS/AL and Florida portion of the North Central Gulf region, but only sites off the coast of Florida were included in effort summaries and subsequent analyses.

Spatial	Habitat	Total	Annual	Selection	Single-	Spherical-
NCN	ALS	72	3	0.042	2	0
NCN	ALM	45	2	0.044	2	0
NCN	ALL	47	2	0.043	1	0
NCN	AMS	289	5	0.017	3	1
NCN	AMM	196	5	0.026	5	1
NCN	AML	179	8	0.045	5	1
NCN	AHS	317	3	0.009	3	0
NCN	AHM	94	4	0.043	2	0
NCN	AHL	175	6	0.034	2	2
NCO	ALS	226	2	0.009	2	0
NCO	ALM	325	3	0.009	3	2
NCO	ALL	106	2	0.019	1	0
NCO	AMS	381	7	0.018	5	2
NCO	AMM	563	7	0.012	6	1
NCO	AML	281	8	0.028	8	3
NCO	AHS	934	8	0.009	1	0
NCO	AHM	276	7	0.025	6	2
NCO	AHL	165	7	0.042	6	1
NCD	ALS	105	2	0.019	3	0
NCD	ALM	203	2	0.010	2	0
NCD	ALL	114	5	0.044	3	0
NCD	AMS	41	5	0.122	4	1
NCD	AMM	51	2	0.039	2	0
NCD	AML	30	4	0.133	3	1
NCD	AHS	142	9	0.063	7	1
NCD	AHM	120	5	0.042	3	0
NCD	AHL	113	4	0.035	2	0
BBN	ALS	28	2	0.071	2	0
BBN	ALM	29	2	0.069	2	0
BBN	ALL	34	5	0.147	5	2
BBN	AMS	153	3	0.020	2	0
BBN	AMM	154	4	0.026	4	0
BBN	AML	159	7	0.044	4	2

Spatial	Habitat	Total	Annual	Selection	Single-	Spherical-
BBN	AHS	24	2	0.083	2	0
BBN	AHM	28	2	0.071	1	0
BBN	AHL	52	6	0.115	5	1
BBO	ALS	2	2	1.000	2	2
BBO	ALM	5	2	0.400	2	0
BBO	ALL	122	12	0.098	9	3
BBO	AMS	1	0	0.000	0	0
BBO	AMM	3	2	0.667	2	1
BBO	AML	0	0		0	0
BBO	AHS	12	2	0.167	2	0
BBO	AHM	13	5	0.385	4	2
BBO	AHL	13	6	0.462	5	1
BBD	ALS	1	0	0.000	0	0
BBD	ALM	0	0		0	0
BBD	ALL	17	7	0.412	6	1
BBD	AMS	0	0		0	0
BBD	AMM	2	0	1.000	0	0
BBD	AML	1	0	0.000	0	0
BBD	AHS	0	0		0	0
BBD	AHM	3	1	0.333	0	0
BBD	AHL	5	3	0.600	2	0
SFN	ALS	10	3	0.300	3	0
SFN	ALM	6	2	0.333	1	0
SFN	ALL	12	2	0.167	2	0
SFN	AMS	31	2	0.065	1	1
SFN	AMM	28	4	0.143	3	1
SFN	AML	35	4	0.114	3	0
SFN	AHS	11	2	0.182	2	0
SFN	AHM	5	2	0.400	2	1
SFN	AHL	50	2	0.040	2	0
SFO	ALS	4	2	0.500	2	0
SFO	ALM	3	2	0.667	2	1
SFO	ALL	13	2	0.154	2	0
SFO	AMS	2	1	0.500	1	0
SFO	AMM	1	1	1.000	1	0
SFO	AML	1	1	1.000	1	0
SFO	AHS	1	0		0	0
SFO	AHM	3	2	0.667	2	0
SFO	AHL	15	3	0.200	2	0
SFD	ALS	0	0		0	0
SFD	ALM	12	2	0.167	2	1
SFD	ALL	1	0		0	0
SFD	AMS	3	2	0.667	2	0
SFD	AMM	3	2	0.667	2	0

Spatial	Habitat	Total	Annual	Selection	Single-	Spherical-
SFD	AML	5	2	0.400	2	1
SFD	AHS	0	0		0	0
SFD	AHM	1	0		0	0
SFD	AHL	4	2	0.500	2	0

Table 11.3.5. Summary of the C-BASS survey design for unconsolidated or uncharacterized habitats. Regional strata were defined as NC: North-Central; BB: Big Bend; SF: South Florida. GAJ abundance strata were defined as NS: Never Sampled; ZC: Zero Catch; MC: Moderate Catch, HC: High Catch. The total number of available grid cells effort allocation are reported for each stratum. Selection probability represents the probability of selection by stratum. The last column represents the total number of C-BASS sites sampled in 2022-2023.

Regional stratum	GAJ abundance stratum	Total grids	Annual allocation	Selection probability	Completed sites
NC	NS	176	7	0.040	0
NC	ZC	95	4	0.042	0
NC	MC	74	6	0.081	3
NC	HC	74	12	0.132	3
BB	NS	178	7	0.039	4
BB	ZC	252	10	0.040	3
BB	MC	55	4	0.073	6
BB	HC	52	8	0.154	5
SF	NS	352	14	0.040	10
SF	ZC	203	8	0.040	2
SF	MC	43	3	0.070	2
SF	HC	43	7	0.163	4

Table 11.3.6. Summary of collapsed natural and artificial reef sampling strata (collapsed over reef area and relief) for which acoustics data were available to produce region-wide acoustics-based estimates of Greater Amberjack abundances.

Reef Habitat	Spatial Strata	Acoustic sites
Natural	NCN	5
Natural	NCO	12
Natural	NCD	2
Natural	BBN	41
Natural	BBO	42
Natural	BBD	22
Natural	SFN	30
Natural	SFO	36
Natural	SFD	7
Artificial	NCN	5
Artificial	NCO	12
Artificial	NCD	3
Artificial	BBN	6
Artificial	BBO	10
Artificial	BBD	1
Artificial	SFN	3
Artificial	SFO	1
Artificial	SFD	3

Table 11.3.7. Summary of sampling effort and total number of individuals of each *Seriola* taxon observed by spatial strata on single camera reads for natural hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

Spatial Strata (N)	*Seriola dumerili*	*Seriola fasciata*	*Seriola rivoliana*	*Seriola zonata*	*Seriola* spp.
NCN (97)	52	0	12	1	13
NCO (95)	143	0	51	12	8
NCD (113)	113	52	60	5	7
BBN (108)	5	0	1	1	6
BBO (155)	55	1	14	134	43
BBD (73)	24	12	26	0	20
SFN (90)	0	0	1	0	8
SFO (134)	44	0	8	10	9
SFD (69)	3	0	12	63	9
Totals (934)	439	65	185	226	123

Table 11.3.8. Summary of sampling effort and total number of individuals of each *Seriola* taxon observed by habitat strata on single camera reads for natural hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Habitat strata labels represent a combination of hardbottom origin (Natural), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large).

Habitat Strata (N)	*Seriola dumerili*	*Seriola fasciata*	*Seriola rivoliana*	*Seriola zonata*	*Seriola* spp.
NLS (65)	53	9	18	8	35
NLM (135)	61	4	16	34	20
NLL (264)	73	45	37	110	32
NMS (45)	9	0	7	8	11
NMM (36)	2	0	3	0	5
NML (16)	0	0	6	63	2
NHS (52)	40	1	18	1	3
NHM (117)	47	3	15	1	5
NHL (204)	154	3	65	1	10
Totals (934)	439	65	185	226	123

Table 11.3.9. Summary of sampling effort and total number of individuals of each *Seriola* taxon observed by spatial strata on single camera reads for artificial hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

Spatial Strata (N)	*Seriola dumerili*	*Seriola fasciata*	*Seriola rivoliana*	*Seriola zonata*	*Seriola* spp.
NCN (25)	12	1	9	0	3
NCO (38)	137	0	35	1	77
NCD (29)	18	0	34	3	4
BBN (27)	15	0	1	0	3
BBO (26)	38	0	44	0	37
BBD (8)	8	0	8	3	7
SFN (19)	7	0	1	0	4
SFO (13)	23	0	8	0	4
SFD (10)	39	0	14	0	8
Totals (195)	297	1	154	7	147

Table 11.3.10. Summary of sampling effort and total number of individuals of each *Seriola* taxon observed by habitat strata on single camera reads for artificial hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Habitat strata labels represent a combination of hardbottom origin (Artificial), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large).

Habitat Strata (N)	*Seriola dumerili*	*Seriola fasciata*	*Seriola rivoliana*	*Seriola zonata*	*Seriola* spp.
ALS (16)	16	0	10	0	4
ALM (16)	5	0	7	4	4
ALL (29)	27	0	17	0	7
AMS (18)	17	0	3	0	3
AMM (25)	39	0	5	0	6
AML (26)	68	0	32	0	61
AHS (17)	10	0	14	0	5
AHM (20)	8	0	23	0	22
AHL (28)	107	1	43	3	35
Totals (195)	297	1	154	7	147

Table 11.3.11. PERMANOVA results for the analysis of reef-fish assemblages based on the Bray-Curtis dissimilarity measure in relation to habitat type (Hab), bait treatment (Bait), and Site.

Source	df	MS	Pseudo-F	P	Sqrt (component of variation)	% of variation
Hab	1	22,888	6.8	0.001	15.8	17.6
Bait	1	1,747	2.8	0.01	3.3	3.7
Site(Hab)	102	4,408	7.2	0.001	43.5	48.0
Hab X Bait	1	854	1.4	0.239	2.2	2.5
Residual	102	615			24.8	27.7
Total	207					

Table 11.3.12. West Florida shelf sampling universe with estimated quantity of natural reef habitat as determined by extrapolating data from mapped areas into unmapped areas. All natural reef habitat classes were combined for these estimates. Proportion Natural reef habitat were calculated from total area mapped and total area of digitized natural hard bottom habitats.

G-FISHER	Stat Zone –	Total Area (km2)	Number side scan	Mean proportion of	Extrapolated Area – Natural	Mean proportion of	Extrapolated Area – Artificial	Unknown & unconsolidated
South Florida (SF)	2 – near	2,438.23	85	0.0790	192.6249	0.000003	0.00801	2245.602
	2 – off	2,700.78	69	0.1442	389.4956	0.000071	0.19048	2311.094
	2 -	4,744.18	158	0.0220	104.2150	0.000002	0.00852	4639.959
	3 – near	5,896.47	28	0.0057	33.6789	0.000106	0.62272	5862.167
	3 – off	9,885.13	143	0.0012	11.5982	0.000006	0.05587	9873.480
	3 -	15,010.87	212	0.0050	75.7508	0.000000	0.00344	14935.115
	4 – near	4,164.59	103	0.0249	103.6792	0.000043	0.17849	4060.735
	4 – off	7,697.36	321	0.0098	75.4880	0.000006	0.04713	7621.822
	4 -	11,972.43	318	0.0045	53.9969	0.000001	0.01731	11918.420
Big Bend (BB)	5 – near	3,665.00	195	0.1307	479.1316	0.000490	1.79668	3184.068
	5 – off	7,066.98	292	0.0184	129.9978	0.000606	4.27911	6932.705
	5 -	9,834.22	389	0.0142	139.8633	0.000083	0.81896	9693.539
	6 – near	5,421.59	74	0.0751	407.0582	0.000000	0.00000	5014.536
	6 – off	10,637.43	219	0.0447	475.6155	0.000000	0.00485	10161.811
	6 -	4,898.06	66	0.0033	16.2248	0.000063	0.30677	4881.533
	7 – near	6,515.14	142	0.1333	868.7491	0.000037	0.24113	5646.146
	7 – off	4,869.66	136	0.0200	97.4535	0.000001	0.00586	4772.198
North Central (NC)	8 – near	2,320.64	39	0.0134	31.0012	0.000057	0.13192	2289.508
	8 – off	4,527.70	180	0.0064	28.8709	0.000116	0.52362	4498.308
	8 -	5,660.65	168	0.0160	90.4124	0.000234	1.32436	5568.911
	9 – near	935.03	27	0.0157	14.6419	0.000046	0.04275	920.349
	9 – off	2,946.13	284	0.0136	39.9770	0.000089	0.26163	2905.895
	9 -	4,634.65	212	0.0079	36.7876	0.000196	0.90824	4596.959
	10 –	1,696.42	28	0.0015	2.5326	0.001509	2.55938	1691.323
	10 – off	4,035.01	177	0.0005	2.0799	0.000229	0.92524	4032.003
	10 -	2,359.98	74	0.0642	151.4882	0.000175	0.41392	2208.081
Totals		146,534.36			4,052.41		15.68	142,466.27

Table 11.3.13. Camera-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites. Estimates were generated using an estimated effective area sampled of 36,644 m2 (108 m buffer). Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

	Artificial		Natural	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
NCN	2,343	2,092.7	3,632	1,334.7
NCO	3,708	1,418.5	14,132	4,117.5
NCD	5,112	2,134.8	33,716	10,158.1
BBN	762	492.5	7,935	5,856.4
BBO	3,027	1,763.9	91,315	44,561.6
BBD	140	129.1	8,812	4107.1
SFN	62	55.7	0	0.0
SFO	806	512.8	35,480	15,370.0
SFD	456	336.8	2,247	1,378.6
Total	16,416		197,269	

Table 11.3.14. Camera-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites. Estimates were generated using an estimated effective area sampled of 7,854 m2 (50 m buffer). Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

	Artificial		Natural	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
NCN	5,601	4,425.8	10,705	3,908.5
NCO	11,021	3,651.5	44,847	12,830.3
NCD	8,956	3,534.4	122,781	37,058.6
BBN	2,033	1,388.1	28,721	21,107.4
BBO	6,757	3,855.3	299,918	144,697.2
BBD	364	332.4	31,469	14,743
SFN	183	157.2	0	0
SFO	1,247	776.4	109,098	46,691.0
SFD	690	569.8	4,543	2,882.2
Total	36,852	7,954.0	652,082	159,190.9

Table 11.3.15. Camera-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites. Estimates were generated using an estimated effective area sampled of 314 m2 (10 m buffer). Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

	Artificial		Natural	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
NCN	22,188	13,947.2	129,606	44,095.2
NCO	77,518	19,712.3	382,865	106,525.7
NCD	32,878	12,235.5	1,568,134	451,830.6
BBN	18,673	12,223.1	428,659	307,738.5
BBO	54,546	34,750.2	3,312,900	1,524,186.8
BBD	4,626	3,733.0	341,770	162,991.1
SFN	1,347	1,159.9	0	0
SFO	3,363	2,184.8	991,107	417,390.1
SFD	1,876	1290.5	22,670	14,431.9
Total	217,015		7,177,711	

Table 11.3.16. Confusion matrix comparing agreement between S-BRUV and acoustic surveys in relation to whether Greater Amberjack (GAJ) were observed on video and large-bodied (>20cmTL), water column (>0.5m off bottom), disaggregated fish (LBWCF or potential Greater Amberjack) were detected via analysis of paired acoustics data for natural and artificial hardbottom surveys.

Natural Hardbottom Surveys			
		Acoustic Survey (LBWCF)	
		Absent	Present
S-BRUV	Absent	136	47
Survey (GAJ)	Present	7	7

Artificial Hardbottom Surveys			
		Acoustic Survey (LBWCF)	
		Absent	Present
S-BRUV	Absent	12	15
Survey (GAJ)	Present	6	10

Table 11.3.17. Acoustics-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites. Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

	Artificial		Natural	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
NCN	2,586	120.0	448	408.8
NCO	4,213	3,078.9	15,726	10,873.3
NCD	21,543	26,228.1	0	0
BBN	63,964	34,562.3	26,708	15,454.8
BBO	8,887	7,817.2	281,524	222,920.6
BBD	0	0	2,830	2,110.2
SFN	83,391	66,853.2	0	0
SFO	0	0	929	710.1
SFD	169	130.5	0	0
Total	184,753	80,140.1	328,165	223,731.5

Figure 11.3.1. Map of spatial stratification of the G-FISHER survey in the northern Gulf of America.

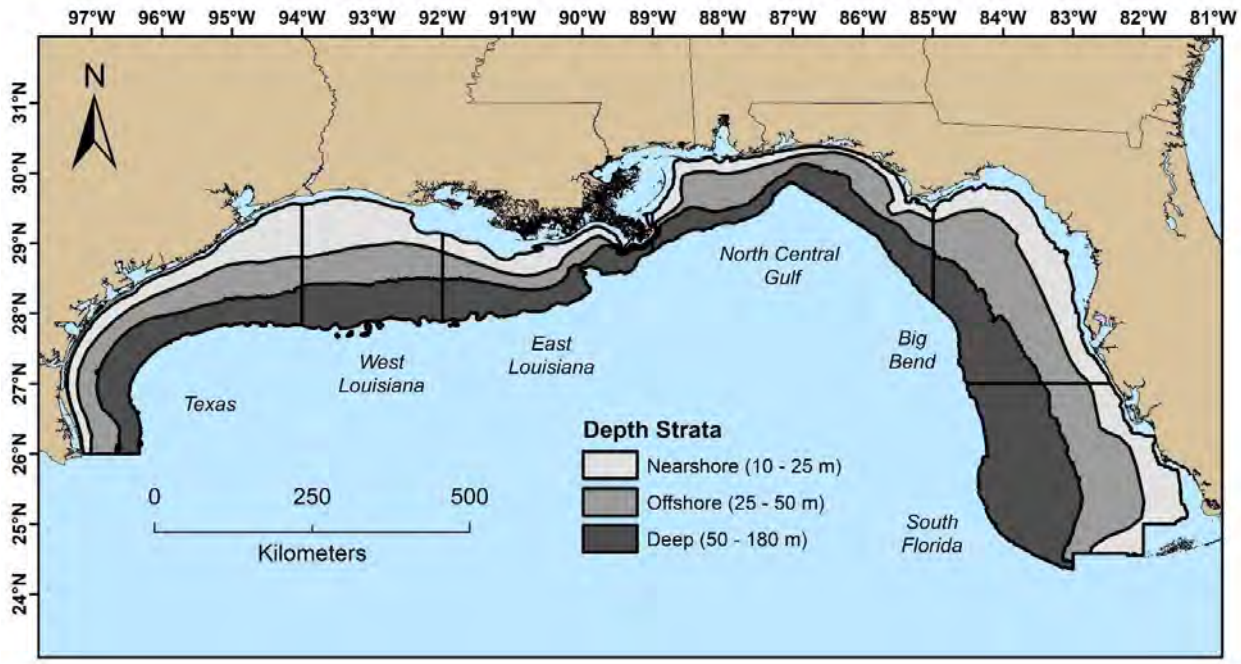


Figure 11.3.2. Example imagery showing variability of natural hardbottom habitats included within the G-FISHER survey. The G-FISHER survey utilizes a 3 x 3 stratification scheme for natural hardbottom habitats based on a combination of relative relief (Low-relief, Moderate-relief, and High-relief) and size of individual reef feature (Small scale, medium scale, large scale).



Figure 11.3.3. Optimal allocation of annual sampling effort among sampling strata for natural hardbottom surveys in the eastern Gulf of America, which was based off a combination of habitat availability and managed species richness (from Switzer et al. 2023). Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

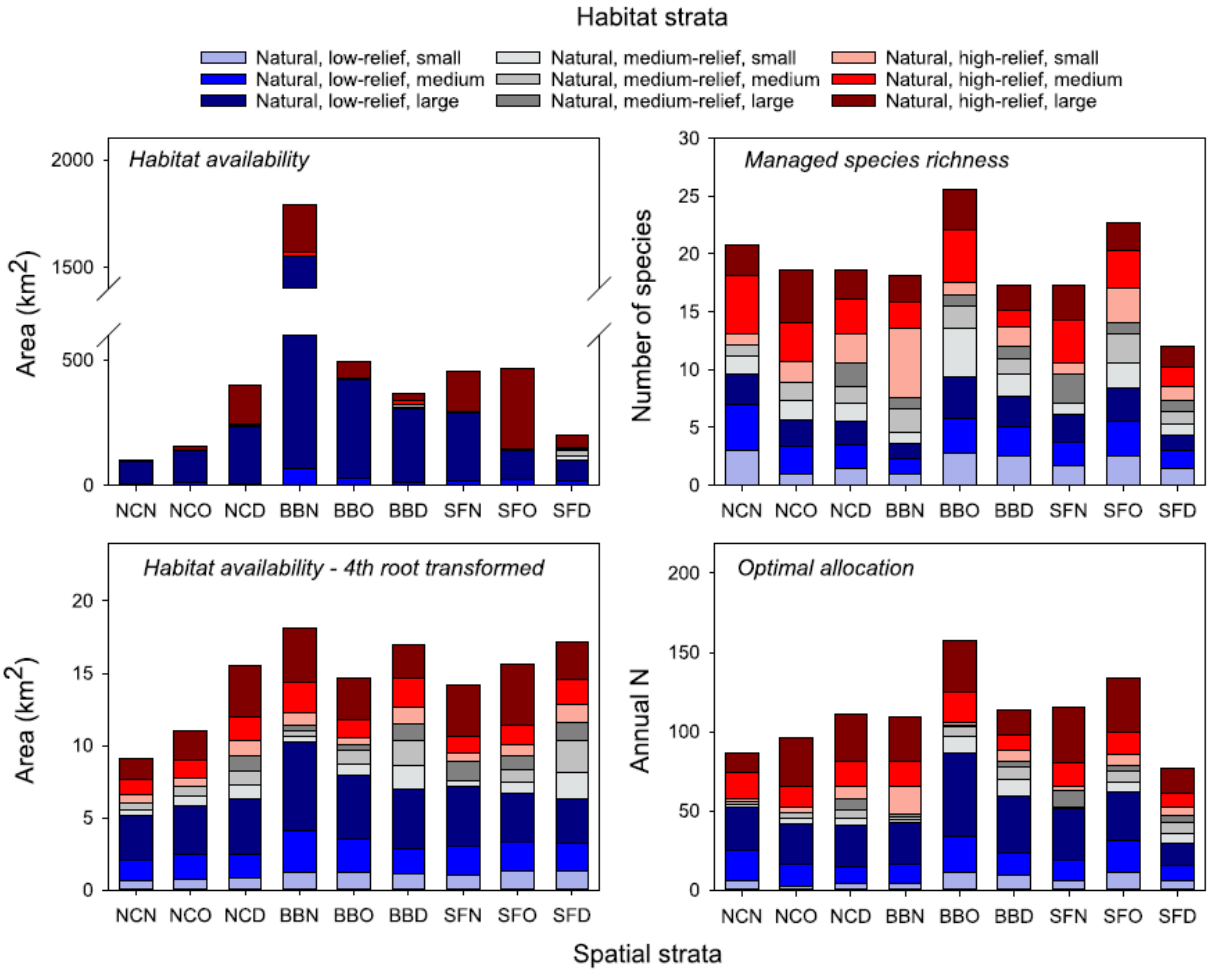


Figure 11.3.4. Example imagery showing variability of artificial hardbottom habitats included within the G-FISHER survey. The G-FISHER survey utilizes a 3 x 3 stratification scheme for artificial hardbottom habitats based on a combination of relative relief (Low-relief, Moderate-relief, and High-relief) and size of individual reef feature (Small scale, medium scale, large scale).



Figure 11.3.5. Optimal allocation of annual sampling effort among sampling strata for artificial hardbottom surveys in the eastern Gulf of America, which was based off a combination of habitat availability and managed species richness (from Switzer et al. 2023). Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

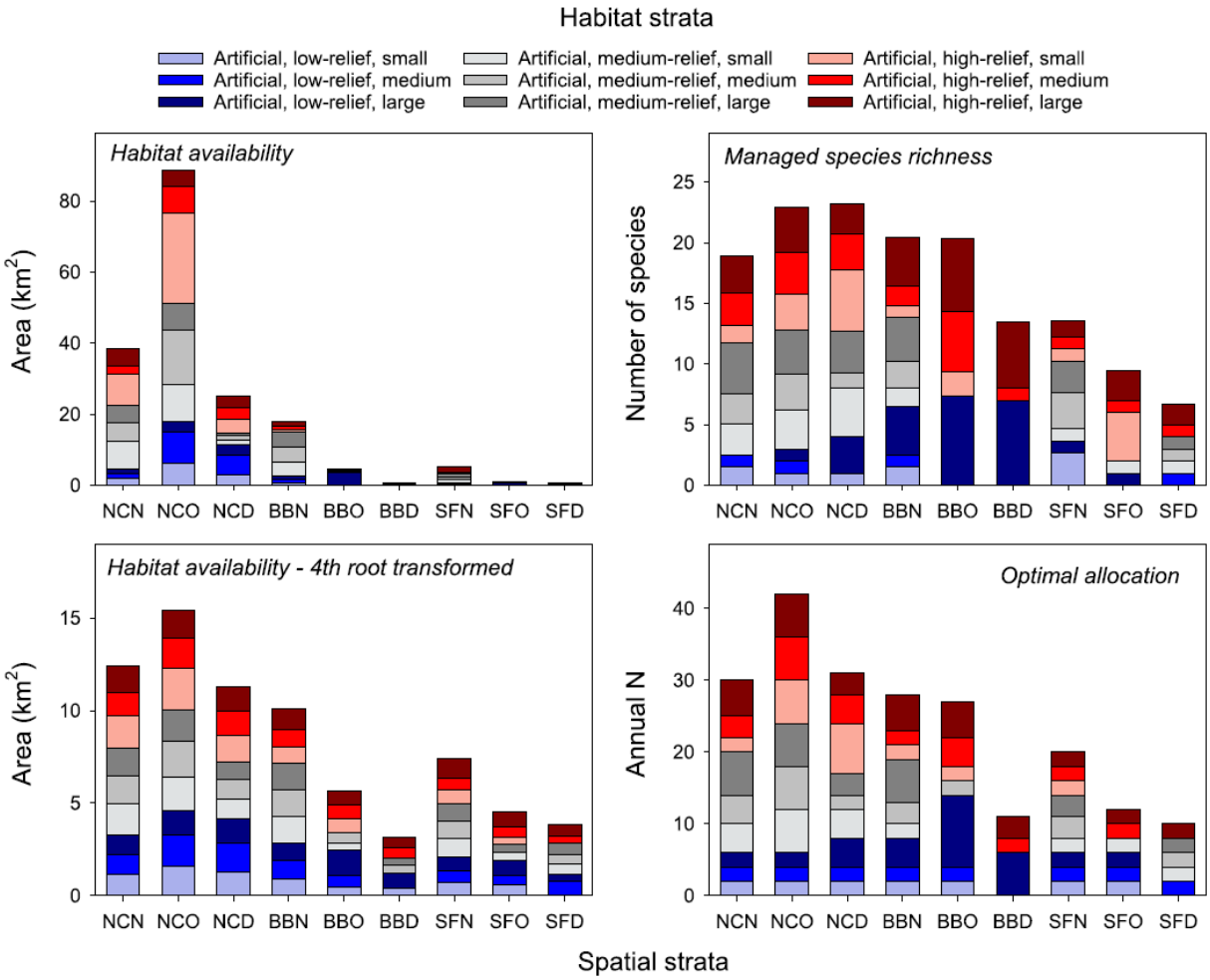


Figure 11.3.6. Spatial extent of the C-BASS survey of unconsolidated and uncharacterized habitats in the central and eastern Gulf of America. Regional spatial stratification followed that of the G-FISHER survey of natural and artificial hardbottom habitats.

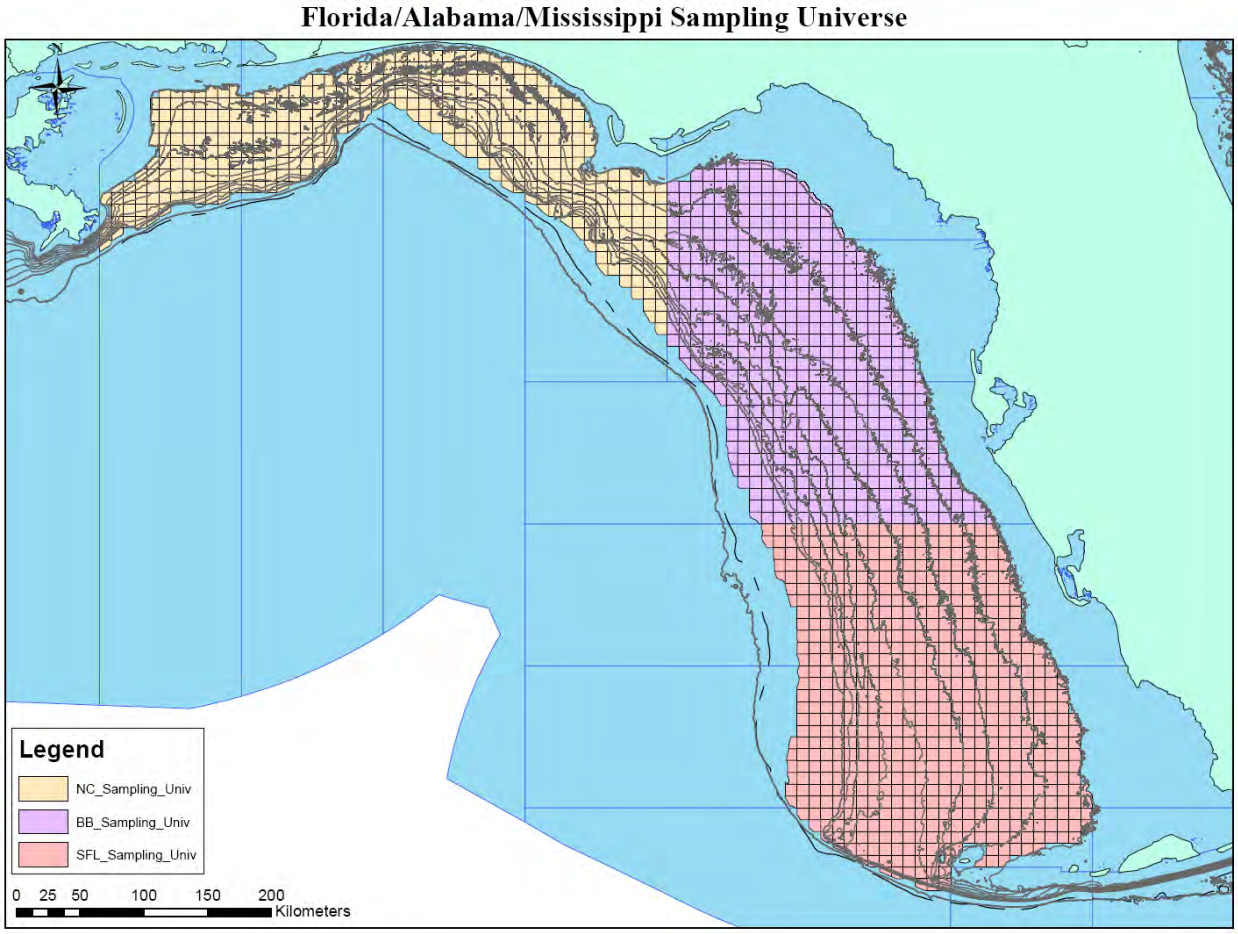


Figure 11.3.7. Map of *Seriola dumerili* observations from the FWRI and NOAA S-BRUV surveys from 2010 – 2020.

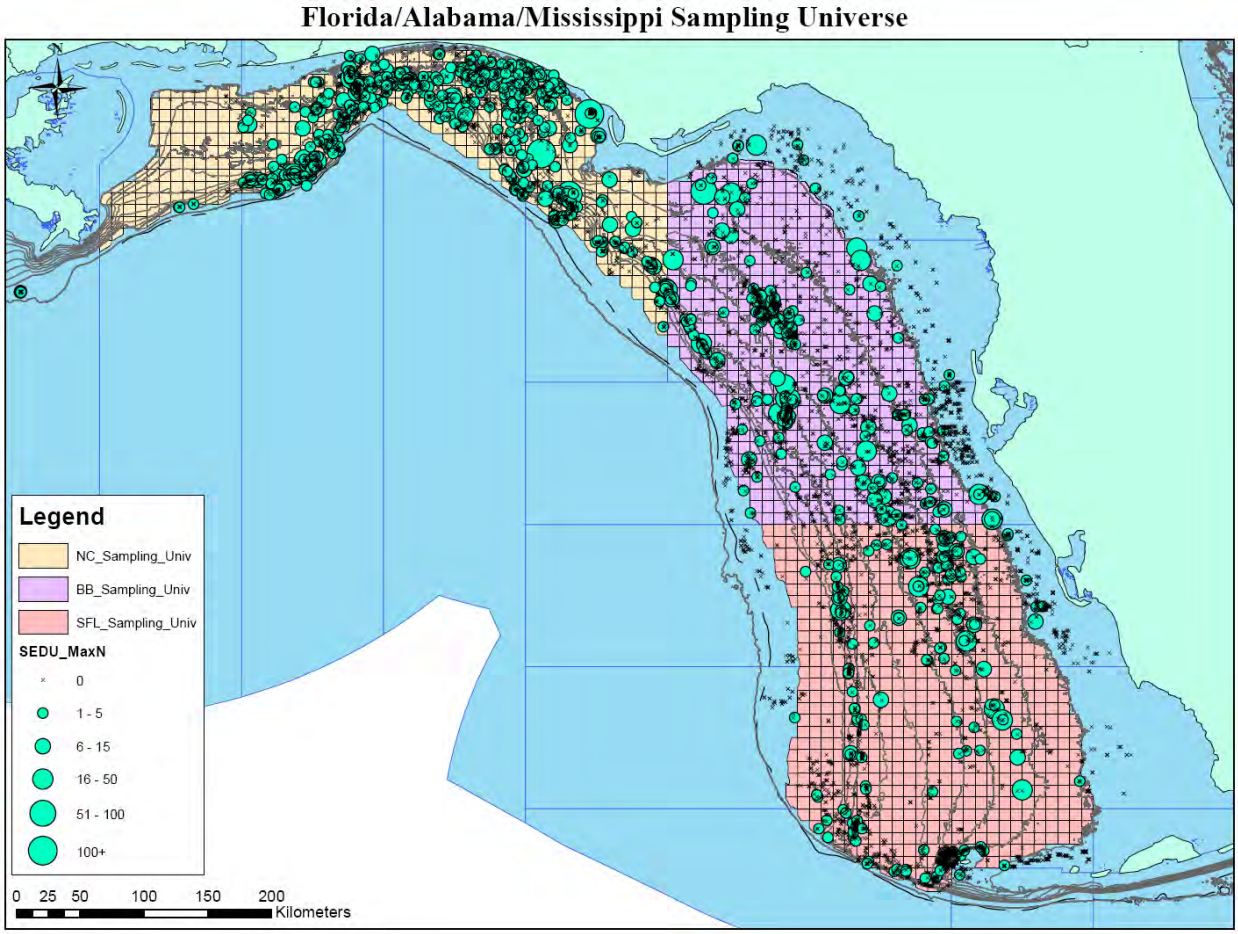


Figure 11.3.8. Map of the total number of FWRI and NOAA S-BRUV surveys conducted within each grid cell from 2010 – 2020.

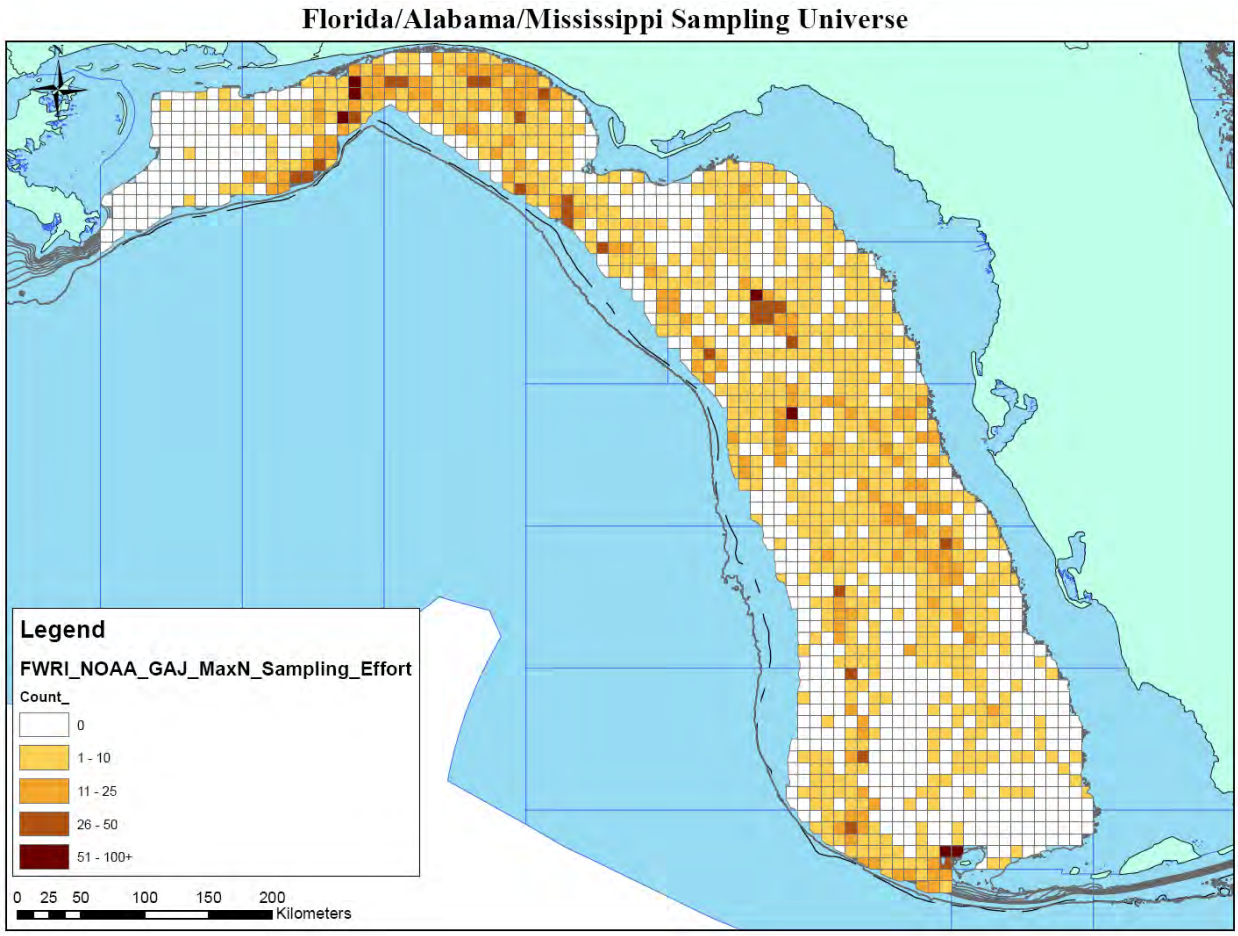


Figure 11.3.9. Map of the average number of *Seriola dumerili* observed during FWRI and NOAA S-BRUV surveys conducted within each grid cell from 2010 – 2020. Based on these results, cells were classified as Zero Effort, Zero Catch (white), Moderate Catch (blue), and High Catch (red).

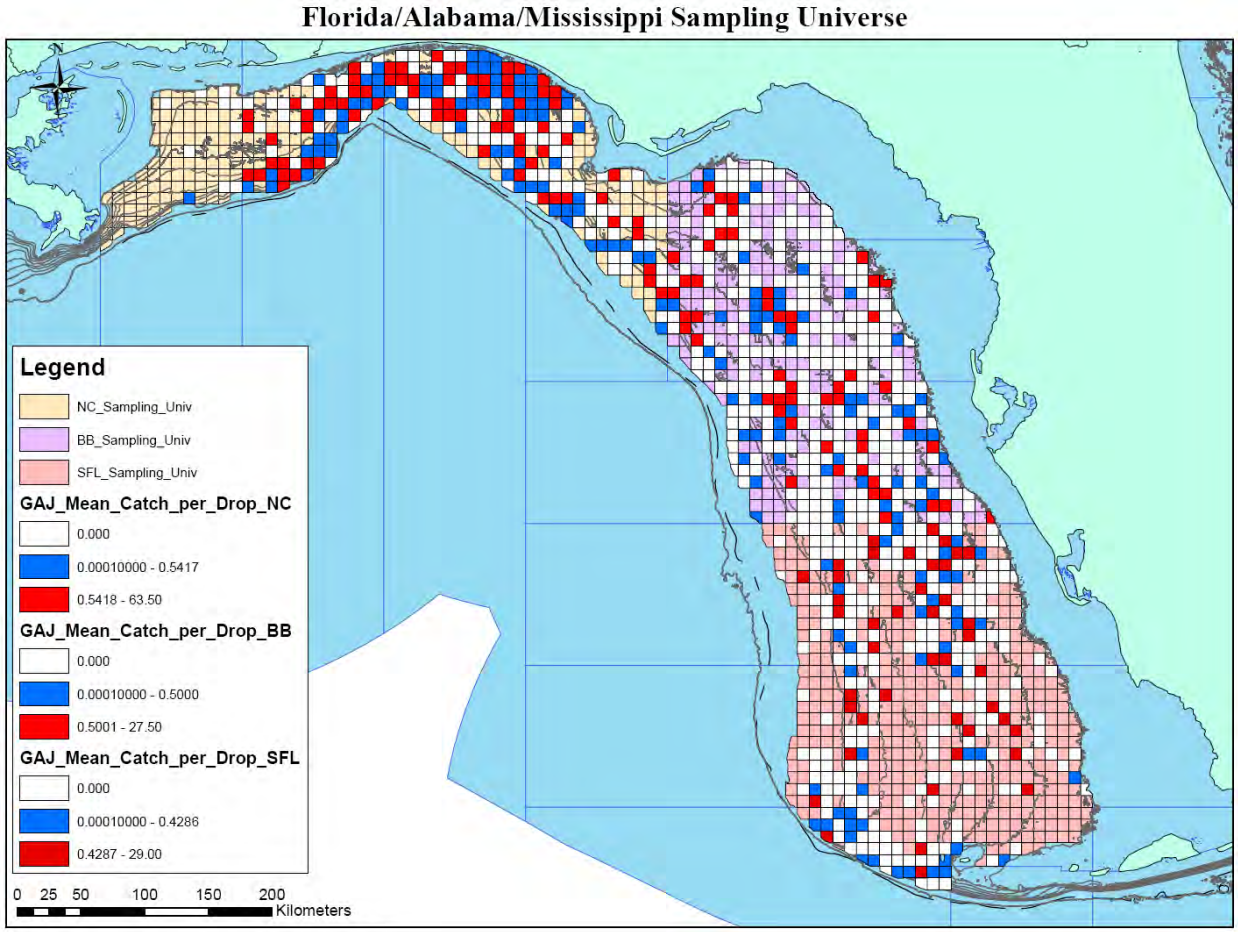


Figure 11.3.10. Map showing the distribution of 4,137 side-scan surveys conducted between 2010 and 2023. Surveys were randomly selected and allocated among National Marine Fisheries Service statistical zones and depth strata. Effort varies by statistical zone due to time differences of funding sources.

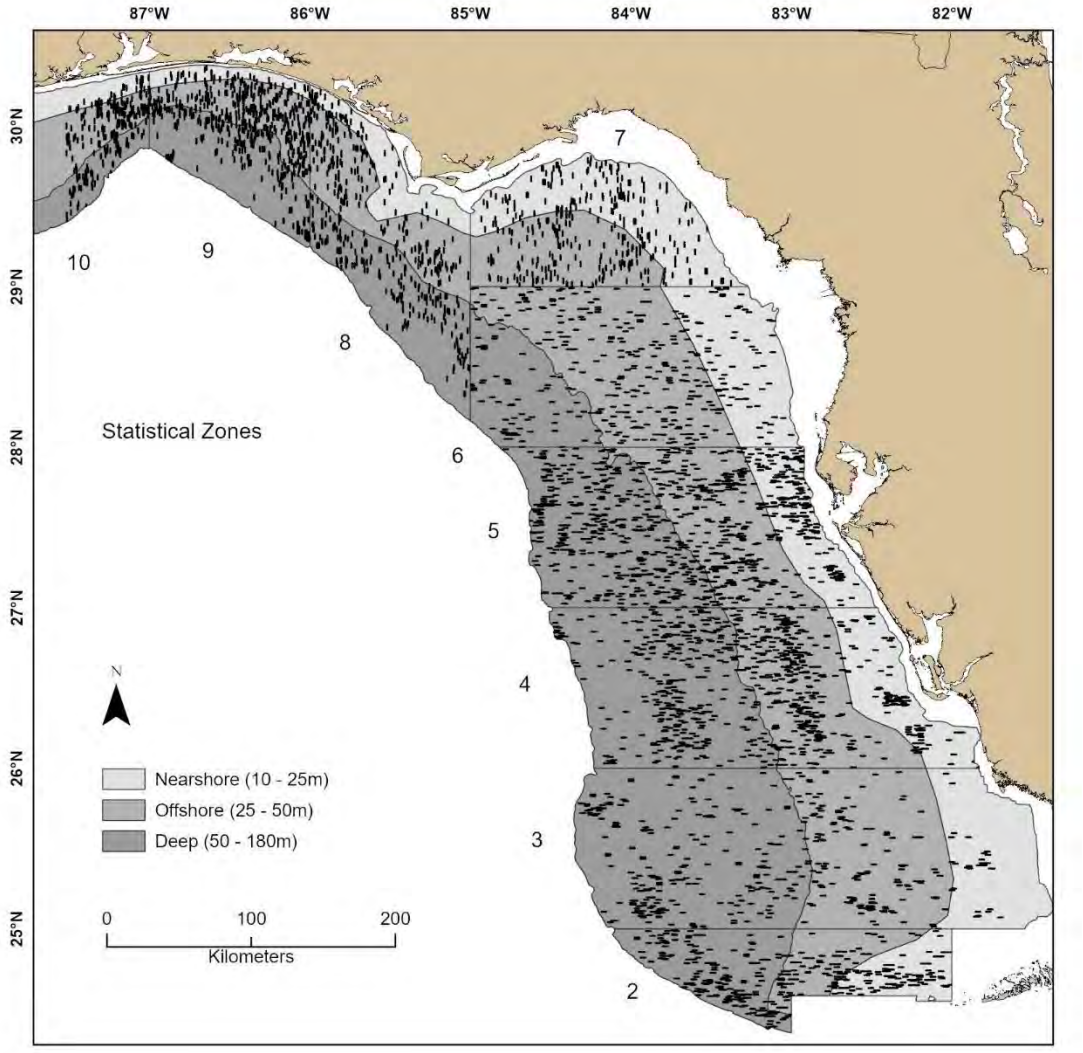


Figure 11.3.11. Spherical stereo-baited remote underwater video (S-BRUV) array used to sample natural and artificial hardbottom sites for the G-FISHER survey.



Figure 11.3.12. Camera-based assessment survey system (C-BASS; Lembke et al. 2017) towed sled configuration for near-bottom fish and habitat assessment.

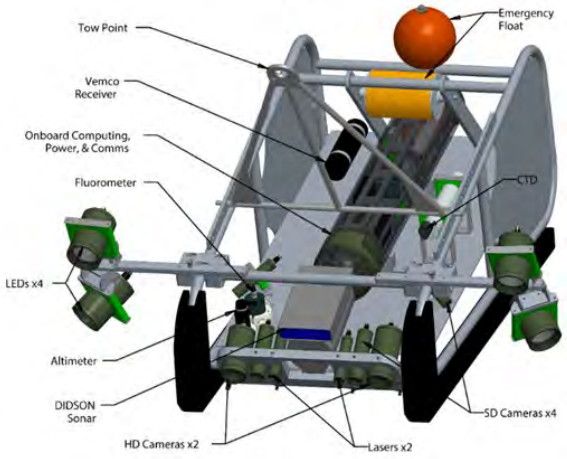




Figure 11.3.13. Conceptual diagram highlighting the key data inputs considered for generating camera-based abundance estimates for Greater Amberjack for natural and artificial hardbottom habitats.

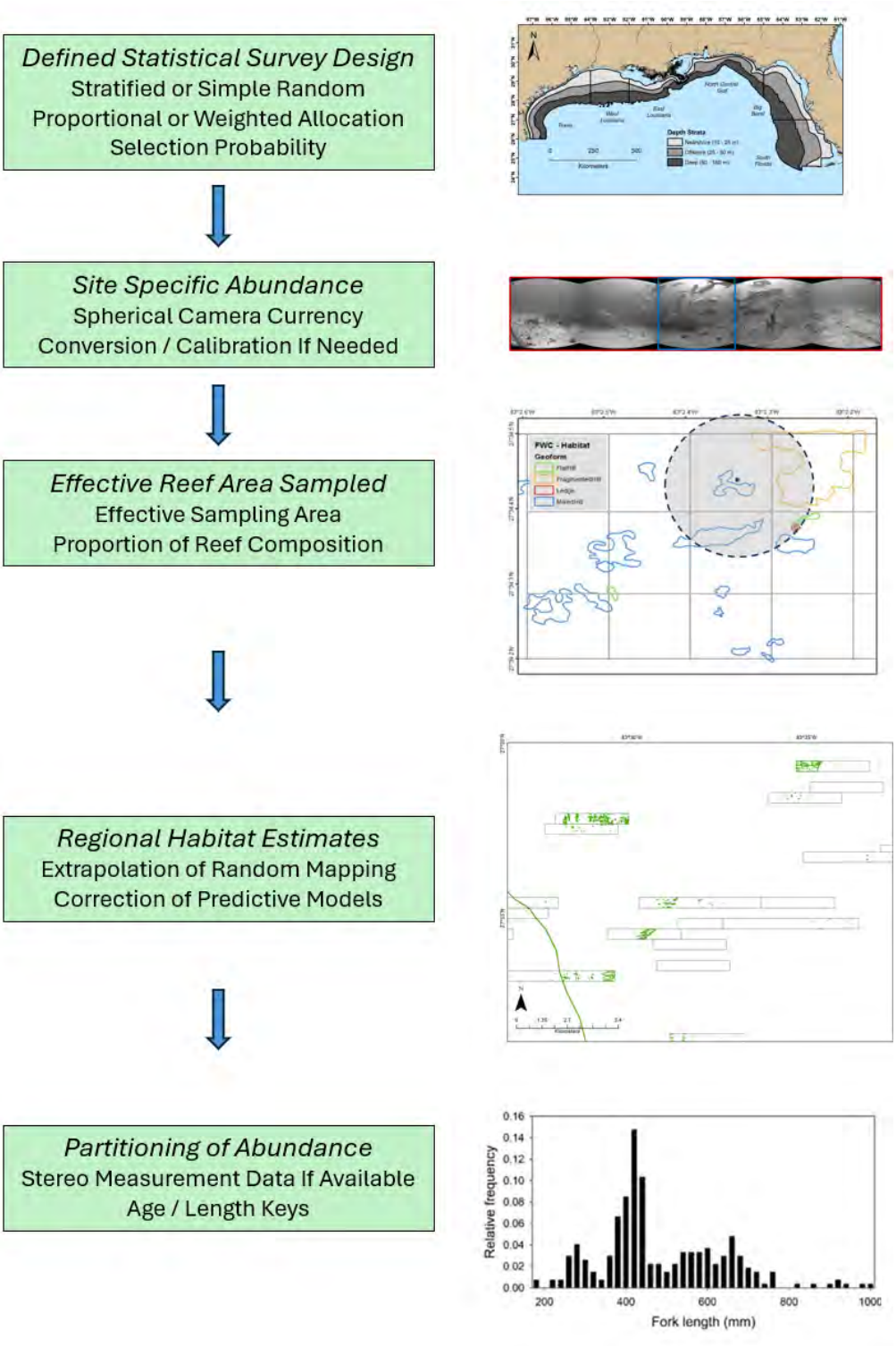


Figure 11.3.14. Schematic diagram of alternative approaches to defining the effective sampling area of the S-BRUV array for Greater Amberjack. The black dot represents the location of S-BRUV deployment. The green shaded area indicates the field of view of an S-BRUV deployment under high visibility conditions (~ 10 m). The black dashed circle represents the estimated average distance that the bait plume could travel by the time the average MaxN was observed (radius of 105 m), while the red dashed ellipse represents a unidirectional bait plume (exact direction unknown). We conservatively chose the black dashed circle as our effective sampling area, which covers an area of 36,644 m2 (see text for rationale).

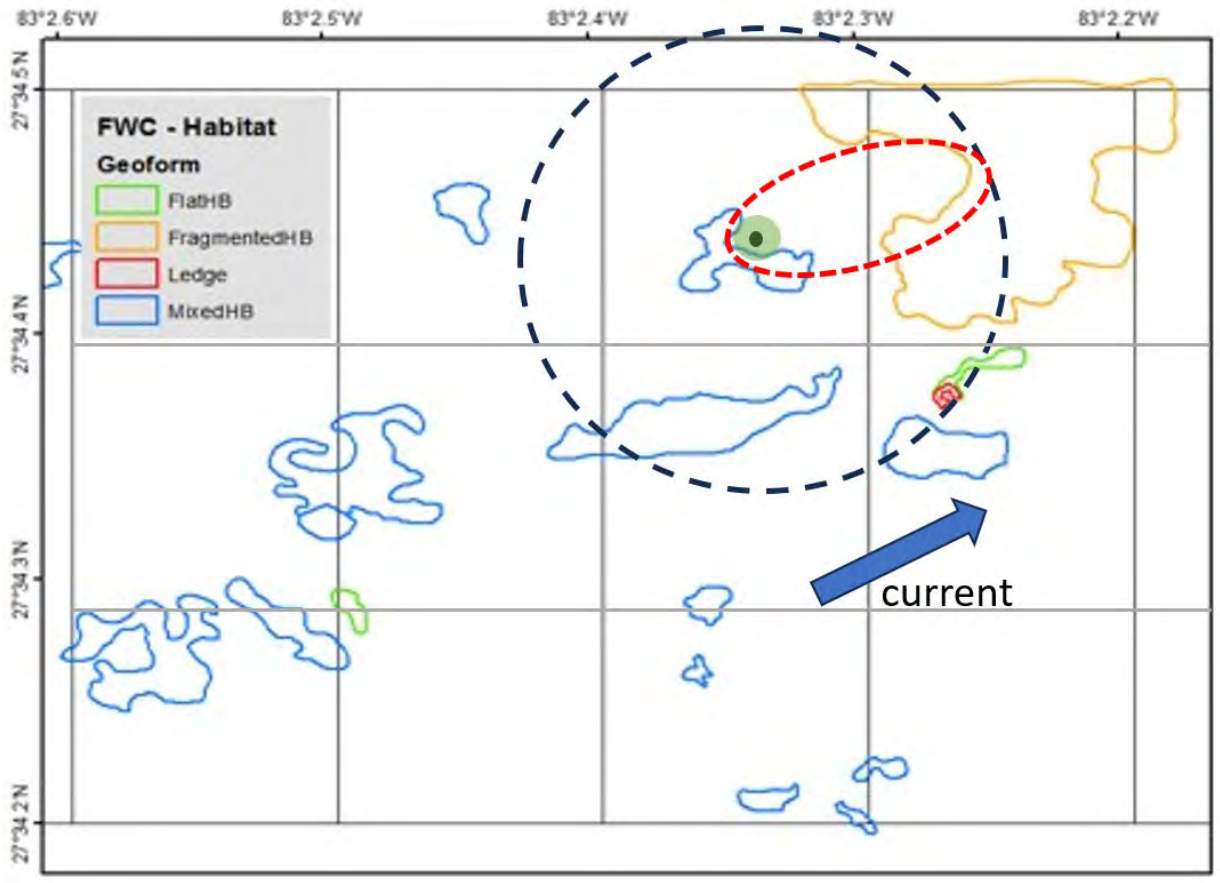


Figure 11.3.15. Frequency distribution of the timing of MaxN of Greater Amberjack on natural (top panel) and artificial hardbottom habitats (bottom panel). Black bars indicate that MaxN was reached at the first time Greater Amberjack were observed, while gray bars indicated that MaxN increased later in the video.

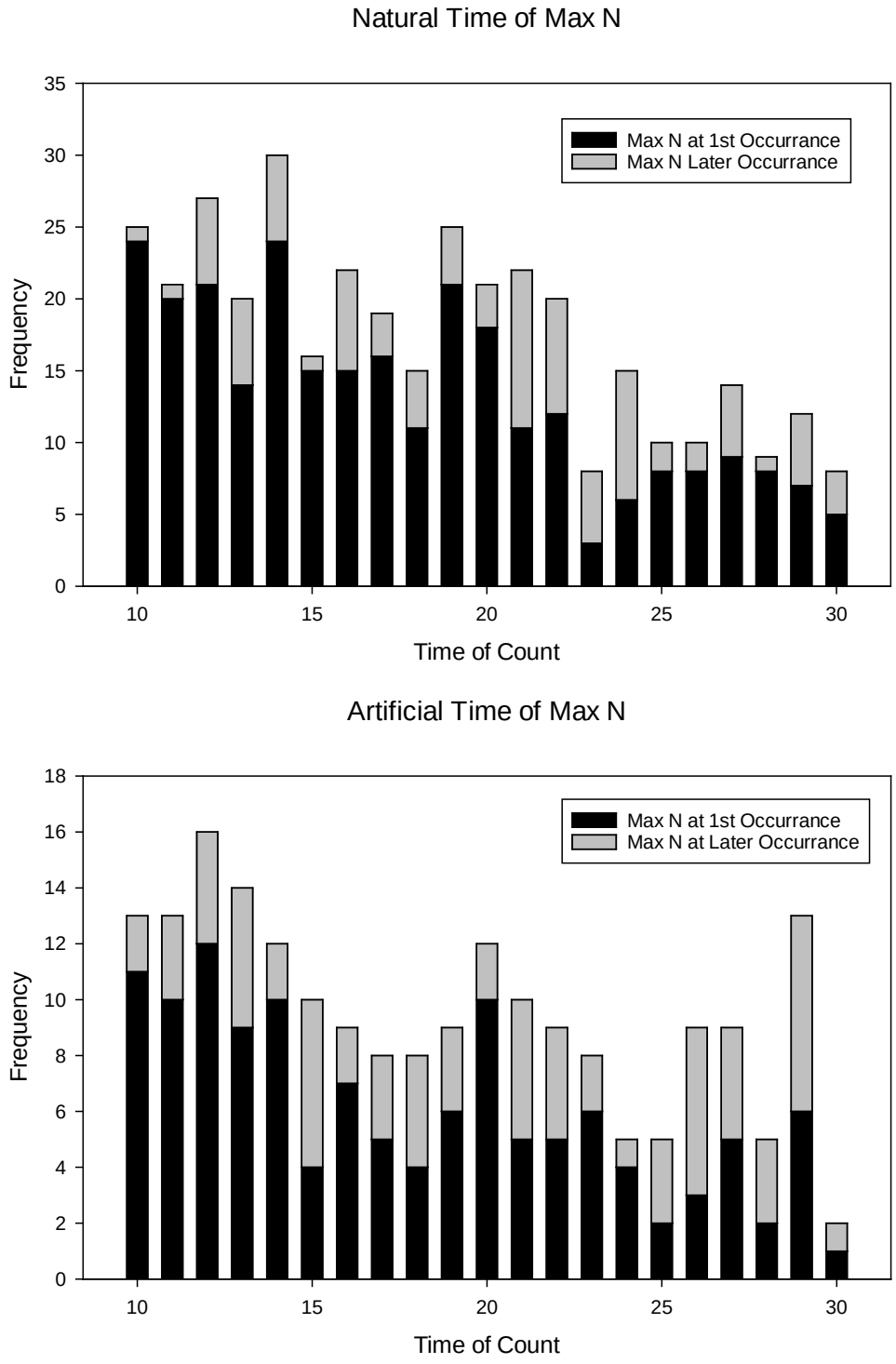


Figure 11.3.16. Example Kernel Density Estimate of effective use area of Greater Amberjack over a 30-minute period (comparable to the duration of S-BRUV deployments) as determined through analysis of VPS location data. Data were generated during a telemetry study conducted off the coast of North Carolina.

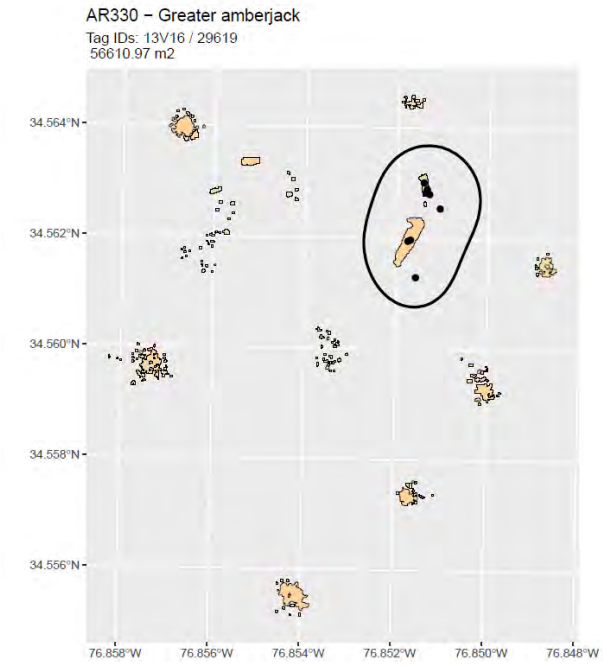


Figure 11.3.17. Map of randomly selected habitat mapping surveys conducted in the eastern Gulf of America. Green polygons represent natural hardbottom habitats, and red polygons represent artificial hardbottom habitats. Percentage coverage of natural and hardbottom habitats were first calculated for each survey, then averaged over all surveys within a statistical reporting zone to provide estimates of total habitat coverage within the study area.

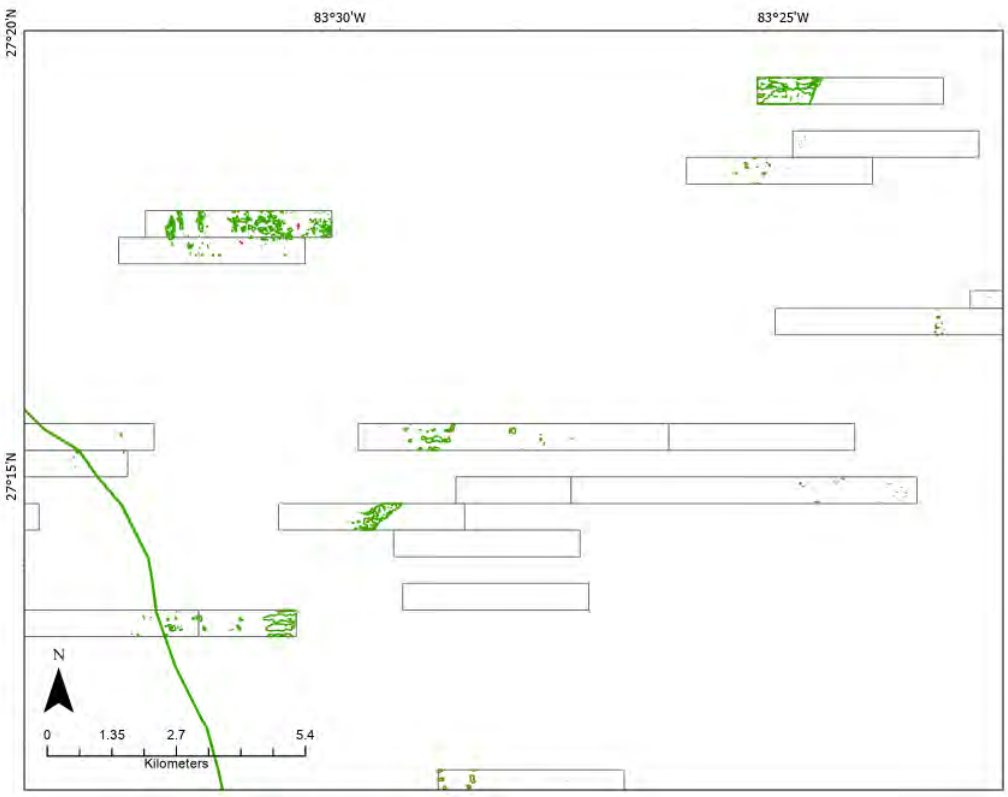


Figure 11.3.18. Map of all natural hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023. Data were obtained from single camera reads, which were completed for all sites. Symbology represents the presence or absence of *Seriola dumerili*.

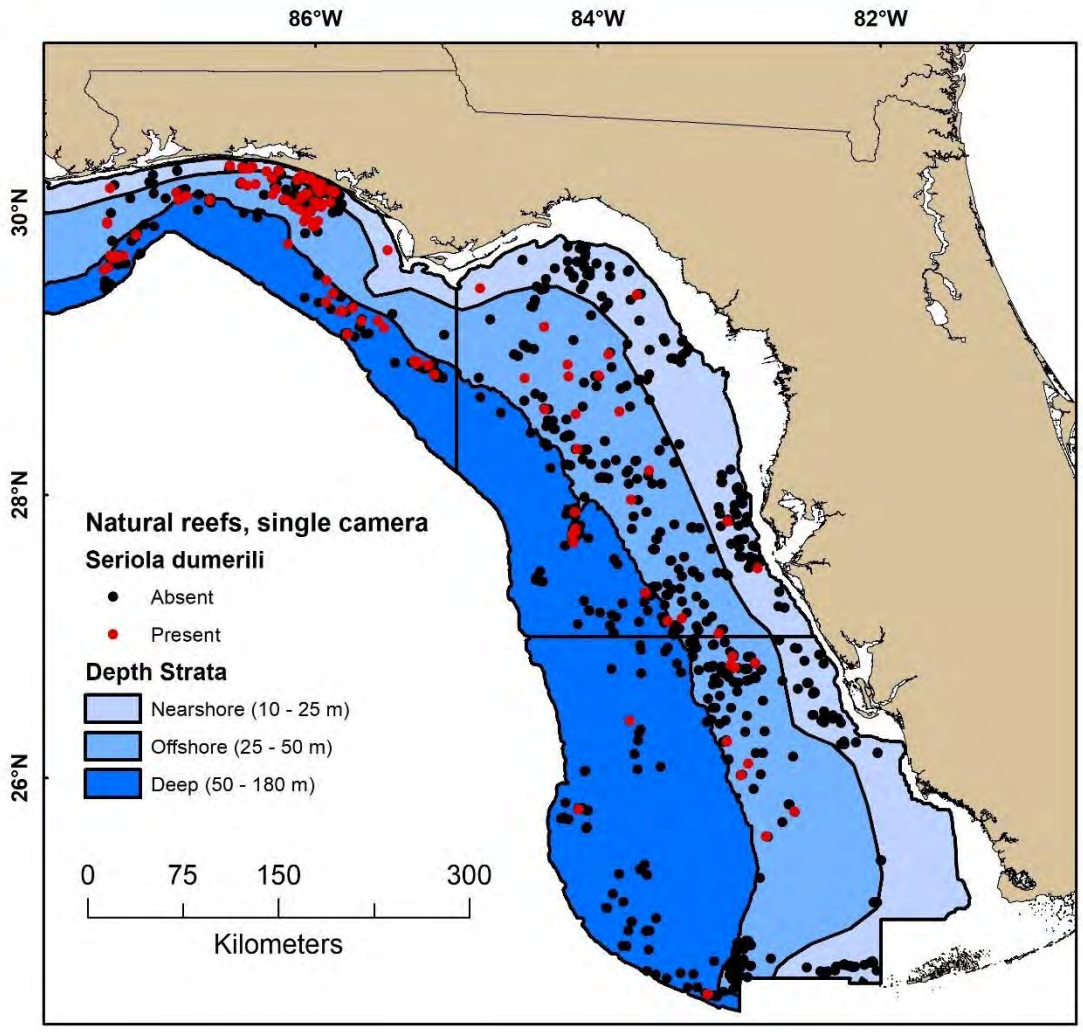


Figure 11.3.19. Map of all natural hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023. Data were obtained from single camera reads, which were completed for all sites. Symbology represents the presence or absence of unidentified *Seriola* spp.

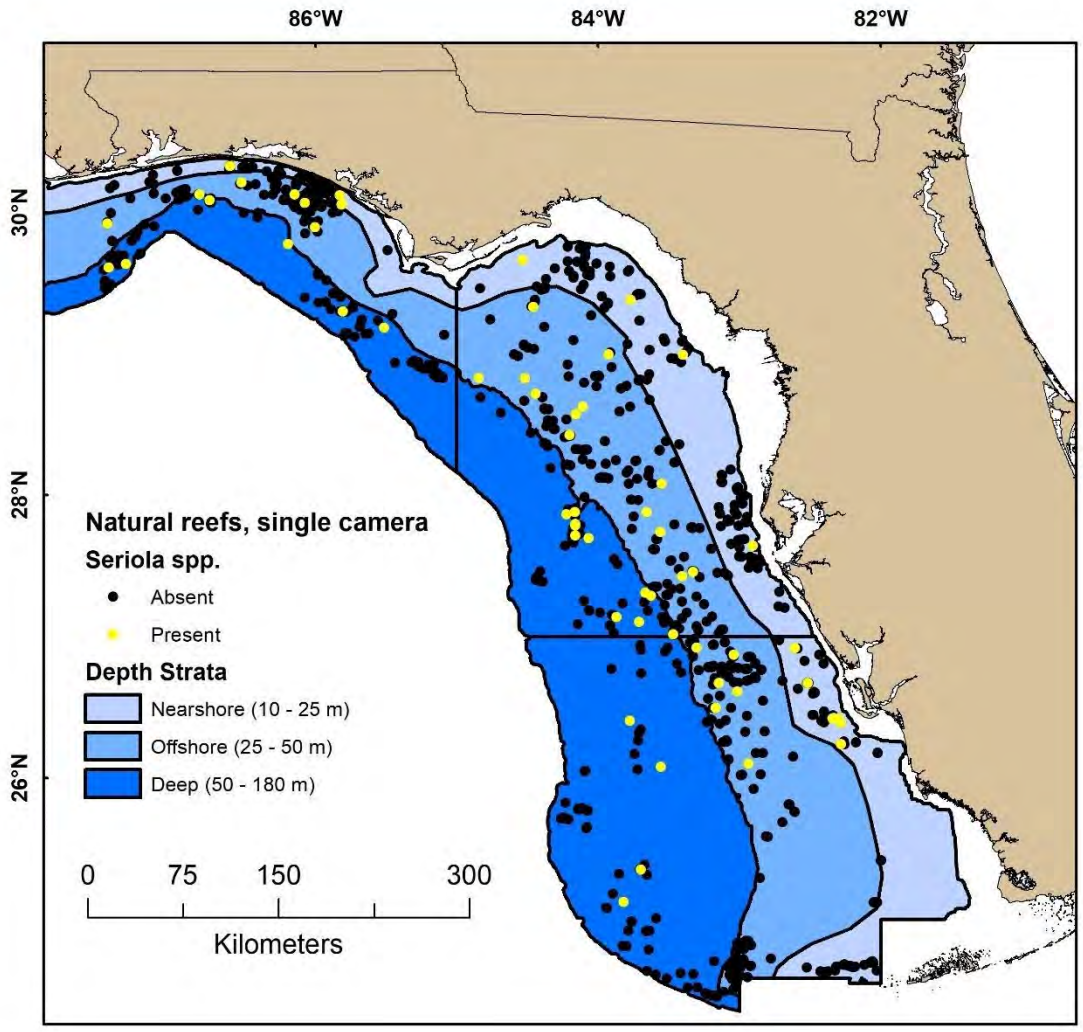


Figure 11.3.20. Average relative abundance (± SE) of *Seriola dumerili* and *Seriola* spp. by spatial strata for single camera reads on natural hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

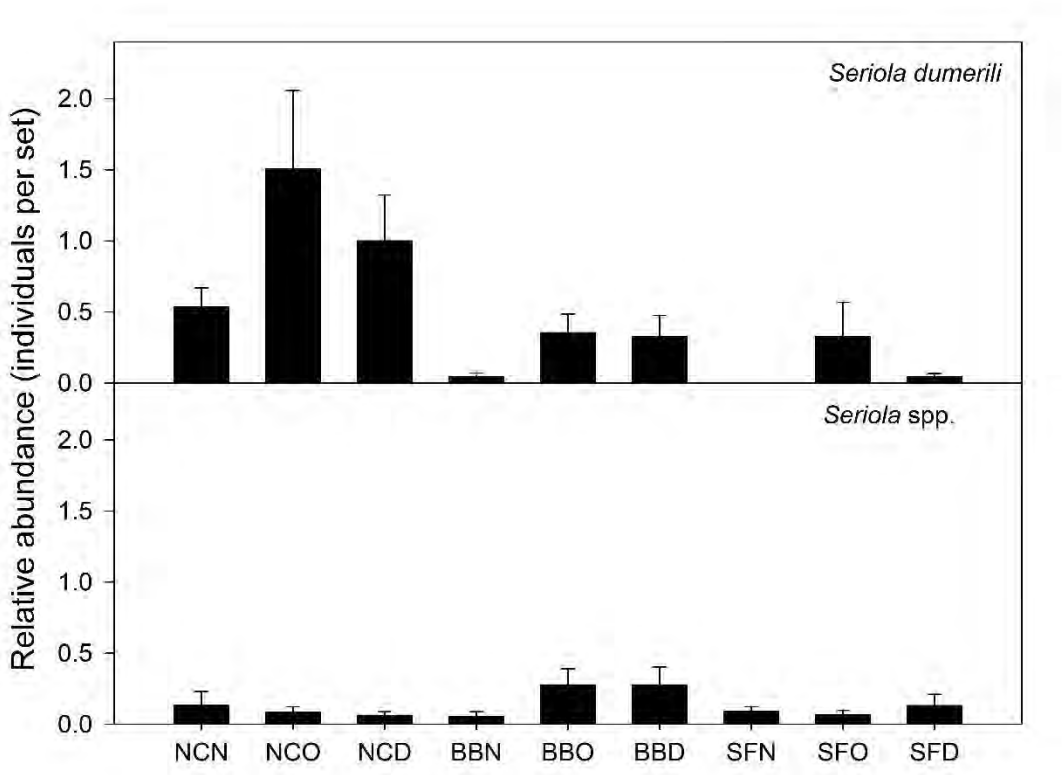


Figure 11.3.21. Average relative abundance (± SE) of *Seriola dumerili* and *Seriola* spp. by habitat strata for single camera reads on natural hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Habitat strata labels represent a combination of hardbottom origin (Natural), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large).

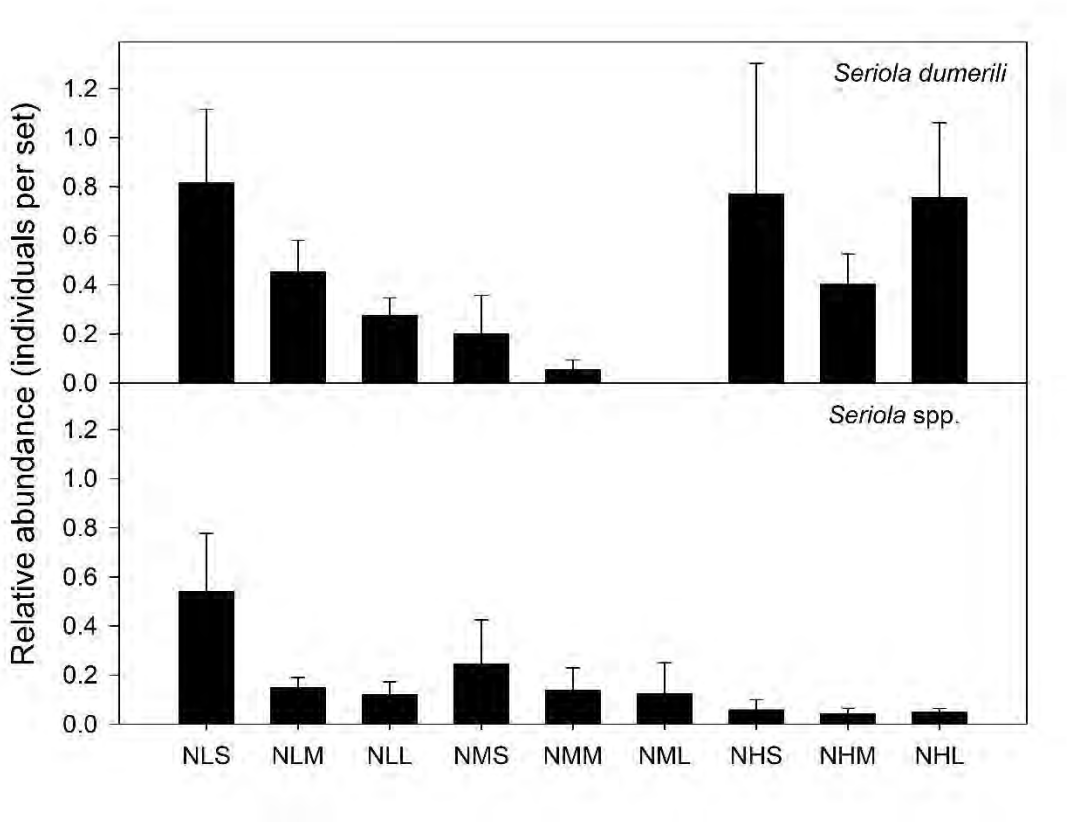


Figure 11.3.22. Size frequency of all *Seriola dumerili* that were observed on natural hardbottom habitats and able to be measured from the 2023 G-FISHER S-BRUV survey.

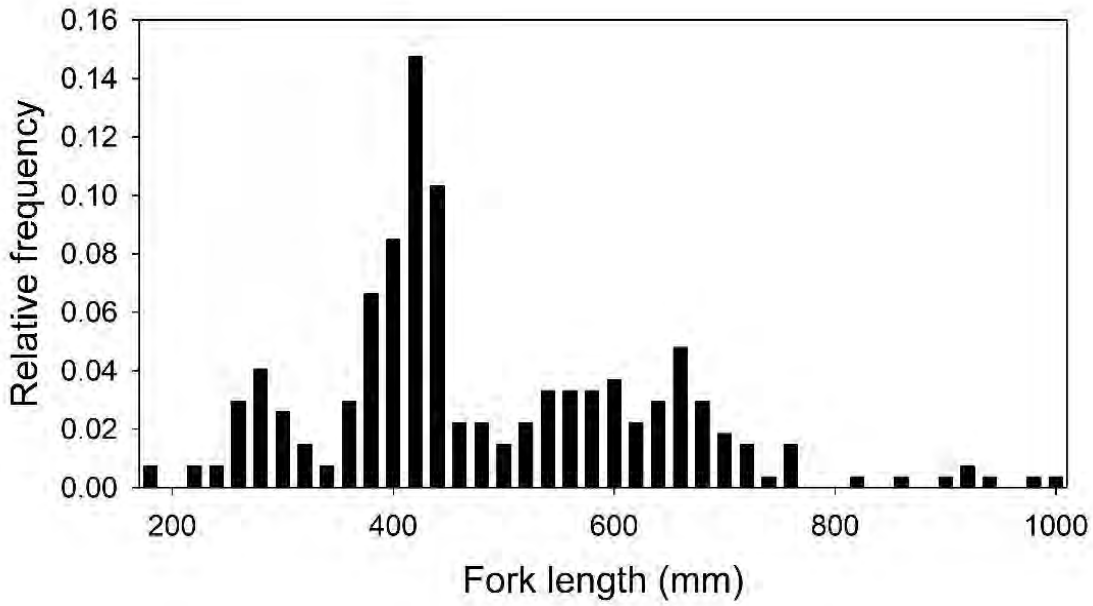


Figure 11.3.23. Map of all natural hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023 for which full spherical camera reads were completed. These are a subset of all sites sampled in 2023 and represent all sites where concurrent hydroacoustic surveys were conducted. Symbology represents the presence or absence of *Seriola dumerili*.

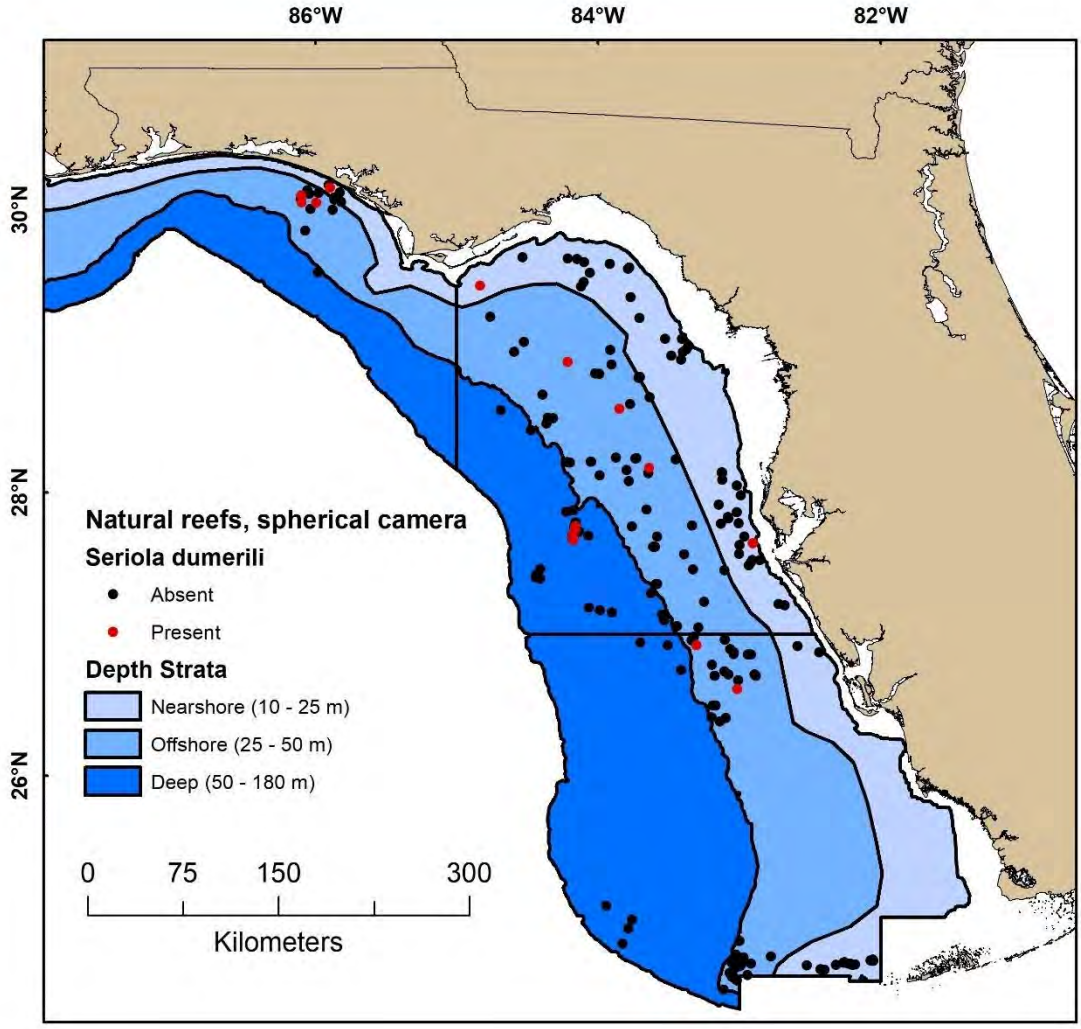


Figure 11.3.24. Map of all natural hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023 for which full spherical camera reads were completed. These are a subset of all sites sampled in 2023 and represent all sites where concurrent hydroacoustic surveys were conducted. Symbology represents the presence or absence of unidentified *Seriola* spp.

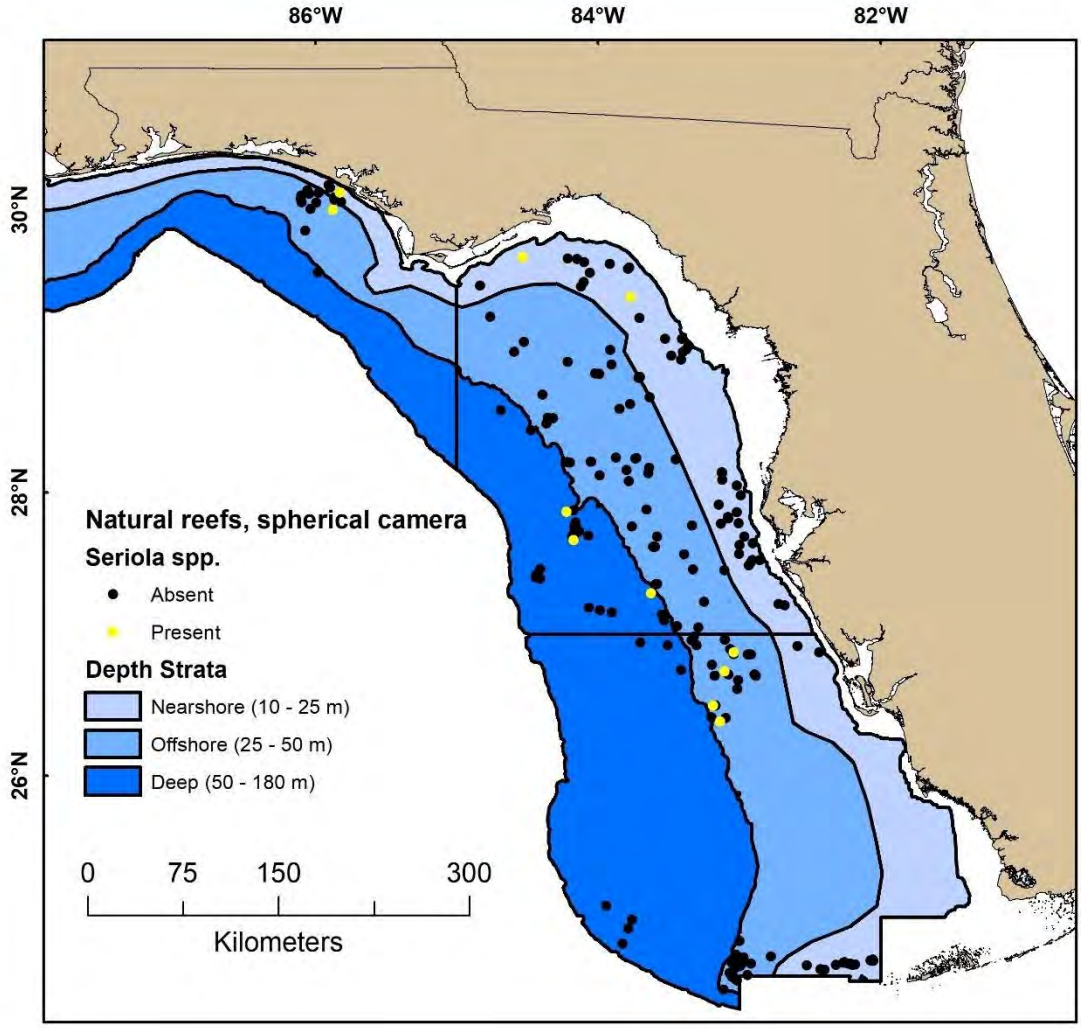


Figure 11.3.25. Map of all artificial hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023. Data were obtained from single camera reads, which were completed for all sites. Symbology represents the presence or absence of *Seriola dumerili*.

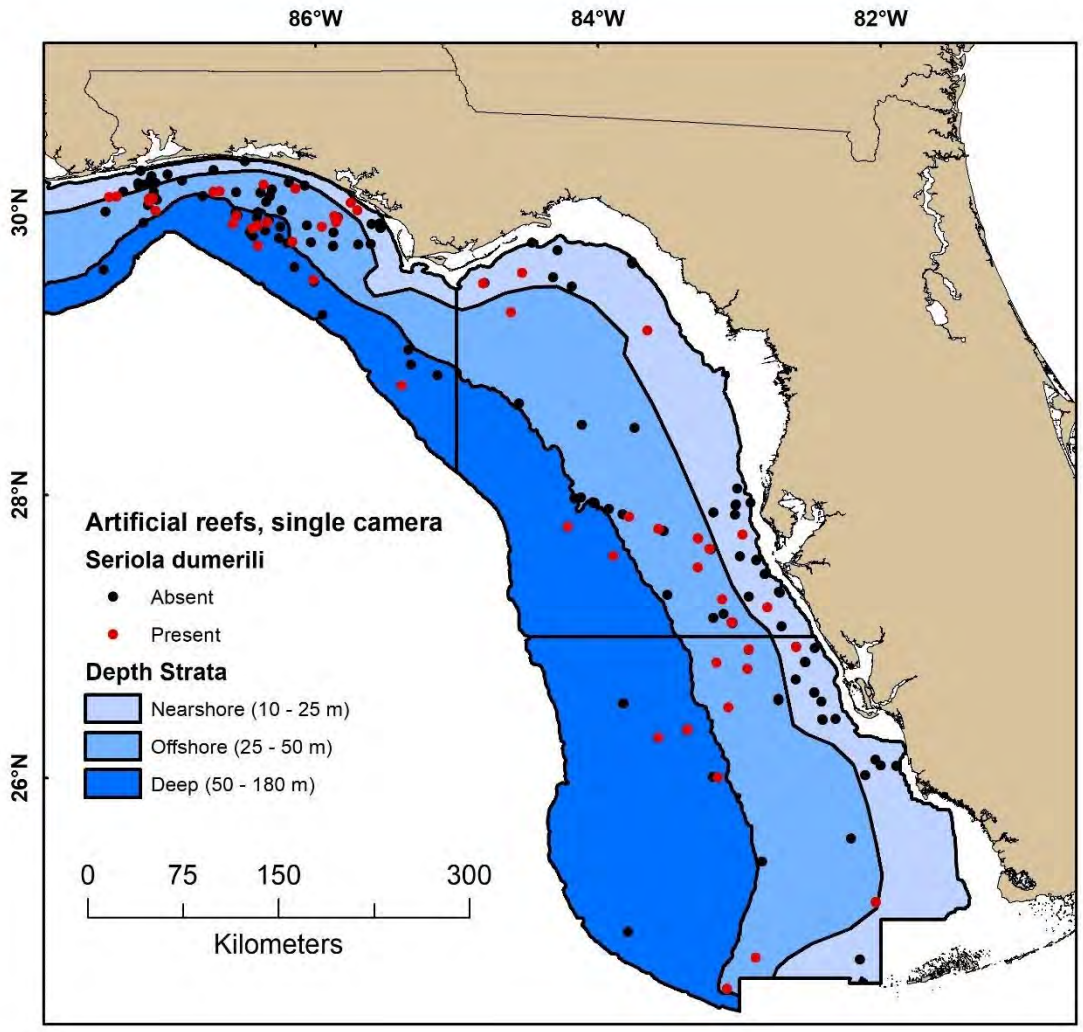


Figure 11.3.26. Map of all artificial hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023. Data were obtained from single camera reads, which were completed for all sites. Symbology represents the presence or absence of unidentified *Seriola* spp.

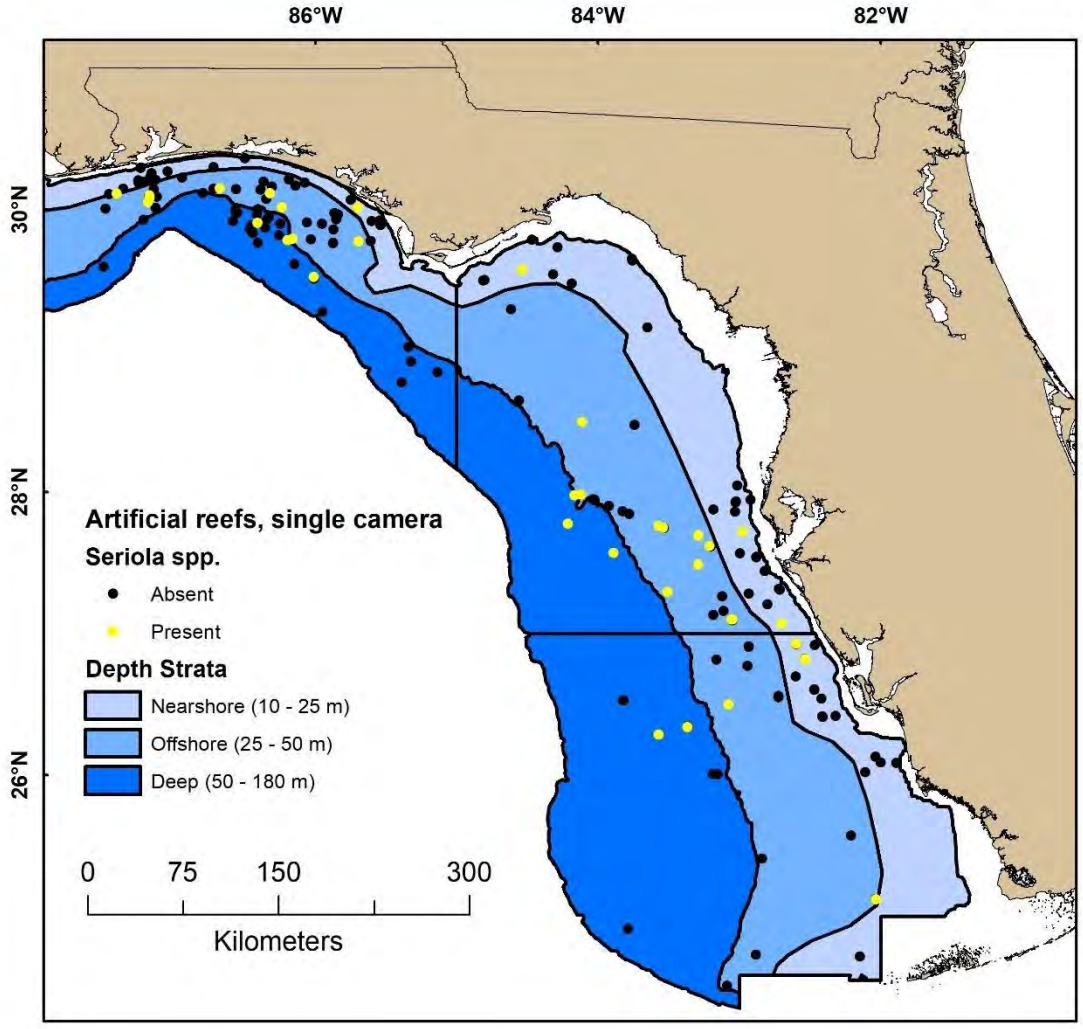


Figure 11.3.27. Average relative abundance (± SE) of *Seriola dumerili* and *Seriola* spp. by spatial strata for single camera reads on artificial hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Spatial strata labels represent a combination of region (NC: North-Central, BB: Big Bend, SF: South Florida) and depth (N: Nearshore, O: Offshore, D: Deep).

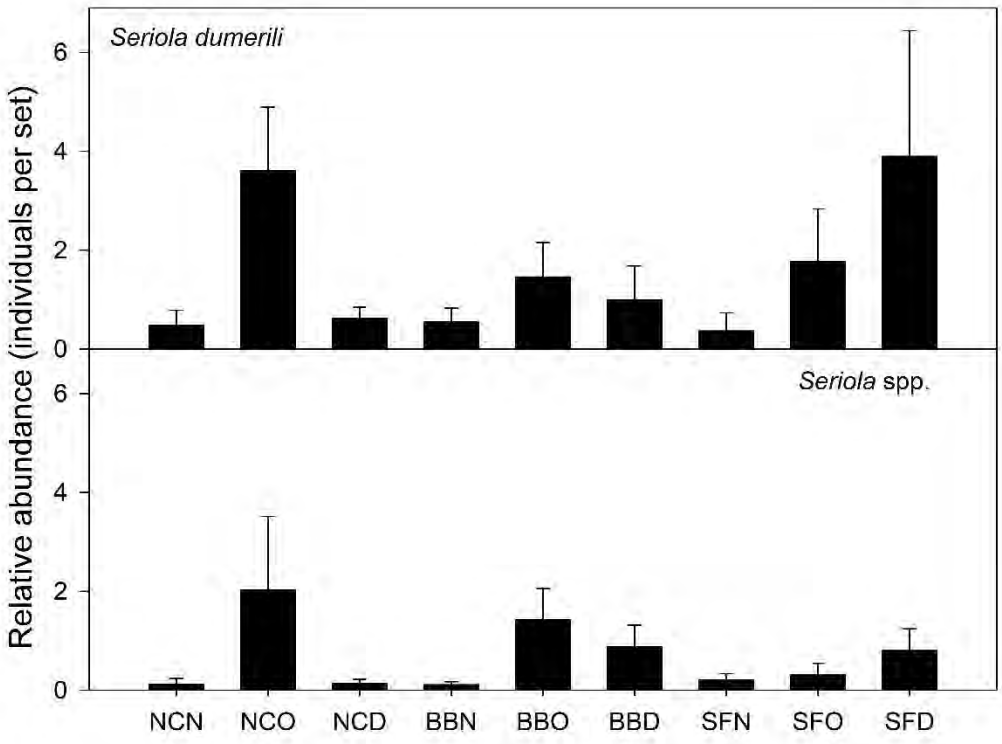


Figure 11.3.28. Average relative abundance (± SE) of *Seriola dumerili* and *Seriola* spp. by habitat strata for single camera reads on artificial hardbottom sites sampled during the 2023 G-FISHER S-BRUV survey. Habitat strata labels represent a combination of hardbottom origin (Artificial), relative relief (L: Low-relief, M: Moderate-relief, H: High-relief), and reef size (S: Small, M: Medium, L: Large).

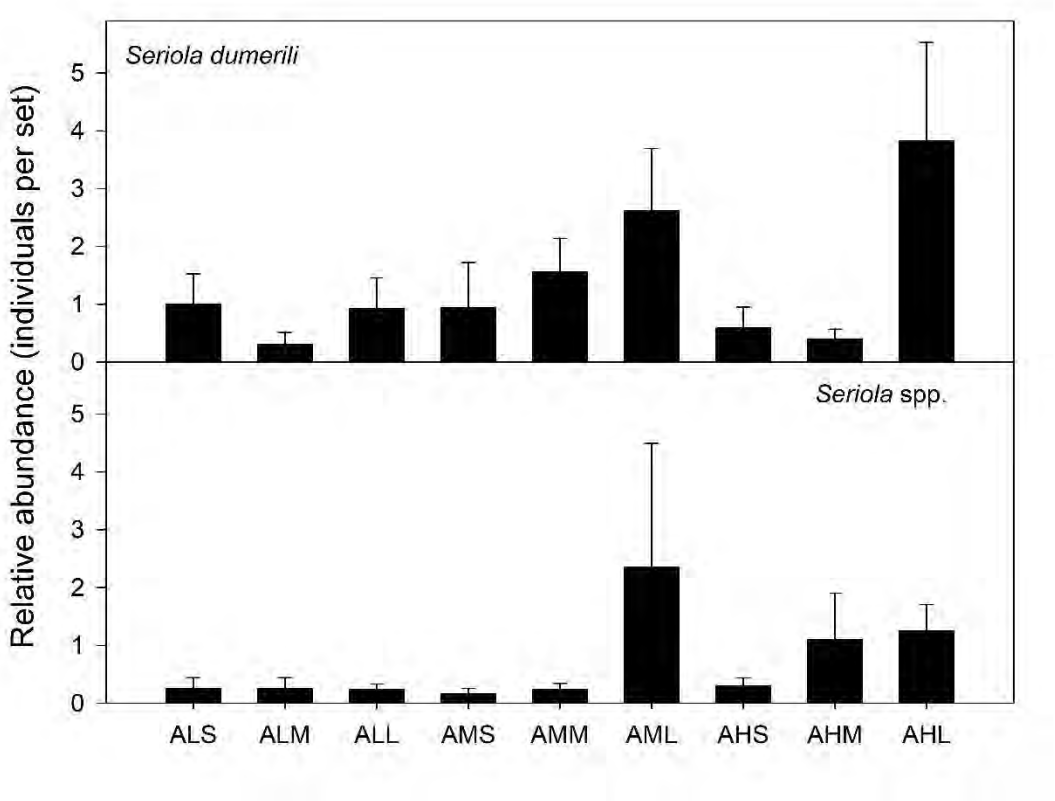


Figure 11.3.29. Size frequency of all *Seriola dumerili* that were observed on artificial hardbottom habitats and able to be measured from the 2023 G-FISHER S-BRUV survey.

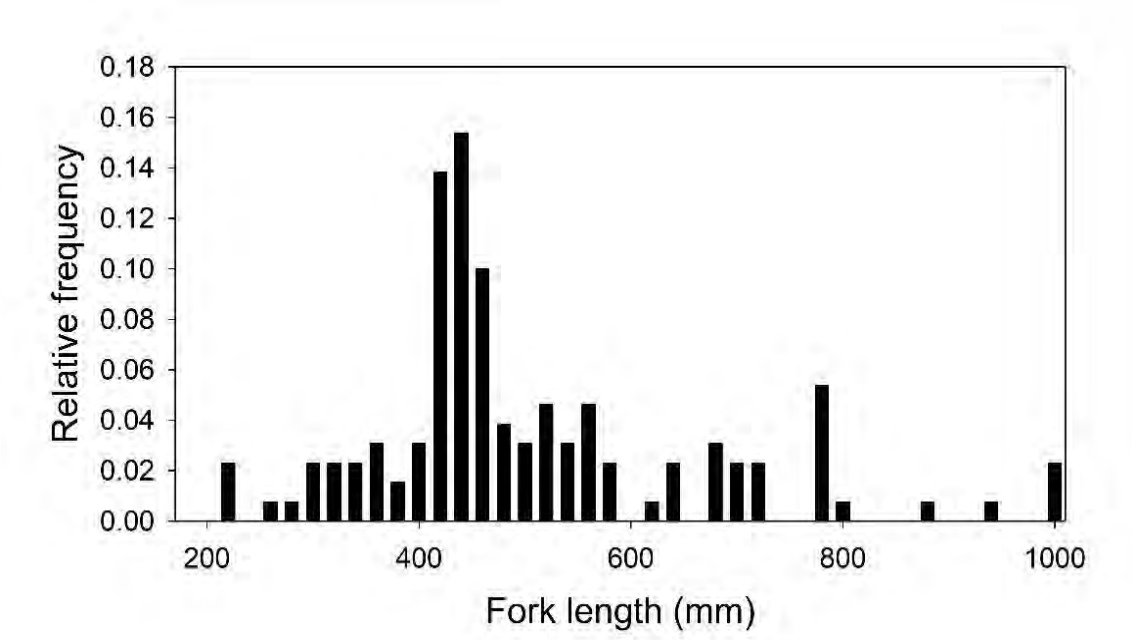


Figure 11.3.30. Map of all artificial hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023 for which full spherical camera reads were completed. These are a subset of all sites sampled in 2023 and represent all sites where concurrent hydroacoustic surveys were conducted. Symbology represents the presence or absence of *Seriola dumerili*.

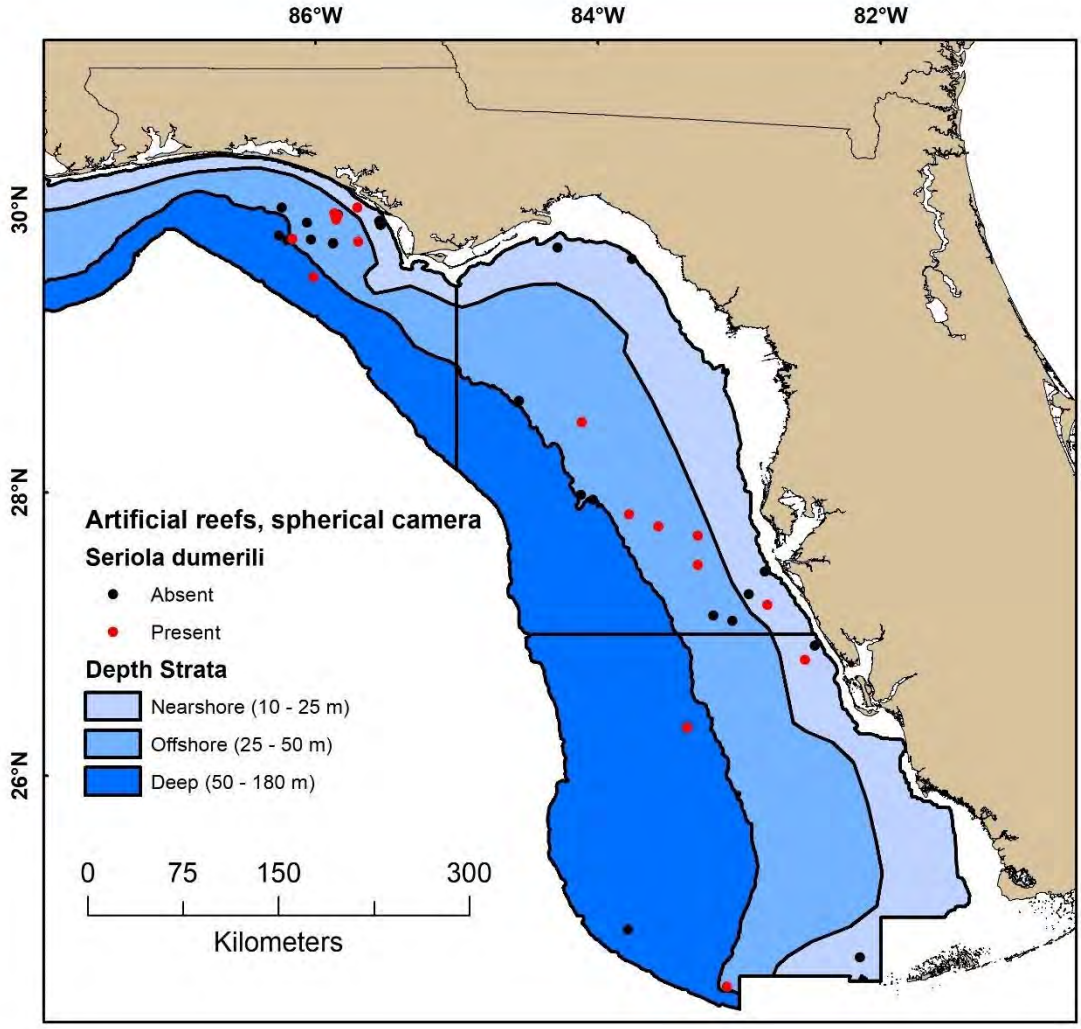


Figure 11.3.31. Map of all artificial hardbottom sites sampled by the G-FISHER S-BRUV survey in 2023 for which full spherical camera reads were completed. These are a subset of all sites sampled in 2023 and represent all sites where concurrent hydroacoustic surveys were conducted. Symbology represents the presence or absence of unidentified *Seriola* spp.

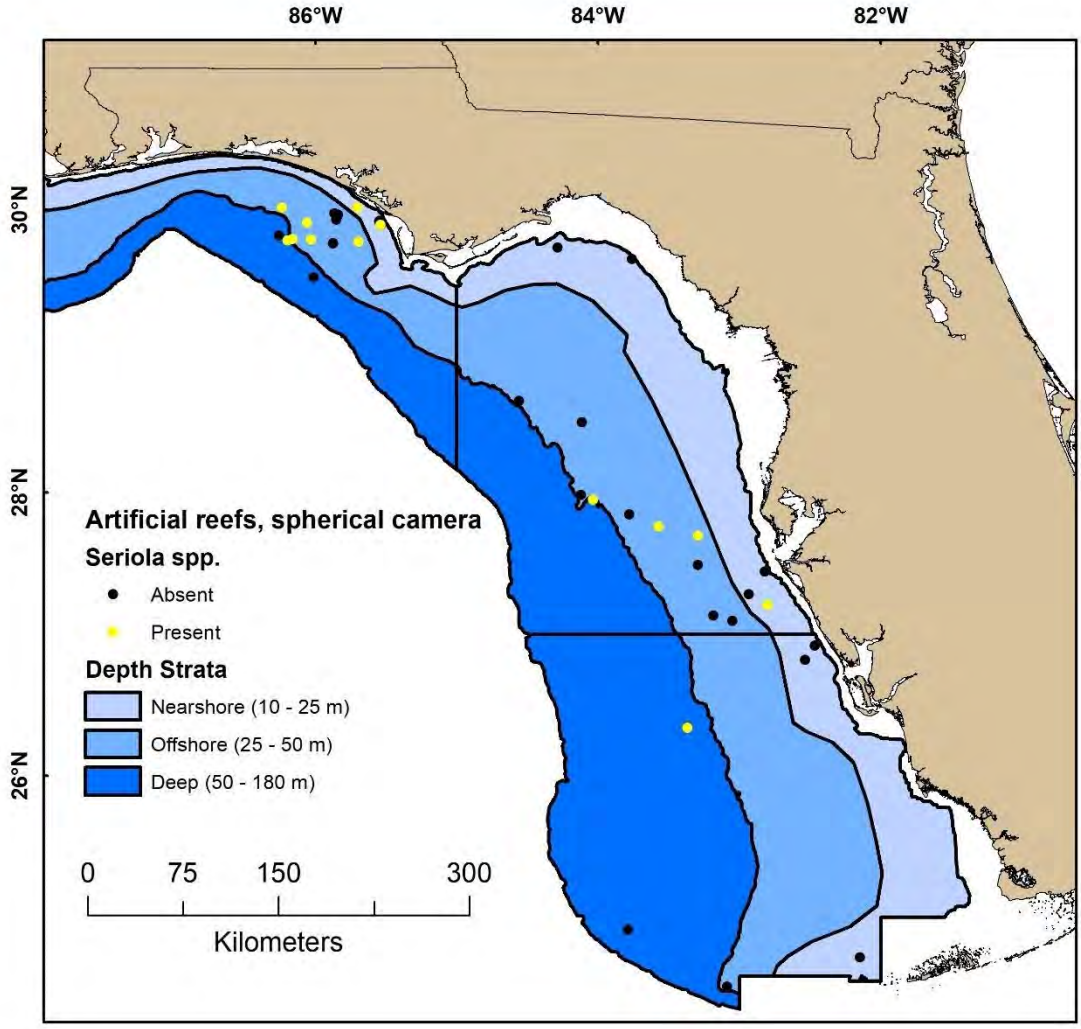


Figure 11.3.32. Map of all towed C-BASS sampling transects conducted targeting unconsolidated sediment habitats in the eastern Gulf of America. The location circled in black is the only site where Greater Amberjack were observed on unconsolidated sediment habitats (N = 4).

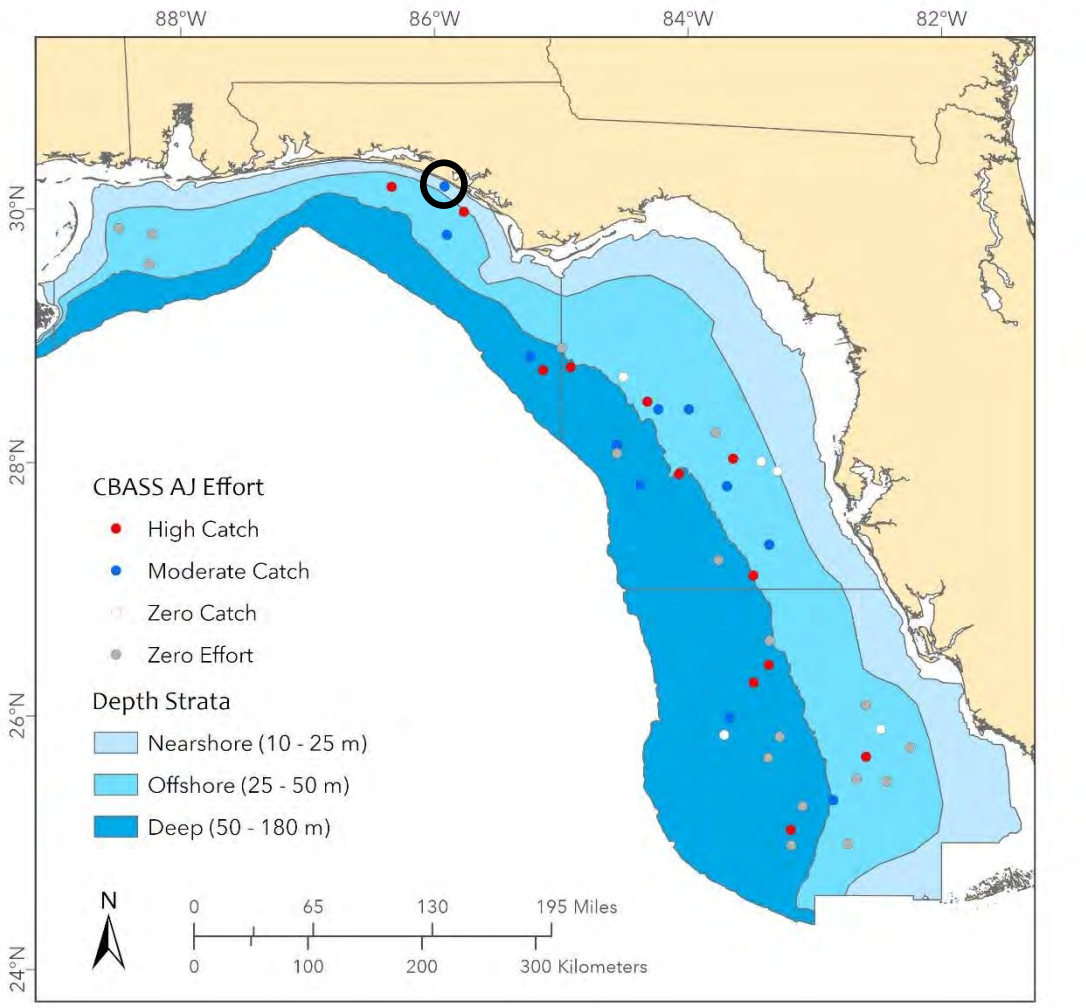


Figure 11.3.33. Map of all paired single-camera to spherical-camera sites (2020 – 2023) analyzed to develop calibration coefficients for conversion from single-camera to spherical-camera counts for *Seriola dumerili*. Red symbols indicate all sites where *Seriola dumerili* were observed on one of the two camera reads.

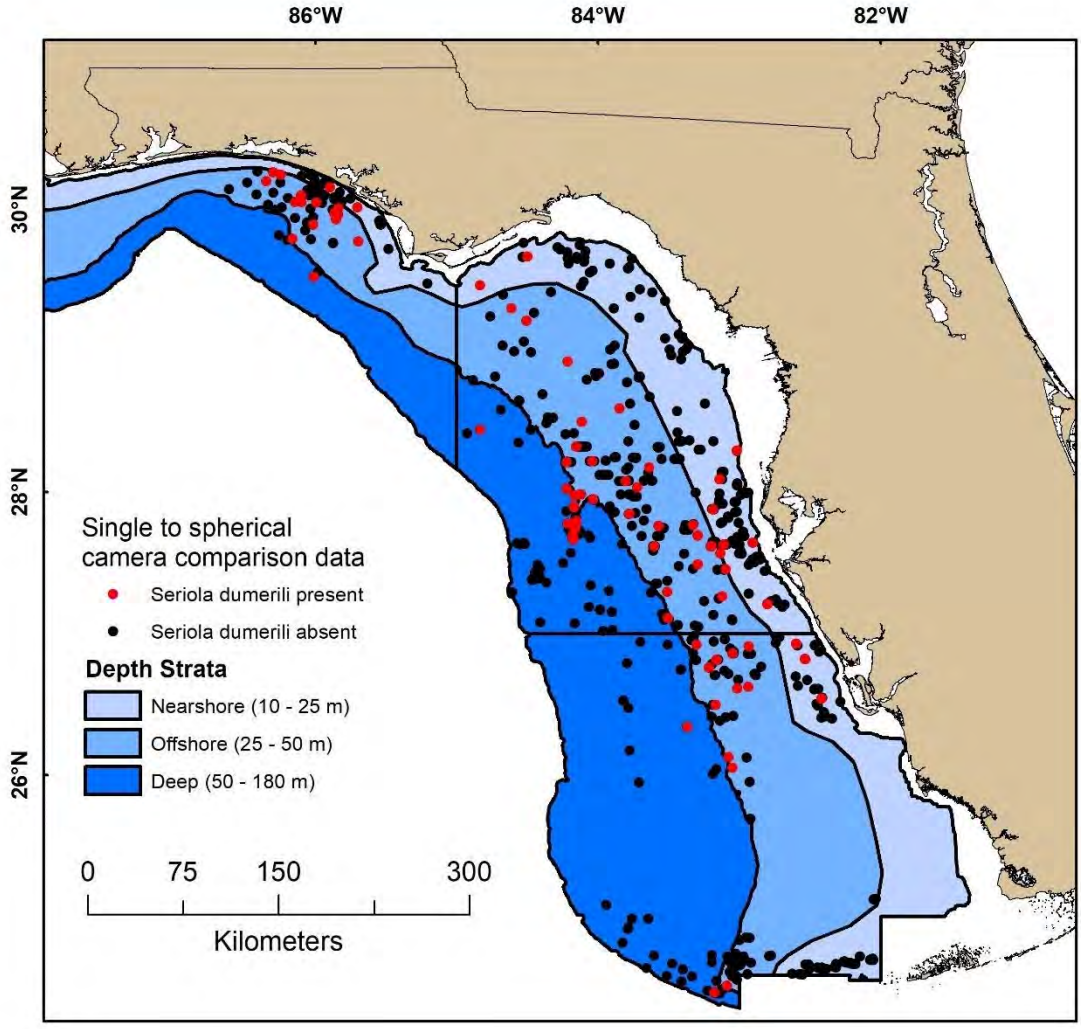


Figure 11.3.34. Plot of spherical-camera and single-camera counts for *Seriola dumerili*. The dashed red line represents the theoretical 1:1 ratio line where counts from the two cameras would be identical.

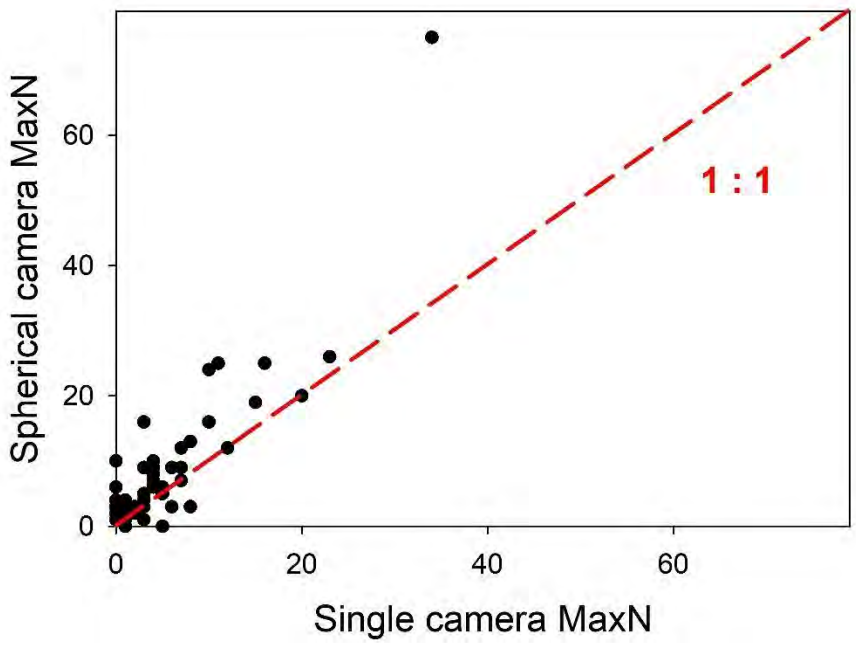


Figure 11.3.35. Estimated spherical camera counts, along with 68% prediction intervals, for single-camera data as generated through Bayesian Poisson regression.

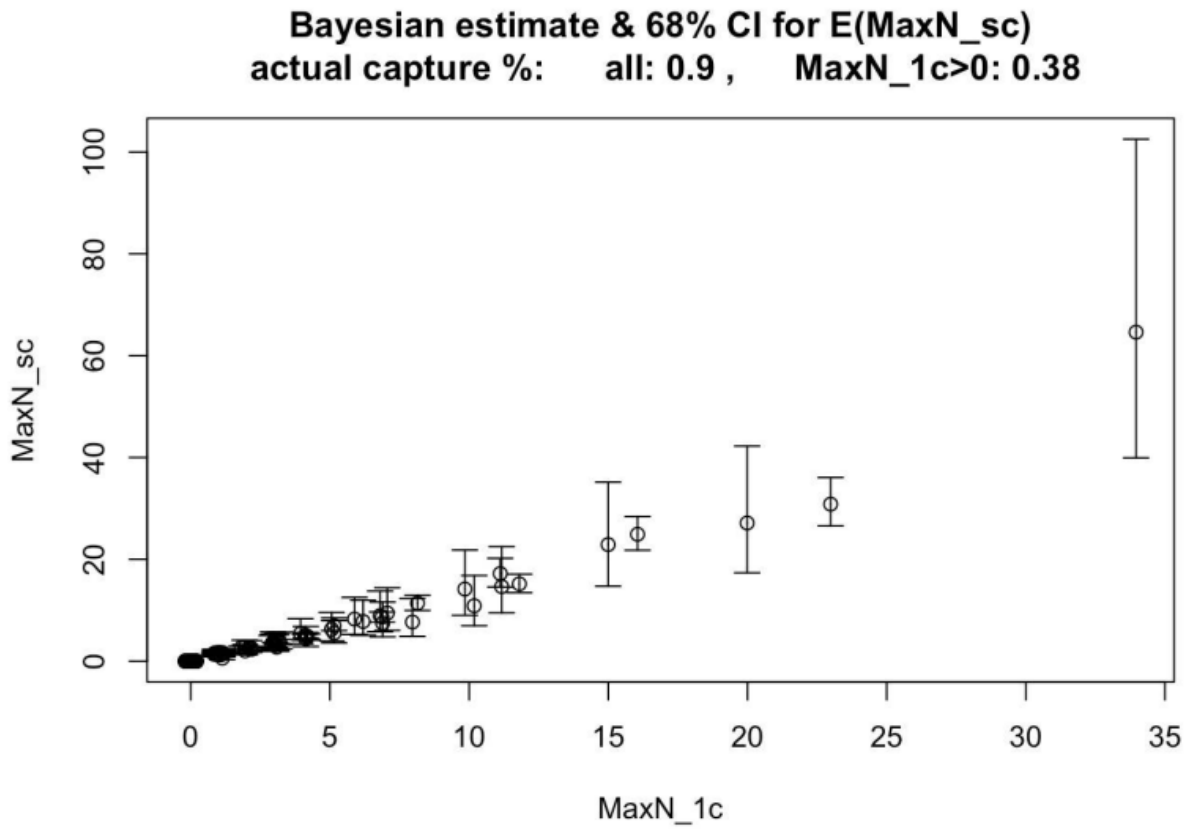


Figure 11.3.36. Map of paired baited / unbaited sampling sites. Artificial hardbottom sites are plotted on the left panel, and natural hardbottom sites are plotted on the right panel. Fill colors identify order in which gear was deployed.

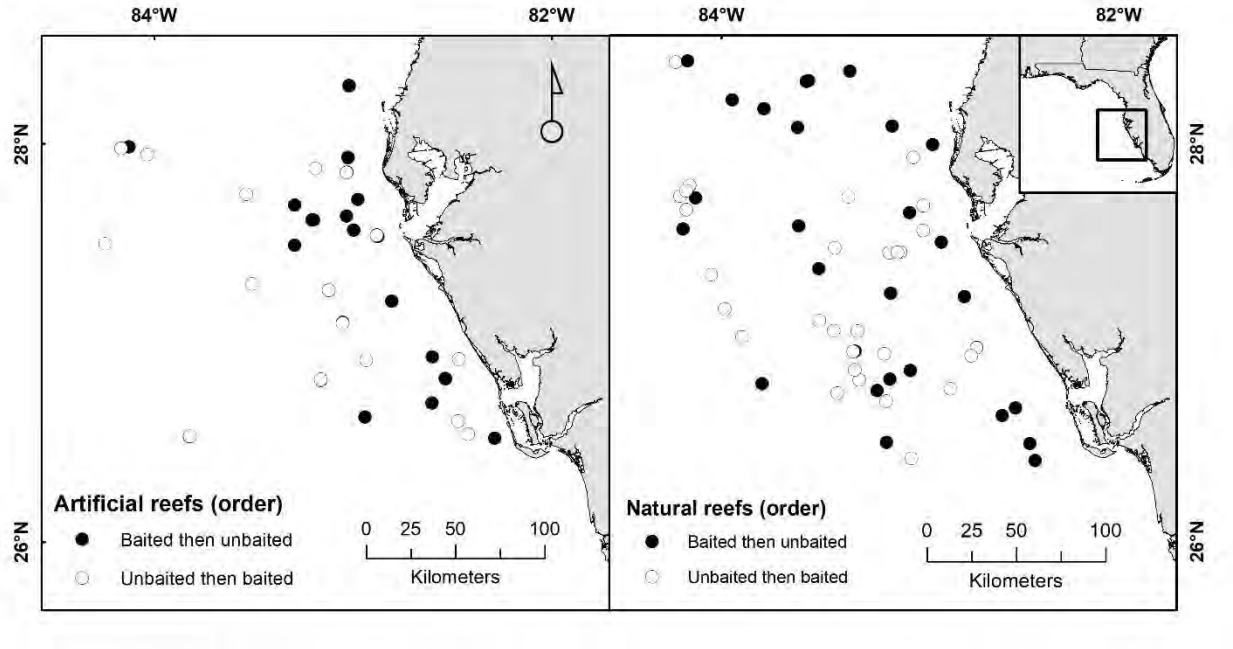


Figure 11.3.37. Results of similarity percentage analyses identifying taxa that contributed to the top 85% of observed differences in assemblage structure between baited and unbaited sets as well as natural and artificial hardbottom habitats. Taxa order are based on the average relative abundance among all sets. Note the differences in scaling for the X-axis between the top six species and those that follow.

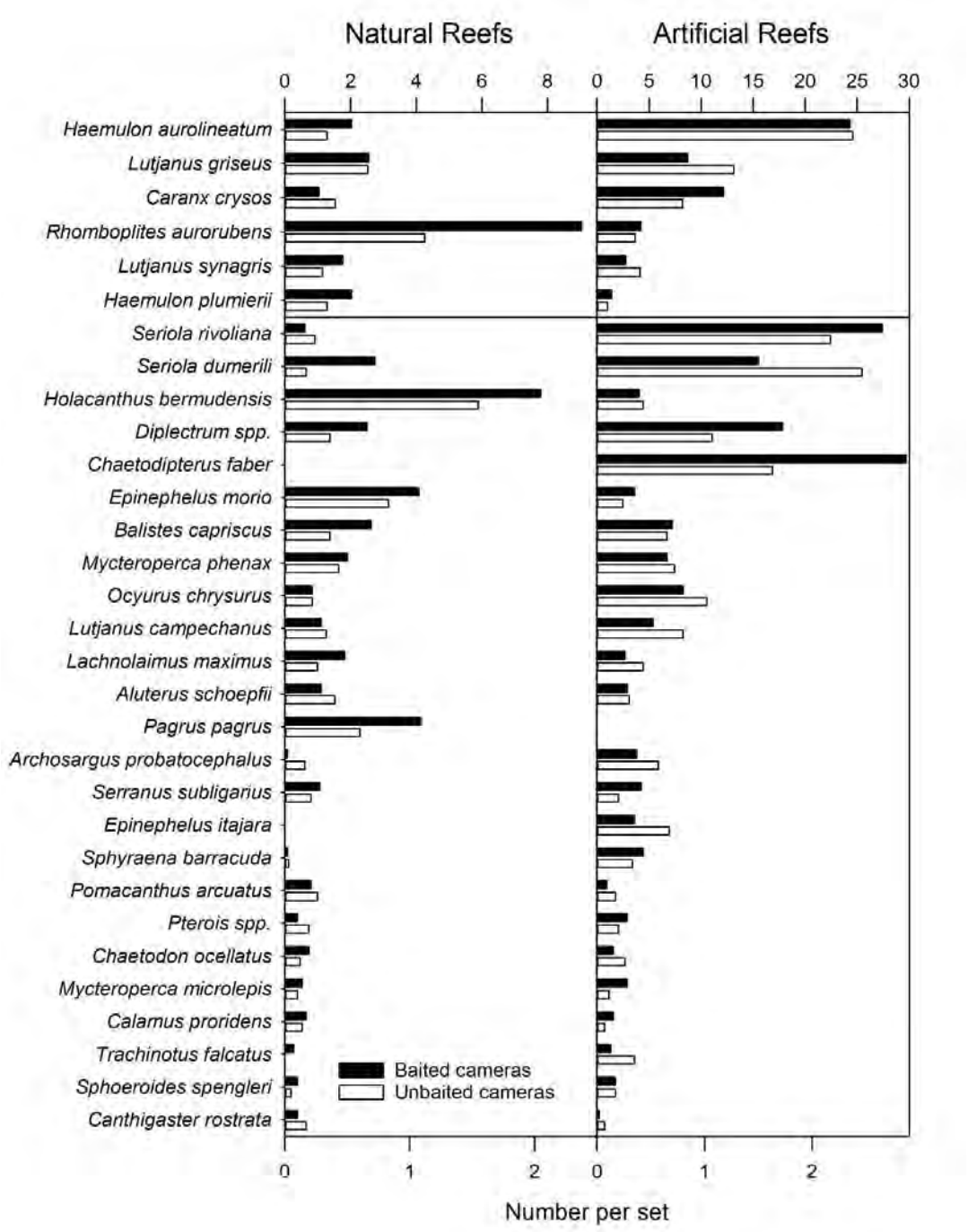


Figure 11.3.38. Scatterplot of Greater Amberjack MaxN values observed during standard S-BRUV reads (Y-axis) and those observed prior to the standard read, including the drop through the water column where available (X-axis). Sites where no Greater Amberjack were observed were excluded.

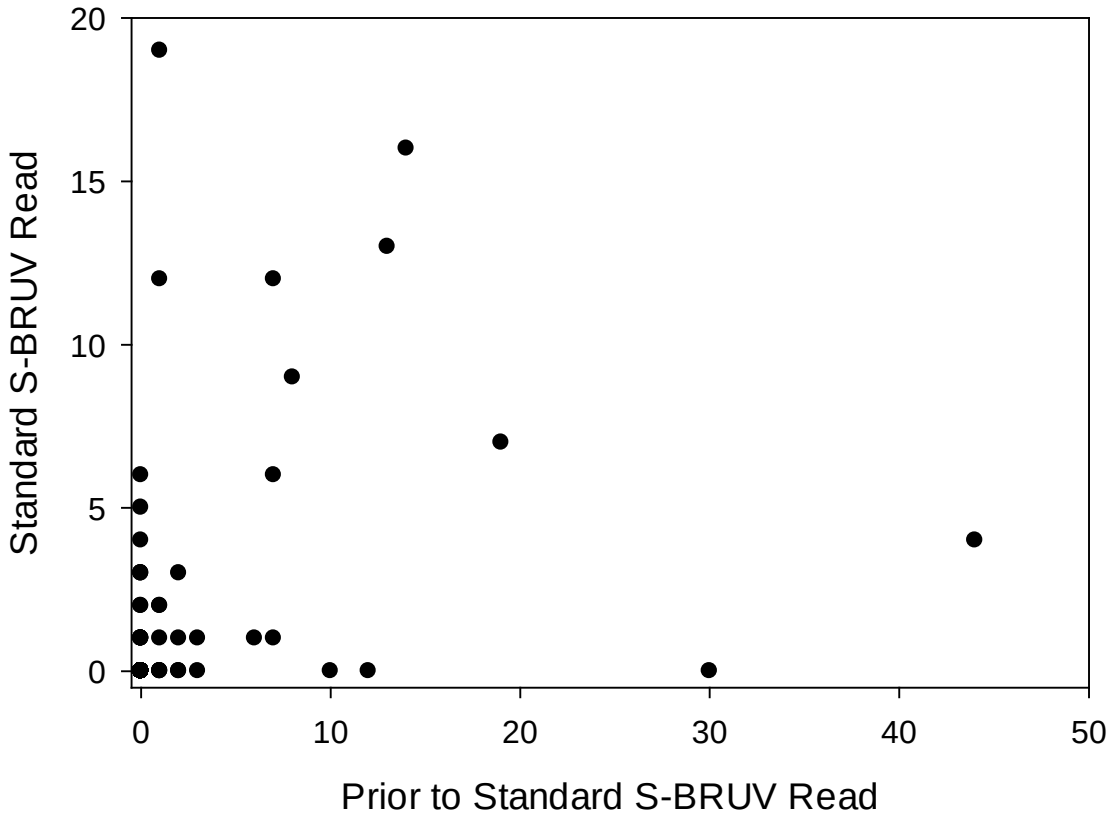


Figure 11.3.39. Scatterplot of large-bodied (>20cmTL), water column (>0.5m off bottom), disaggregated fish (LBWCF or potential *Seriola*), as determined through echo integration analysis of acoustics data, with Greater Amberjack MaxN values observed during standard S-BRUV reads (bottom panel) and those observed prior to the standard read, including the drop through the water column (top panel).

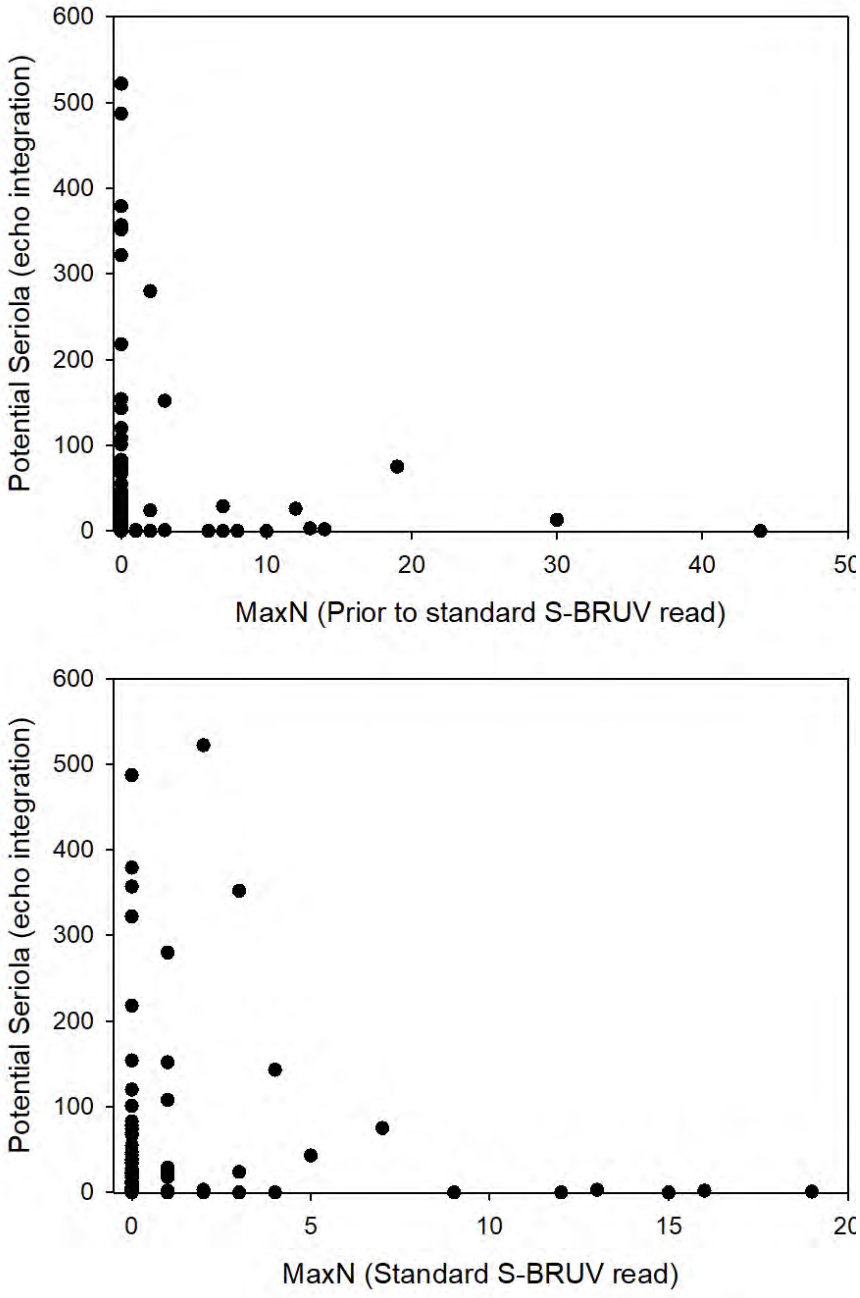
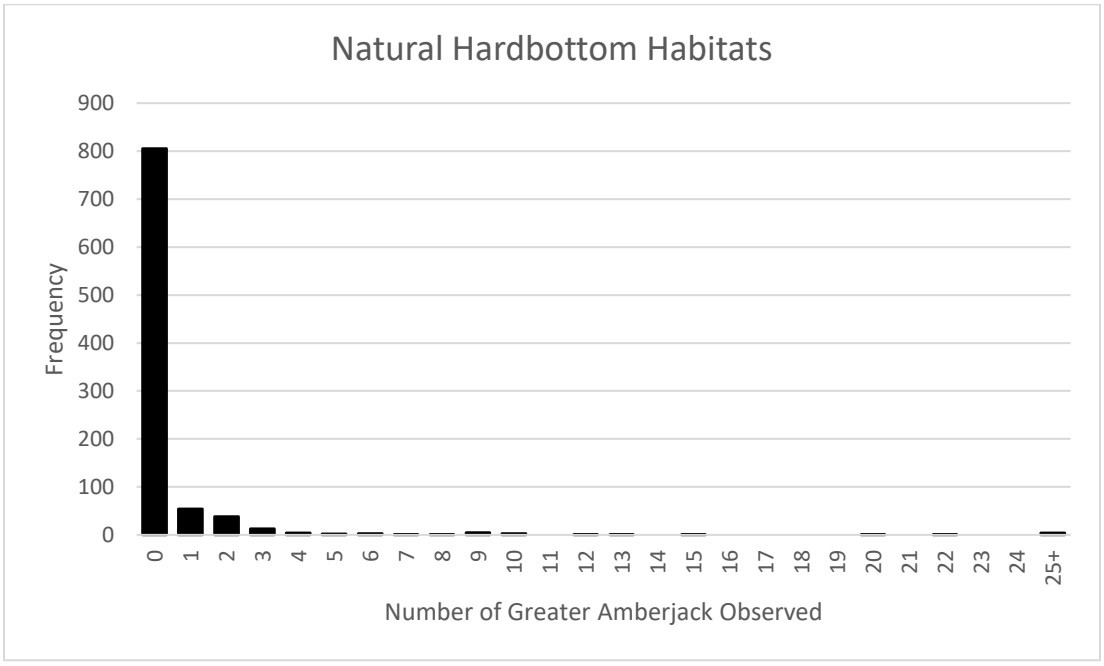


Figure 11.3.40. Frequency distribution of site-specific Greater Amberjack abundance for natural and artificial hardbottom habitats, respectively. Counts are presented in spherical-camera currency, using count calibration formulas where only single-camera count data were available.



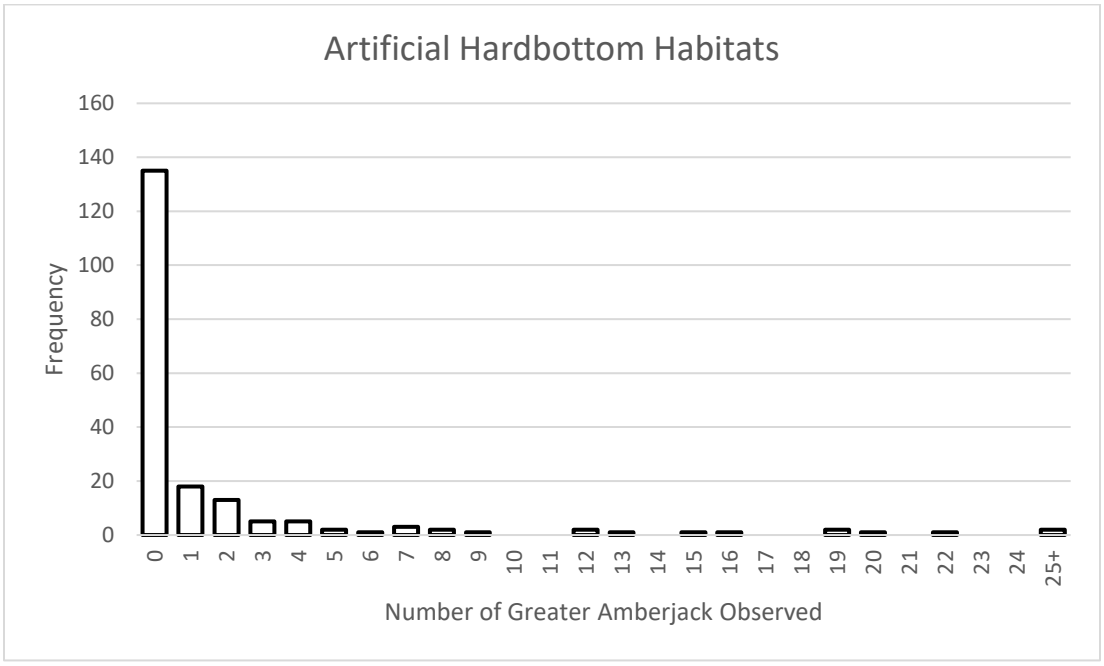


Figure 11.3.41. Frequency distribution of effective reef area sampled during S-BRUV surveys of natural and artificial hardbottom habitats, respectively.

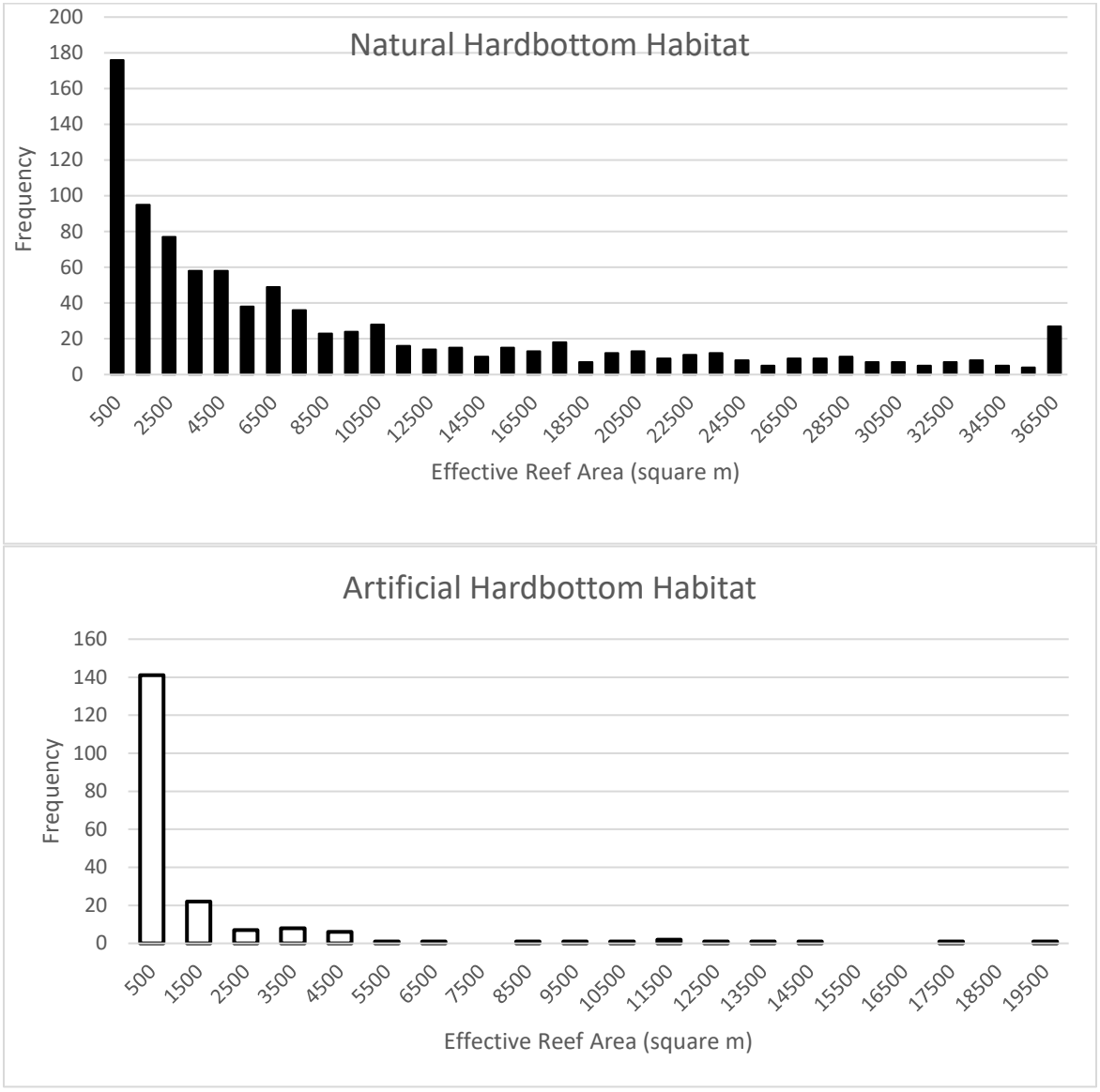


Figure 11.3.42. Frequency distribution of site-specific large-bodied (>20 cmTL), water column (>0.5 m off bottom), disaggregated fish (LBWCF or potential Greater Amberjack) abundance for natural and artificial hardbottom habitats, respectively.

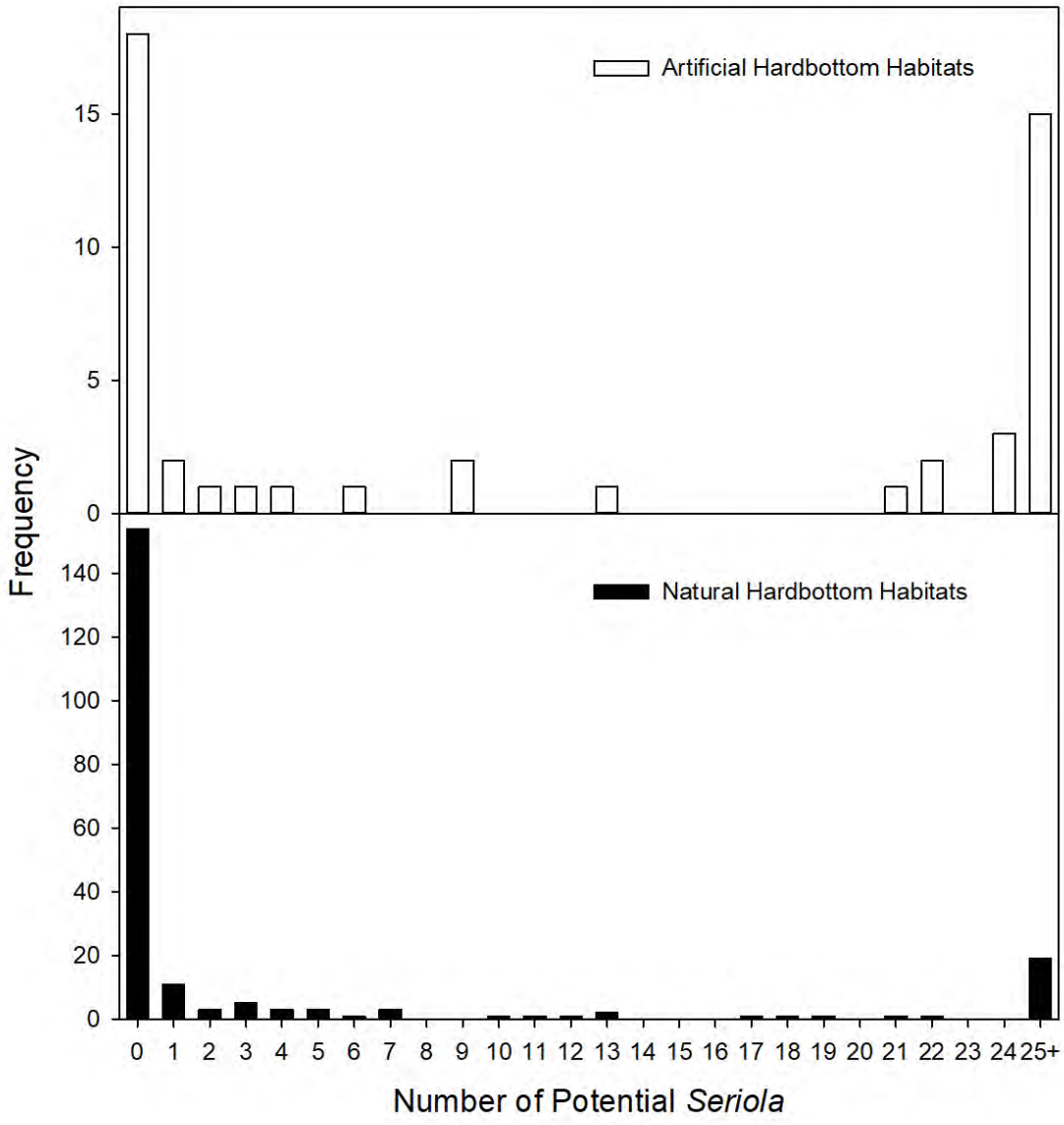
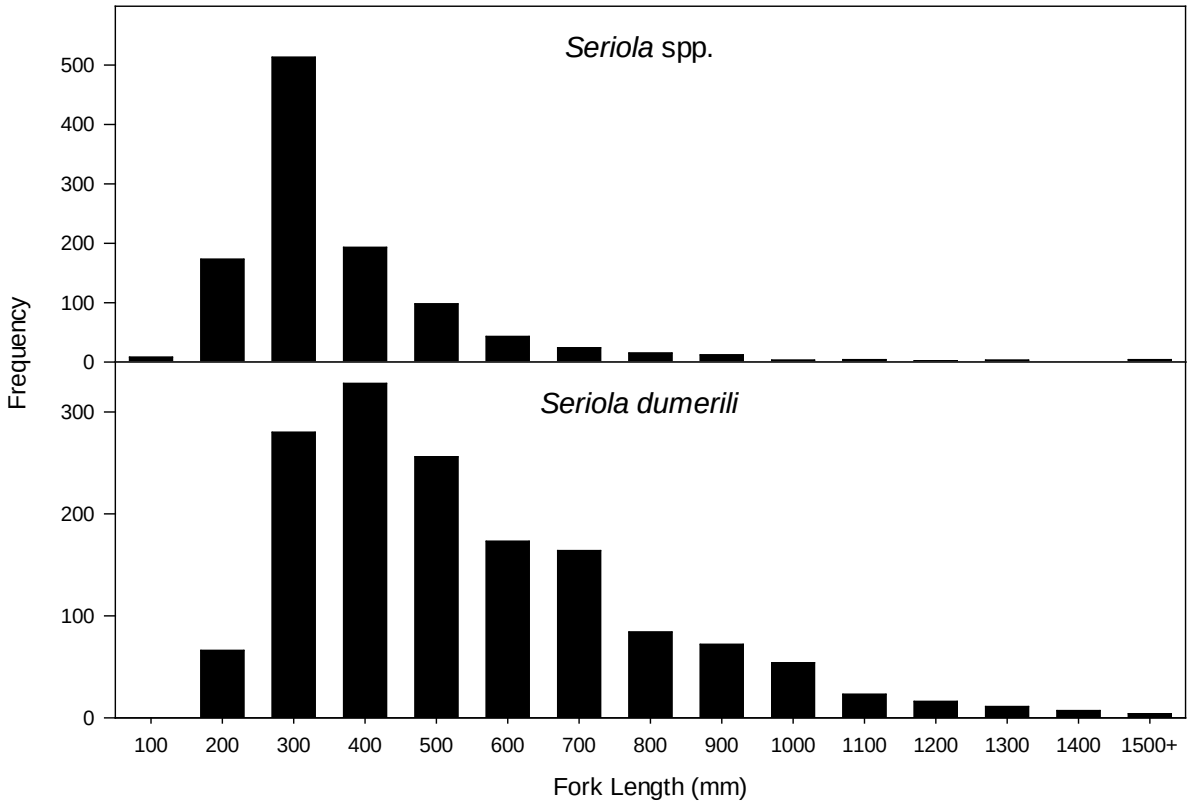


Figure 11.3.43. Historical size composition of unidentified *Seriola* spp. and *Seriola dumerili* from the G-FISHER S-BRUV survey.



References

Campbell, M.D., A.G. Pollack, C.T. Gledhill, T.S. Switzer and D.A. DeVries. 2015. Comparison of relative abundance indices calculated from two methods of generating video count data. Fisheries Research 170:125-133.

Gardner, J. V., P. Dartnell, and K. J. Sulak. 2002. Multibeam Mapping of the Pinnacles Region, Gulf of Mexico. USGS Open-File Report. OF02-006. https://pubs.usgs.gov/of/2002/0006/site/

Gardner C, Goethel DR, Karnauskas M, Smith MW, Perruso L and Walter III JF. 2022. Artificial Attraction: Linking Vessel Monitoring System and Habitat Data to Assess Commercial Exploitation on Artificial Structures in the Gulf of Mexico. Front. Mar. Sci. 9:772292. doi: 10.3389/fmars.2022.772292

Grasty, S. E. 2014. Use of a towed camera system for estimating reef fish population densities on the West Florida shelf. M.S. thesis, x p. University of South Florida, St. Petersburg, Florida.

Hine, A.C., Locker, S.D., 2011. Florida continental shelf: great contrasts and significant transitions. In: Buster, N.A., Holmes, C.W. (Eds.), 101-127, Gulf of Mexico: origin, waters and biota. Volume 3, Geology. Texas A&M University Press, College Station, Texas.

Keenan, S., T. Switzer, A. Knapp, E. Weather, and J. Davis. 2022. Spatial dynamics of the quantity and diversity of natural and artificial hard bottom habitats in the eastern Gulf of Mexico. Continental Shelf Research 233: 104633. https://doi.org/10.1016/jcsr.2021.104633.

Lembke, C., S. Grasty, et al. 2017. The Camera-Based Assessment Survey System (C-BASS): A towed camera platform for reef fish abundance surveys and benthic habitat characterization in the Gulf of Mexico. Continental Shelf Research 151: 62-71.

Liu, Y. and R.H. Weisberg. 2012. Seasonal variability on the West Florida Shelf. Progress in Oceanography. 104: 80-98. http://dx.doi.org/10.1016/j.pocean.2012.06.001

Paxton, A.B and 22 co-authors. 2024. Artificial reef footprint in the United States ocean. Nature Sustainability. Vol. 7, 140 – 147. https://doi.org/10.1038/s41893-023-01258-7

Schobernd, Z.H., N.M. Bacheler, and P.B. Conn. 2014. Examining the utility of alternative video monitoring metrics for indexing reef fish abundance. Canadian Journal of Fisheries and Aquatic Sciences 71:464-471.

Switzer, T.S., S.F. Keenan, K.A. Thompson, C.P. Shea, A.R. Knapp, M.D. Campbell, B. Noble, C. Gardner, and M.C. Christman. 2023. Integrating assemblage structure and habitat mapping data into the design of a multispecies reef fish survey. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 15(4): e10245. https://doi.org/10.1002/mcf2.10245.

Thompson, Kevin A., Theodore S. Switzer, Mary C. Christman, Sean F. Keenan, Christopher Gardner, Katherine E. Overly, Matt Campbell. 2020. Indices of abundance for Greater Amberjack (*Seriola dumerili*) using combined data from three independent video survey. SEDAR70-WP-07. SEDAR, North Charleston, SC. 21 pp.

11.4. South Atlantic

Regional Estimate IV (Southeast FL to NC)

Ted Switzer, Sean Keenan, Julie Gross, Lynne Stokes, John Hoenig, Kevin Boswell, Allison White, Jeff Buckel, Steve Murawski, Josh Kilborn

Introduction

In the South Atlantic region, the estimation of absolute abundance of Greater Amberjack was completed to a large degree by leveraging efforts from the Southeast Reef Fish Survey (SERFS) chevron-video trap (CVT) sampling conducted in 2023. SERFS (Bacheler et al. 2018, 2025) is an annual survey comprised of three programs (the marine Resources Monitoring, Assessment, and Prediction program or MARMAP, the South Atlantic SouthEast Area Monitoring and Assessment Program or SEAMAP-SA, and the SouthEast Fishery Independent Survey, or SEFIS) conducted collaboratively by the South Carolina Department of Natural Resources (SCDNR) and the National Marine Fisheries Service. The survey utilizes baited chevron traps equipped with trap-mounted video cameras to collect data on reef fish abundance and associated habitats. As the SERFS survey is restricted to natural reef habitats at depths of approximately 15 – 100 m between Cape Hatteras, North Carolina and the St. Lucie Inlet, FL, several complementary efforts were conducted to fill important data gaps in the South Atlantic study area. First, supplemental sampling of natural and artificial hardbottom habitats was conducted using stereo-baited remote underwater videos (S-BRUVs) identical to those used in the eastern Gulf of America/West Florida region (Chapter 11.3) to provide data on relative abundance, size composition, and habitat characterization. Second, standardized hydroacoustic surveys were conducted in association with S-BRUV deployments (when conditions allowed for S-BRUV deployment), to produce an estimate of the distribution and abundance of large-bodied pelagic fishes in waters around the S-BRUV (White et al. 2022). And finally, since large-scale hardbottom habitat mapping was unavailable for much of the study area, side scan sonar mapping was completed in conjunction with the S-BRUV sampling around a subset of randomly selected natural and artificial hardbottom sites to validate and adjust estimates of habitat quantity based on existing predictive habitat maps.

To complement the SERFS/S-BRUV survey of natural and artificial hardbottom habitats, a survey of unconsolidated / uncharacterized bottom habitats was conducted using the Camera-Based Survey Assessment System (C-BASS; Lembke et al. 2017; Grasty et al. 2019). A towed underwater video was designed to estimate the distribution and abundance of Greater Amberjack throughout the vast areas of presumably non-reef habitat throughout the South Atlantic shelf waters.

Combined, these research and survey activities were designed to address the following objectives in the South Atlantic (North Carolina/South Carolina and Georgia/East Florida Regions):

(1) Provide data on site-specific abundances of Greater Amberjack and other *Seriola* spp. on natural hardbottom habitats from the SERFS Chevron Video Trap (CVT) surveys.
(2) Provide data on site-specific abundances of Greater Amberjack and other *Seriola* spp. on natural and artificial hardbottom habitats from S-BRUV surveys (where available) and concurrent hydroacoustic surveys.
(3) Provide data on transect-specific abundances of Greater Amberjack and other *Seriola* spp. on unconsolidated / uncharacterized bottom habitats from C-BASS towed camera surveys.
(4) Produce estimates of the absolute abundance of Greater Amberjack by region (NC/SC and GA/FL), habitat (natural and artificial hardbottom habitats and unconsolidated sediment) and size/age class.

Method

Survey design: Survey of natural hardbottom habitats

The survey domain for the South Atlantic region extended from North Carolina south through Miami and encompassed waters from 10 – 150 m deep (Figure 11.4.1). Due to funding limitations, the portion of the South Atlantic south of the Florida Keys was excluded from survey efforts; this decision was informed in part by the general low abundance of Greater Amberjack observations within the reef fish visual census survey (Dr. Robert Muller, Florida Fish & Wildlife Research Institute, personal communication).

The primary source of fishery-independent data on Greater Amberjack was provided by the SERFS survey; as SERFS represents an ongoing, long-term monitoring program, the underlying design could not be modified for this project. SERFS follows a simple random sampling design where a subset of sites containing known hardbottom habitat are selected at random to sample each year; the number of stations available in the sampling frame has increased over time as more natural hardbottom habitats have been identified. For the 2023 sampling season, the sampling frame included 4,370 total natural hardbottom sites; of these, a total of 1,500 stations were randomly selected for sampling (Table 11.4.1; Figure 11.4.2). At-sea operations involve sampling sites in groups of up to six sites concurrently; if less than six randomly selected sites fell within a sampling group, additional nearby stations were often sampled opportunistically (these sites are termed 'nonrandom' in the SERFS database) to increase overall sampling effort during research cruises.

Although the SERFS survey does not incorporate any spatial stratification into their underlying survey design, we first developed a spatial stratification scheme, based on a combination of region and depth, to post-stratify data from the SERFS survey and to stratify supplemental sampling effort. The underlying sampling frame extended from 25° 40' N to 36° 33'N, and was subdivided into six pseudo-states: North Carolina, South Carolina, Georgia, North Florida, Mid-Florida and South Florida (description in chapter 3; Figure 11.4.1). Each pseudo-state was then sub-divided into four depth strata: Inshore (10 – 20m), Shallow (20.1 – 40m), Mid (40.1 – 70m) and Deep (70.1 – 150m) using available bathymetry isobaths imported into a GIS project. Due to a general lack of high-resolution habitat mapping data, the supplemental S-BRUV sampling frame was developed by compiling an inventory of known natural hardbottom habitat points that included both natural hardbottom sites from the SERFS sampling frame (https://seamap.org/seamap-reef-fish-survey/) as well as additional natural hardbottom sites compiled by FWC-FWRI FIM (Florida Fish and Wildlife Conservation Commission – Fish & Wildlife Research Institute's Fisheries Independent Monitoring) for all pseudo-states except for South Florida. These geospatial data were then queried to remove duplicate points as well as all points that were known to contain unclassified or artificial reef habitat, resulting in a final sampling frame of 11,124 known natural hardbottom sites (Table 11.4.2). For the South Florida region (<27.0^{o} lat; Figure 11.4.3), habitat mapping data were available from the Florida Unified Reef Map (https://myfwc.com/research/gis/fisheries/unified-reef-map/); these data were used to develop a hardbottom habitat sampling frame. Available habitat data represented classified polygons of reef/hard bottom features designated within the Geoform field as "Coral Reef and Colonized Hardbottom" to indicate natural habitat present. The Source field within the Unified Reef Map shapefile stated that "SE Florida NSU, v2015" as the origin of the data interpretation. Following an approach like that used in the eastern Gulf, the South Florida region was gridded with a 0.1 x 0.1nm framework and intersected with the habitat polygons as well as artificial reef habitat points from the inventory (Figure 11.4.4). The number of grids containing habitat were summarized in Table 11.4.2.

Supplementary sampling of natural hardbottom habitats with S-BRUV arrays and acoustics was conducted within all regions, and within all depth strata, except for the Mid depth stratum, which was only sampled within

the South Florida region due to the high percentage of SERFS sampling effort that falls within this stratum throughout the remainder of the South Atlantic survey domain. Based on historical sampling efficiency in the eastern Gulf, we estimated that we would be able to sample a total of 234 natural hardbottom sites given available ship time (Figure 11.4.5; Table 11.4.2). Sampling effort was allocated disproportionately among three spatial surveys: Inshore/Nearshore depth strata excluding South Florida (N = 131); Deep depth stratum excluding South Florida (N = 65); and the South Florida region (N = 38). Sampling within the Inshore/Nearshore depth strata was conducted primarily to quantify differences in Greater Amberjack abundance between natural and artificial hardbottom habitats and, secondarily, serve as an indirect method to calibrate counts between S-BRUV and SERFS data processing protocols. Sampling within the Deep depth stratum was conducted at a higher sampling intensity because the SERFS survey does not cover the full range of depths for Greater Amberjack. The South Florida region was treated independently because it is not covered by the SERFS survey, and the underlying sampling domain was developed using habitat mapping data, and not individual points. Although allocation of sampling effort among these three spatial surveys was disproportionate, allocation of effort within each survey was generally proportional to the habitat (defined as the total number of available natural hardbottom sampling sites within each spatial stratum for all regions except South Florida; for South Florida, habitat availability was defined as the total number of microgrids containing natural hardbottom habitat within each stratum), ensuring that a minimum of three sites were selected within each spatial stratum where available.

Survey design: Survey of artificial hardbottom habitats

The SERFS CVT survey does not sample artificial hardbottom therefore sampling was only conducted using S-BRUV and acoustics. To develop the artificial hardbottom sampling frame, an inventory of artificial reef habitat points was assembled by compiling information from the FWC-FWRI DMFM (Division of Marine Fisheries Management) database and from Dr. Avery Paxton (NOAA National Centers for Coastal Ocean Science (NCCOS). All point data were given unique ID/names and summarized by spatial strata (Figure 11.4.6; Table 11.4.3). The distribution of known artificial reef sites was heavily skewed towards the Inshore and Shallow depth strata; therefore, artificial hardbottom sampling was restricted to these two strata. Artificial hardbottom sampling sites were randomly selected with proportional allocation among spatial strata proportional to habitat availability (defined as the total number of available artificial hardbottom sampling sites within each spatial stratum), ensuring that a minimum of three sites were selected for each spatial stratum. Based on historical sampling efficiency in the eastern Gulf, we estimated that we would be able to sample a total of 150 artificial hardbottom sites given available ship time (Table 11.4.3).

Survey design: Survey of unconsolidated / uncharacterized habitats

The South Atlantic survey domain for the survey of unconsolidated or uncharacterized habitats extended from the North Carolina – Virginia Border south to 27° N, including waters 20 – 150 m deep. Unlike the eastern Gulf of America, where the sampling frame was subdivided into 5’ by 5’ blocks, sampling in the South Atlantic was designed based as proportional allocation of sampling effort among subregion (NC: North Carolina; SC: South Carolina; GA: Georgia; NF: North Florida; MF: Mid Florida) and depth (Shallow: 20 – 40 m; Mid/Deep: 40 – 150 m). Unlike the stationary video surveys of hardbottom habitats, the survey of unconsolidated habitats was not conducted within the inshore stratum (10 – 20 m), or in South Florida (south of 27° N).

To inform the sampling design for the C-BASS survey, estimates of total sampling area within each subregion and depth stratum were generated in a geographic information system (Table 11.4.4). It was estimated that a total of 60 C-BASS transects could be conducted during the study, and the location of each of these was

proportionally allocated (based on area) and randomly selected using a geographic information system (Table 11.4.5). For each randomly-selected site, a randomly-selected direction (0 – 359°) was determined for each C-BASS tow.

Existing habitat knowledge / data availability

Similar to the West Florida region, a comprehensive high-resolution map of hard bottom habitats (and thus unconsolidated sediments) does not exist for the South Atlantic survey domain (including the NC/SC and GA/FL regions). Nevertheless, prior efforts have been made to synthesize existing information to generate estimates of natural (Steward et al. 2022) and artificial hard bottom habitat coverage (Steward et al. 2022; Paxton et al. 2024) in the region. Steward et al. (2022) integrated several existing sources of natural and artificial hardbottom data and used multiple approaches to generate independent estimates of region-wide habitat coverage. For natural hardbottom habitats, total estimates ranged from 1,506 – 3,602 km^2, with an average of 2,654 km^2 in natural hardbottom habitat coverage. For artificial hardbottom habitats, total estimates ranged from 2 – 4 km^2, with an average of 3 km^2 artificial hardbottom habitat coverage. While this study filled important information gaps in the availability of hardbottom habitats in the southeastern U.S., it also highlighted the continued need for both broadscale and fine scale mapping efforts throughout the region.

Field methods: SERFS chevron-video trap (CVT) survey

At each sampling site, a 1.7 x 1.5 x 0.6m chevron trap (Figure 11.4.7) was baited with 24 menhaden (*Brevoortia* spp.) and deployed for a targeted soak time of 90 minutes. All traps were equipped with a GoPro Hero 4 camera over the trap entrance to quantify fish abundance, and a Nikon Coolpix S210/S220 camera mounted over the nose of the trap to characterize microhabitat features in the opposite direction. Traps were typically deployed in groups of six, and cameras were turned on and set to record before trap deployment and turned off after trap retrieval.

Field methods: S-BRUV survey

Supplemental sampling of both natural and artificial hardbottom habitats was conducted using an S-BRUVs array consisting of five independent stereo-video recorders mounted orthogonal to each other inside a 0.8m wide × 1.0m tall aluminum frame (Figure 11.4.7). Cameras were positioned to provide a near-360-degree view of the surrounding seafloor. Two of the five cameras were paired with a second satellite camera (one at a baseline of 231 mm and the other at a baseline of 300 mm) mounted in parallel to provide stereo video measurements; each S-BRUV was calibrated before and immediately after the sampling season to ensure accuracy of stereo-video measurements. All cameras were synched and recorded with an 86° horizontal field of view and 61° vertical field of view at a pixel resolution of 1,920 H x 1,200 V. Data were recorded to an onboard computer system which also provided power via an external battery. An In-Situ© Aqua Troll 500 water quality datasonde was mounted to each S-BRUV to record temperature (°C), salinity (psu) and dissolved oxygen (mg L^{-1}). Each S-BRUV was baited with four mackerel (*Scomber* spp.) halves (approx. 450 g) and approximately 225 g of squid (*Illex* sp.) and deployed for 35 minutes prior to retrieval. All sites were sampled with S-BRUVs when conditions allowed, although in some instances, current velocity was too strong for safe S-BRUV deployment and retrieval.

A paired hydroacoustic survey was conducted at all sites concurrent to the S-BRUV deployment. Full details of hydroacoustic sampling methods are provided in Chapter 9, but in brief, hydroacoustic equipment was deployed using either a towed, pole-mounted, or hull-mounted system depending on the vessel. Regardless of deployment method, a combination of three (38 kHz, 70 kHz, 120 kHz) Kongsberg EK 80 wideband autonomous transceivers was used to map the three-dimensional distribution of fish surrounding the S-BRUV array. At each

site, the hydroacoustic survey consisted of a series of seven 200-m parallel transects; the center transect passed over the site of the S-BRUV deployment, while each successive parallel transect was increasingly further away from the center line (two transects at a targeted 7-m distance, two transects at a targeted 18-m distance, and two transects at a targeted 35-m distance). If the S-BRUV array could not be deployed due to high current velocity, sites were sampled with hydroacoustics only.

Field methods: C-BASS sampling

Unconsolidated sediment habitats were sampled using the Camera-Based Assessment Survey System (C-BASS; Grasty 2014; Lembke et al. 2017) which consisted of a series of six high-definition video cameras used in stereo and individually (Figure 11.4.7). The system also has a range of environmental sensors (temperature, depth, salinity, CDOM) and navigation-related sensors (e.g., pitch, roll, yaw, altitude). The system is "flown" in real time using imagery supplied from the tow sled to the vessel deck. Depth alterations are done via adjusting the towing cable length to affect the height over bottom. The system is generally flown at an altitude of 2-5 m above the sea floor, depending on water visibility and bottom topography, and is towed at speeds of 1.5 – 2 m s^{-1}. All efforts were made to sample the selected transects as planned; however, unforeseen constraints to sampling (e.g., presence of oil rigs or obstacles in the transect path, unsafe proximity to high traffic areas, presence of actively fishing vessels) resulted in transects that were either cut short of the 6.5 nm target, moved to an adjacent grid cell, or towed in an alternate randomly selected compass direction.

Data processing methods: SERFS chevron-video trap (CVT) survey

All SERFS videos were read during an interval of 20 total minutes, beginning 10 minutes after the trap landed on the seafloor to allow time for the trap to settle. Relative abundance of reef fish is estimated using the SumCount approach where one-second video snapshots are read every 30 seconds for the 20-minute read time interval, totaling 41 snapshots read for each video. For each taxon, all individuals observed across the 41 snapshots are summed. To maintain comparability to data provided by the supplemental S-BRUV survey, a MaxN proxy was extracted from the SERFS video dataset, which represents the maximum number of individuals observed across all 41 snapshots. Video readers also conduct a characterization of observed microhabitat features, including measures of biotic density and substrate composition. Water clarity was also scored for each sample as poor (bottom substrate could not be seen), fair (bottom habitat could be seen but the horizon was not visible), or good (the horizon was visible). A CTD cast was also taken for each simultaneously deployed group of traps.

Data processing methods: S-BRUV survey

For all S-BRUV deployments, the relative abundance of all fishes was estimated as MaxN, or the maximum number of individuals of each taxon (note that counts from different taxonomic levels are not mutually exclusive, for example *Seriola dumerili* and *Seriola* spp., so are not additive) observed in a single screen frame to avoid the possibility of double-counting. Full spherical camera reads were conducted for all sites, and MaxN was limited to counts of 299 or less, as beyond 299, it is difficult to estimate the number of individuals observed accurately. Each video was viewed for a total of 20 minutes, beginning approximately 10 minutes after the gear reached the bottom; in addition, MaxN of *Seriola* spp. was determined prior to the standard read time (e.g., during the camera drop) to provide information on the presence and abundance of *Seriola* spp. outside of the standard read time. All fishes were identified to the lowest possible taxonomic level, most commonly species. A variety of metrics of habitat quality were also characterized for each of the five cameras, including various measures of relief, substrate coverage and composition, biota coverage and composition, and habitat complexity among others.

Stereo-video measurements were obtained for a subset of taxa, including all managed species as well as any species that attains a maximum size of greater than 300 mm. Measurements were taken at a single point during the video in which the greatest number of individuals were measurable (both the snout and tail were visible on both cameras and the individual was generally oriented perpendicular to the video camera). Stereo-video measurements represent fork length (FL, in mm). Often only a subset of individuals observed on video were measurable, and there are instances where no individuals were measurable due either to stereo camera malfunction or fish orientation.

Data processing methods: C-BASS surveys

Imagery from the C-BASS video transects were processed in 15 s intervals, although scrolling a few sections in each direction was allowed to provide context and ensure that classification adequately characterized the area. Observed habitats were classified according to the substrate and biotic components of the Coastal and Marine Ecological Classification Standard (CMECS) scheme. Characterized metrics included seafloor induration, cover and composition of primary attached biota, and primary and secondary components of hardbottom and softbottom habitat, respectively. In addition, all observed fish were enumerated and identified to the lowest possible taxon; if no taxonomic distinction was possible, observed individuals were classified as unidentified small and large fishes, depending on size. Positional data, including roll, pitch, yaw, altitude, and tow speed were analyzed to estimate area viewed by each transect.

Habitat calculation methods

Acknowledging both the value and limitations of the Steward et al. (2022) study, we developed an alternate approach to estimate habitat coverage in the southeastern U.S. survey domain. The foundation of our analyses was the Southeast Conservation Adaptation Strategy (SECAS) Atlas; Atlantic Coral and Hardbottom (Southeast Blueprint Indicator; https://secas-fws.hub.arcgis.com/maps/fws::atlantic-coral-hardbottom-southeast-blueprint-indicator-2024/about) raster dataset which included complete coverage of the entire South Atlantic survey domain (Figure 11.4.8). The SEACAS Atlas was developed through analyses of multiple input mapping datasets, including the South Atlantic Bight Marine Assessment synthesis developed by the Nature Conservancy that was also a key component of the Steward et al. (2022) study (Conley et al. 2017). The outcome of the SECAS Atlas was a raster data layer (30m pixels) where each pixel was assigned an indicator value representing relative likelihood of containing natural or artificial hardbottom habitats (Table 11.4.6 and 11.4.8). There were nine indicators in the original Atlas output, but the cold-water coral mounds class did not occur within the bounds of this survey and was not included. Further the Wrecks and Artificial Reef indicators were combined into one class. The final list of seven indicators used here included: Medium Probability of hardbottom, High Probability of hardbottom, Highest Probability of hardbottom, Artificial reef habitat, Confirmed natural hardbottom, Confirmed natural hardbottom with associated benthic fauna, No habitat prediction.

Rather than directly summarizing spatial coverage by indicator value, a limited number of habitat mapping surveys were conducted to validate habitat coverage and generate habitat proportion correction factors for each indicator class. Locations of habitat mapping surveys were randomly selected from the same universe of natural and artificial hardbottom sampling sites used for regional camera surveys. At each randomly selected location, habitat-mapping surveys were conducted using an L3-Klein 3900 side scan sonar operating at 445 kHz. All surveys included three passes with 10% overlap to correct for potential positional inaccuracies. For surveys centered on natural hardbottom habitats, each pass covered approximately 4 nm for a total coverage of approximately 3.5 km^2. For surveys centered on artificial hardbottom habitats, each pass covered approximately 0.5 nm for a total coverage of approximately 0.56 km^2. Side scan sonar data were post-processed following

identical methods to those used in the Gulf where corrections were made for navigation offsets, layback of the towfish, and time varied gain. Sonar imagery (geotiffs) were exported at 0.25m resolution and imported into a GIS project. Individual reef features were identified and classified through interpretation of the raw imagery, and polygons for each feature were manually digitized as a vector layer superimposed on the geotiff within GIS (Figure 11.4.9).

To develop correction factors for the SECAS Atlas indicators, we first clipped the SECAS raster to the spatial footprint of each side scan sonar survey. Within each side scan survey, we overlaid both the SECAS Atlas indicator data and the habitat polygons generated from interpretation of the side scan sonar mapping data (considered "truth"). For each indicator value, we calculated (1) the total coverage within the side scan sonar survey footprint, and (2) the proportion of this area that was confirmed as hardbottom habitat. Separate analyses were conducted for the natural and artificial reef surveys, respectively. We then calculated average proportion of confirmed habitat for each SECAS Atlas indicator by weighting by the total area of that indicator within individual surveys (giving more weight to surveys where an indicator had more extensive coverage). Due to the limited number of surveys conducted, analyses were conducted across the entire survey domain, acknowledging that these values likely vary spatially. Indicator-level habitat proportions (Table 11.4.6 and 11.4.8) were then used to estimate hard bottom habitat coverage for all subsequent analyses.

Analytical methods: Developing calibration coefficients between SERFS and S-BRUV reads

While MaxN (or a MaxN proxy) were produced from both the SERFS and S-BRUV surveys, MaxN values from the two surveys are not directly comparable. The MaxN proxy from the SERFS survey was obtained from a single camera with a field of view of 123°, whereas the MaxN obtained from the S-BRUV survey was obtained from a camera with a full spherical field of view. In the eastern Gulf, a Bayesian Poisson regression was applied to convert single-camera counts to spherical-camera counts (Chapter 11.3; Figure 11.3.34). Unfortunately, this approach could not be applied to convert SERFS MaxN values to S-BRUV MaxN currency because of differences in the single-camera field of view between SERFS and the S-BRUV cameras (86°). With very limited paired data from the Alabama calibration studies, we developed an alternative approach to convert SERFS MaxN values to S-BRUV MaxN currency. We first restricted both the SERFS and S-BRUV natural reef survey data to a common spatial footprint, including data from the Inshore and Nearshore depth strata, and from all regions except South Florida. As these data were collected following very similar designs and represent the same population, we then generated cumulative frequency distributions for each survey independently (Figure 11.4.10). We then converted all SERFS MaxN values into S-BRUV spherical MaxN values based on their corresponding quantiles. When quantiles fell between two MaxN integer values, the calibrated value was conservatively rounded down to the nearest integer.

Analytical methods: Estimation of absolute abundance of Greater Amberjack – camera-based estimate

Site-specific abundance:

The abundance of Greater Amberjack at each sampling site was estimated in spherical camera abundance currency. For sites where spherical camera reads were conducted, site-specific abundance represented MaxN from the spherical camera reads. For sites sampled by the SERFS survey, the SERFS MaxN proxy was converted to an S-BRUV spherical camera count where only single camera reads were conducted, single camera counts were converted to spherical camera counts using the comparative quantile method as described previously.

Effective sampling area of reef habitat:

Analysis of videos yields a qualitative measure of relative visibility which can be used to estimate the area of the seafloor viewed for each camera site (on the order of meters). However, it is likely that some proportion of individuals observed originate from areas beyond that which can be directly observed on video. As the SERFS and supplemental S-BRUV surveys all use baited camera systems, there is likely some degree of directional bait attraction which would attract fish from an unknown distance. It is also likely that there is a measure of reactiveness to gear moving through the water. Accordingly, using the area observed by the camera as a denominator for density estimation would likely result in a significant overestimation of true Greater Amberjack density at each sampling site.

Following the approach developed for estimating effective sampling area in the eastern Gulf, we chose to apply an effective sampling area of 36,644 m^2 for subsequent analyses. Within the effective sampling area for each site, both reef and non-reef habitats are often present. As the objective of the S-BRUV survey was to estimate Greater Amberjack abundance on natural and artificial hardbottom habitats, we calculated the total area of reef habitat within the effective sampling area (assuming that the contribution of non-reef habitats to observed Greater Amberjack abundances was minimal). We first clipped the Southeast Conservation Adaptation Strategy Atlas; Atlantic Coral and Hardbottom (Southeast Blueprint Indicator) (https://secas-fws.hub.arcgis.com/maps/fws::atlantic-coral-hardbottom-southeast-blueprint-indicator-2024/about) raster dataset predictive habitat map within a 108 m radius for each sampling site. We then calculated the total amount of habitat within each predicted habitat class that occurred within the 108 m buffer. These estimates were then adjusted by the correction factors developed using random habitat mapping surveys to estimate the total amount of reef habitat that occurred within the 108 m buffer. Results from these analyses produced estimates of natural, artificial, and total hardbottom habitat within the effective sampling area of each camera site, and these estimates of reef habitat coverage, not the total effective sampling area of 36,644 m^2, were used in calculations to convert site-specific Greater Amberjack counts to densities.

Estimating site selection probabilities:

Site selection probabilities for all surveys and strata are provided (Tables 11.4.1; 11.4.2; 11.4.3). The SERFS survey followed simple random selection from a universe of points. The S-BRUV survey on natural and artificial habitats followed a proportional allocation by number of sites available within each spatial stratum. Proportion of total sites for each stratum were multiplied by the anticipated number of sites to be sampled by S-BRUVs.

Estimating Greater Amberjack size composition:

Size data for Greater Amberjack were available only from the S-BRUV surveys equipped with stereo-video cameras. Not all Greater Amberjack observed were able to be measured due to factors such as distance from the lenses, excessive angle or turbidity. Greater Amberjack were measured from both natural and artificial hardbottom habitats when possible, and data were summarized to generate habitat-specific size frequency distributions for each region.

Estimating Greater Amberjack abundance:

A design-based estimate of Greater Amberjack absolute abundance was generated from an analysis of S-BRUV (natural and artificial reef habitats), SERFS trap-video data (natural only) and C-BASS data in the South Atlantic. An overview of general analytical methods was provided in Chapter 10.

Analytical methods: Estimation of absolute abundance of Greater Amberjack – C-BASS estimate

Site-specific abundance and sample probability:

Any 15 sec video segment with positive observation of Greater Amberjack was evaluated for habitat to confirm association with unconsolidated sediment. The metric for site specific abundance was the number of observed Greater Amberjack within the area encompassed by the 15 sec video segment to generate a density (num. / m^2). Stratum-specific densities were then calculated that accounted for site selection probabilities.

Estimating Greater Amberjack abundance:

Final stratum-specific abundance estimates were generated by extrapolating density estimates to the total estimated are of non-reef habitats as generated from interpretation of the mapping data (Table 11.4.4).

Analytical methods: Estimation of absolute abundance of Greater Amberjack – acoustics-based estimate

Data from paired camera and hydroacoustic surveys were integrated to develop an acoustics-based estimate of Greater Amberjack. Details of estimation methods follow below.

Acoustic-based estimates of potential Greater Amberjack abundance:

Hydroacoustic data were processed according to protocols described in full detail in Chapter 8. These analyses produced estimates of potential *Seriola*, or all targets thought to be greater than approximately 200 mm FL, observed more than 0.5 meters above the sea floor, and not tightly aggregated. These estimates likely contain targets from large-bodied fishes other than Greater Amberjack; therefore, estimates of potential Greater Amberjack then need to be corrected based on species composition from coincident video surveys.

Estimating Greater Amberjack partitioning ratios:

As acoustic estimates of potential *Seriola* include all large-bodied fishes, we generated site-specific estimates of the percentage of large-bodied fishes observed that were confirmed to be Greater Amberjack. Estimates of total large-bodied fishes per site include all species that routinely reach sizes of 200 mm or greater. Notable exceptions to this included two species (Blue Runner and Vermilion Snapper) who often occur in large numbers, a significant percentage of which are small individuals. For these two species, we calculated the percentage of individuals measured that were greater than 200 mm, then adjusted counts by multiplying by this percentage. Stereo measurements are often not possible for a particular sampling site, therefore these corrective percentages were calculated and applied globally.

Estimating Greater Amberjack abundance:

Because few sites had both observed Greater Amberjack on video and potential *Seriola* identified via acoustics, we needed to aggregate camera data to estimate the percentage of potential *Seriola* that were in fact Greater Amberjack. For all sites where the number of large-bodied fishes observed was > 30 individuals, we used site-specific correction ratios. For other sites, we aggregated by spatial strata (collapsing over all habitat strata).

For each acoustics survey, estimates of the total area swept were calculated. As surveys cover both reef and non-reef habitats, we estimated the total reef area within the acoustics survey by multiplying the total area swept by the estimated percentage of sampling area comprised of natural reef habitat as determined from the 108 m buffer analyses. These data were then used to generate design-based estimates of Greater Amberjack.

Results

Survey accomplishments: Survey of natural hardbottom habitats (SERFS)

During 2023 sampling, the SERFS survey sampled 979 of the 1,500 randomly selected sites (65.3%; Table 11.4.1). Sites were distributed across the two GAJ regions as well as all depth strata (Figure 11.4.2). Greater Amberjack

were observed in 23.5% of all deployments with frequency distribution patterns similar between regions (Figure 11.4.11). After adjustment into S-BRUV MaxN currency, a total of 1,004 Greater Amberjack were observed over all deployments. The greatest number observed was 77 from a single deep site in Georgia, however most of the positive observations were from one or two individuals.

The effective sampling area of reef habitat within the 108m buffer of each deployment ranged from 896 m^2 to a maximum of 10,009 m^2 as estimated from the Southeast Conservation Predicted Habitat map. The majority of sites (~55%) were estimated to have approximately 5,000 m^2 (13%) of hard bottom within the buffer area (Figure 11.4.12) and this pattern was consistent between the NC/SC and the GA/FL regions.

There were 482 deployments completed in the NC/SC region and there were 389 Greater Amberjack observed. Most observations occurred within the mid-shelf depth stratum in South Carolina (Figure 11.4.13) followed by the shallow stratum. From North Carolina, mean CPUE values (individuals per site) were greatest from the shallow depth stratum and decreased in deeper strata. No Greater Amberjack were observed from the inshore depth stratum in Georgia/Florida region and the greatest mean CPUE was from the Georgia deep stratum (Figure 11.4.14). Mean CPUE were similar among the other depth strata from the other regions.

Survey accomplishments: Survey of natural hardbottom habitats (S-BRUV survey)

Of the 234 natural hard bottom sites selected for sampling with S-BRUVs, 140 sites were completed and 22 sites were only sampled with hydroacoustics due to high current conditions that prevented S-BRUV deployment (Table 11.4.2; Figure 11.4.5). Greater Amberjack were observed in 22.8% of S-BRUV deployments on natural hardbottom sites with most observations comprising of one or two individuals (Figure 11.4.15). A total of 144 Greater Amberjack were observed from S-BRUV deployments on natural hardbottom habitats with one observation of 46 fish from a South Carolina site.

The effective sampling area of reef habitat within the 108m buffer of each deployment ranged from 896 m^2 to a maximum of 24,975 m^2 as estimated from the Southeast Conservation Predicted Habitat map. The amount of reef area varied for the S-BRUV sites with many sites in the 5,000 m^2 (Figure 11.4.16) range which was similar to the SERFS survey. While the distribution was similar between the NC/SC and GA/FL regions, the GA/FL region had more sites with less effective sampling area containing natural hardbottom.

There were 39 S-BRUV deployments within the NC/SC region on natural hardbottom habitats and nine of these (23.1%) observed Greater Amberjack. The greatest CPUE was from the shallow (20 – 40.1m) depth stratum and none were observed from inshore strata (no sampling occurred in mid-shelf strata) (Figure 11.4.17). Of the 101 S-BRUV deployment in GA/FL region, 20 observed Greater Amberjack (19.8%) from natural hardbottom sites. Greater Amberjack were observed from the shallow stratum as well as inshore in Georgia and in the deep stratum in North Florida (Figure 11.4.18), however no sampling was conducted in the mid-shelf stratum.

Survey accomplishments: Survey of artificial hardbottom habitats (S-BRUV survey)

Of the 148 selected artificial hardbottom sites selected, a total of 78 sites were successfully sampled in the South Atlantic (Table 11.4.3; Figure 11.4.6). Greater Amberjack were observed from 22 S-BRUV deployments (28.2%) on artificial hardbottom and many of the non-zero counts were large numbers (Figure 11.4.19). A total of 291 Greater Amberjack were observed from S-BRUVs with an observation of 86 individuals from a North Florida inshore site.

The effective sampling area of artificial reef habitat within the 108m buffer ranged from 0 to 1,259 m^2 of predicted artificial hardbottom habitat. Not all artificial reefs available within the selectable sites were represented in the Southeast Conservation Predicted Habitat map therefore did not have an area available. Of

all sampled sites, 65 contained 1,259 m^2 of predicted artificial habitat which comprised 3% of the buffer area (Figure 11.4.20).

Greater Amberjack were present in eight of the 28 sites completed from NC/SC region and the greatest CPUE was from the shallow stratum for both North and South Carolina (Figure 11.4.21). No sampling on artificial reefs occurred in the mid-shelf or deep strata for NC/SC. The greatest CPUE for the GA/FL region was observed from inshore stratum in North Florida driven primarily by the n=86 observation. Greater Amberjack were observed on artificial reef habitats in the majority of the sampled strata (Figure 11.4.22).

Survey accomplishments: Survey of unknown/uncharacterized habitats (C-BASS)

A total of 31 towed C-BASS transects were conducted on unconsolidated sediment habitats in the South Atlantic waters (Table 11.4.5; Figure 11.4.23). These transects covered a total of 316,668 m (average distance per transect of 10,556 ± 2,186 m) and sampled a total of 4,431,795 m^2 (average area per transect of 147,723 ± 43,012 m^2). Among all C-BASS transects conducted in the South Atlantic, *Seriola dumerili* were only observed once in association with unconsolidated sediment habitats; one individual was observed in the shallow depth stratum in South Carolina (Figure 11.4.23).

Results: Estimation of absolute abundance of Greater Amberjack – camera-based estimate

Estimates were generated following a similar approach to that developed for West Florida (Figure 11.3.13) although, we integrated data from SERFS videos and S-BRUV video data for one estimate.

Site-specific abundance

The process used to determine abundance for the South Atlantic regions followed a nearly identical process to the west Florida/eastern Gulf chapter (Chapter 12.3). For the SERFS videos as well as the S-BRUV videos, we used the standard read data which occurs after the lander reaches the seafloor and the silt plume dissipates. For both gear types, a standard spherical abundance metric was used following a conversion for SERFS videos as described above. For natural hardbottom habitats, counts ranged from 0 – 77 Greater Amberjack per site, with 781 sites where Greater Amberjack were not observed. For artificial hardbottom sites, counts ranged from 0 – 86 individuals per site with 56 sites where Greater Amberjack were not observed.

Effective area of reef habitat sampled

Estimates for reef habitat were derived from evaluating a predicted hardbottom habitat map to habitat data collected from small-scale, high resolution side scan sonar surveys conducted during the Greater Amberjack project. The frequency distribution of the effective area of reef habitat that fell within the 108 m buffer was presented for natural hardbottom habitat from the SERFS survey (Figure 11.4.12), natural hardbottom habitat from the S-BRUV survey (Figure 11.4.16) and artificial hardbottom from the S-BRUV survey (Figure 11.4.20). For natural reef sampling sites, the effective area of natural reef habitat in proximity to the S-BRUV array ranged from 0 to 24,975 m^2, with an average of 4,946 m^2 for S-BRUVs and from 0 to 10,202 m^2, with an average of 3,968 m^2 for SERFS videos. For artificial reef sampling sites, the effective area of artificial reef habitat in proximity to the S-BRUV array ranged from 0 to 1,259 m^2, with an average of 1,114 m^2.

Estimating site selection probabilities

Site selection probabilities for natural SERFS, natural S-BRUV and artificial hardbottom sampling strata are provided in Tables 11.4.1, Table 11.4.2 and 11.4.3, respectively.

Estimating natural and artificial habitat coverage

For each pseudo-state and depth bin strata, the weighted mean ratio was multiplied by the area of each predicted habitat confidence bin within the spatial strata bounds for that pseudo-state (Table 11.4.6 and 11.4.8).

While the greatest weighted ratio was in the Confirmed natural hard bottom bin (68% of predicted area contained hard bottom as identified within side scan surveys), overall contribution of this bin was low due to lower area within pseudo-states with the exception of South Florida. The predicted bin termed "Highest probability of hard bottom" had a 13% weighted ratio with side scan identified habitat, however it covered such a large area, that it contributed to most of the overall habitat coverage for mid-Florida, north-Florida and Georgia. For natural habitats, the estimated total was 3,250.8 km^2, artificial habitats was 2.446 km^2 and uncharacterized habitats was 42,459.57 km^2 (Table 11.4.7).

Similar to Florida/Georgia, the predicted bin termed "Highest probability of hard bottom" contributed the most to aerial coverage with over a square km in both North and South Carolina states (Table 11.4.8). The final estimated total of natural hard bottom was 4,529.2 km^2, artificial reef habitat was 1.201 km^2 and uncharacterized habitat was 59,485.98 km^2 (Table 11.4.9).

Estimating Greater Amberjack abundance

Camera-based estimates of abundance of *Seriola dumerili* were made using the stratified ratio estimator with area-based sampling units. The count values used for abundance were observed MaxN spherical camera counts on sampled sites from S-BRUV deployments and proxy MaxN from SERFS trap-camera deployments. The sample base weights used in the estimator was the reciprocal of the sampling probability. The sampling probability was the number of sites selected to sample over all available sites within the spatial stratum except for the south Florida strata. In south Florida, the sampling probability was the number of microgrids selected to sample for the specific habitat divided by the total number of microgrids containing that habitat within the specific spatial/depth stratum. Estimates and their standard errors were computed using the R package *survey* (10.32614/CRAN.package.survey), which appropriately handles variance estimation for complex sample designs.

Estimates of abundance for natural and artificial habitat from the camera count data from the use of buffers of 108 m are shown in Tables 11.4.10 and 11.4.11. Each table displays the estimates aggregated to the spatial stratum level, as well as the overall estimate of *Seriola dumerili* for natural and artificial habitats, respectively. There were 855,267 (± 354,455) *Seriola dumerili* estimated from natural habitats in the Georgia/Florida region while 8,692 (± 5,678) were estimated from artificial habitats (Table 11.4.10). The greatest abundance was estimated from the North Florida Shallow stratum (n = 287,951 ± 104,962) while Mid-Florida Deep stratum was unable to generate an estimate due to no samples collected. There were 2,171,199 (± 1,427,723) *Seriola dumerili* estimated from natural habitats and 2,745 (± 1,475) from artificial habitats in North/South Carolina region. The stratum with greatest abundance was North Carolina Shallow with 1,201,665 (± 845,729) (Table 11.4.11). For the South Atlantic, there were 3,037,903 ± 1,076,814 Greater Amberjack estimated from both habitat strata.

Partitioning Greater Amberjack abundance among size/age classes

The SERFS survey does not use stereo-lens or measurement lasers as part of sample collection, so measurement data was not available for any SERFS sites. The S-BRUV survey uses stereo-lens video therefore measurements are available when Greater Amberjack are observed in paired lenses, have aspect angles to the camera not too high, and turbidity/visibility levels are not elevated. During the S-BRUV sampling, 117 Greater Amberjack were measured (Figure 11.4.24) with 69 from natural hardbottom (47.9% of observed) and 48 from artificial habitats (16.5% of observed). Our ability to partition eventual estimates of Greater Amberjack abundance by size/age classes will likely be limited by such a small sample size.

Results: Estimation of absolute abundance of Greater Amberjack – C-BASS estimate

There was a single occurrence of Greater Amberjack observed by all C-BASS video surveys; which occurred in the South Carolina - shallow stratum. Density estimate was determined by calculating average density over all surveyed portions within each state and depth strata and extrapolating to the total amount of estimated non-reef habitat. The final estimate was 41,732 ± 28,076.5 Greater Amberjack (all from the North Carolina/South Carolina shallow stratum).

Results: Estimation of absolute abundance of Greater Amberjack – acoustics-based estimate

Acoustic-based estimates of potential Greater Amberjack abundance

Acoustic data were collected concurrently with many S-BRUV deployments in the South Atlantic region (methods described in Chapter 9). Processing of acoustic data resulted in a count of potential Seriola at each S-BRUV site. Similar to the eastern Gulf, agreement between paired S-BRUV and acoustic surveys, in terms of presence/absence of Greater Amberjack, was generally poor (Table 11.4.12 and 11.4.13). For natural hardbottom surveys, potential Greater Amberjack were identified with acoustics on 16 sites from NC/SC and 28 sites GA/FL where no Greater Amberjack were observed on video, while Greater Amberjack were observed on video at 1 site NC/SC and 7 sites GA/FL where no potential Greater Amberjack were identified through acoustics analyses. For artificial hardbottom surveys, potential Greater Amberjack were identified with acoustics on 8 sites NC/SC and 23 sites GA/FL where no Greater Amberjack were observed on video, while Greater Amberjack were observed on video at 3 sites NC/SC and 4 sites GA/FL where no potential Greater Amberjack were identified through acoustics analyses. Agreement sites where Greater Amberjack were observed and positive for acoustics was greater in the South Atlantic regions with 36 total sites across habitats and regions.

Estimates of Greater Amberjack for the South Atlantic region, as determined through analyses of acoustics data, are provided in Table 11.4.14 and 11.4.15. Acoustic estimates were only available for S-BRUV associated deployments as the SERFS survey did not conduct concurrent acoustic surveys. There were spatial strata within the South Atlantic region which were not sampled with S-BRUVs due to limited funds and acknowledgement of sufficient coverage from the SERFS survey. Additionally, many of the deep depth strata were only able to be sampled with acoustics as the current was too strong for S-BRUV deployment. This precluded the application of camera-based ratios of Greater Amberjack to Large bodied-fishes for acoustic data; resulting in no ability to generate abundance values from these strata. For the Georgia/Florida region, estimates from artificial reef habitats were low with 745 Greater Amberjack estimated from North Florida shallow stratum. There were seven spatial strata with estimates of zero Greater Amberjack and five strata unable to generate estimates due to no acoustic data available. Estimates from natural habitats were higher, however only three spatial strata contained positive values. North Florida shallow (N=827,974), North Florida deep (N=100,165) and Mid-Florida shallow (N=51,410) contained all the estimated Greater Amberjack. Other spatial strata contained estimates of zero fish or no estimate could be generated. Sites from North and South Carolina were characterized by larger estimates of abundance with 15,043 Greater Amberjack from artificial habitats and 2,051,885 from natural habitats. The greatest spatial strata was North Carolina natural with 1,979,472 estimated Greater Amberjack.

Discussion

Camera-based Greater Amberjack abundances were generated for both regions following modified design-based estimation procedures established for the West Florida region (Chapter 12.3) with the notable exception that partitioning of Greater Amberjack abundances into size / age classes will not be possible due to the extremely sparse availability of size composition data. Camera-based abundance estimation in the South Atlantic was

facilitated greatly by the availability of data from the Southeast Reef Fish Survey (SERFS), an ongoing broad-scale, fishery-independent survey focused on natural hardbottom habitats from North Carolina south to Fort Pierce, FL. Data from SERFS are collected following a statistically-rigorous simple random survey design, and represent abundance data using underwater video methodologies that typically exhibit less species and size selectivity than traditional capture gears. In the South Atlantic, this survey was complemented by an ancillary survey utilizing stereo-baited remote underwater video (S-BRUV) arrays. The S-BRUV survey, which overlapped with the SERFS survey to a large extent, was designed to (1) provide data on artificial hardbottom habitats that are not currently sampled by SERVS, (2) provide stereo-video length data (SERFS does not broadly utilize stereo cameras and Greater Amberjack are very rarely captured within traps), and (3) provide data beyond the southern terminus of the SERFS survey. The SERFS and S-BRUV surveys of hardbottom habitats were further complemented by a towed camera survey designed to provide broad coverage on areas of presumed unconsolidated sediment, or non-reef habitat. Combined, data from these surveys should be capable of providing Greater Amberjack abundance estimates on natural hardbottom habitats, artificial hardbottom habitats, and non-reef habitats, respectively.

Our proposed approach (Figure 12.3.13) to camera-based abundance estimation relies on some key assumptions, several of which were based largely on experience and intuition along with some very limited empirical data. As was the case for West Florida, we acknowledge that there is some uncertainty in the ability of video readers to accurately identify Seriola spp. on video, especially at small sizes. To compensate for this, video reading protocols require that video readers only identify individuals to species if they are certain of their identification, and all individuals that were identified to genus only (*Seriola* spp.) were excluded from analyses. Second, the choice of our abundance metric, MaxN, is in itself conservative. Some evidence suggests that MaxN may underestimate true abundance, especially among schooling species. Nevertheless, the use of MaxN where all individuals were observed on a single screen shot ensures that double counting is not possible. Third, there is some uncertainty in our conversion of single-camera counts from the SERFS survey to comparable spherical-camera counts produced by the S-BRUV survey. For the West Florida region, we were able to utilize paired single-camera to spherical-camera comparisons from the same S-BRUV array to inform these calibrations. As the horizontal and vertical field of view of the SERFS and S-BRUV single cameras are not identical, we could not directly utilize the West Florida calibration protocols. Similarly, although both gears were included in the calibration studies conducted off Alabama, results from these studies were limited and likely not representative of the types of habitats in the South Atlantic. Ideally, we would have been able to conduct a paired calibration study in the South Atlantic where both systems were deployed either concurrently or simultaneously, however such efforts were not possible given the limited resources available for this project. Instead, we relied on comparisons of the distribution of count data from SERFS and S-BRUVS in an overlapping area to develop a quantile-based abundance conversion. While we feel this approach is reasonable, the low sample sizes, especially from the S-BRUV survey, add a great deal of uncertainty. Fourth, instead of using the area viewed by the camera in our density estimates (~ 10 m radius under ideal conditions), we estimated a conservative and much greater effective area of sampling (~ 108 m radius). Greater Amberjack are known to be reactive to sampling vehicles moving through the water, and results from the Gulf indicated at least a modest influence of bait attraction, especially on natural hardbottom sites. Therefore, our estimates represent the likely furthest distance we would expect Greater Amberjack to be drawn from. And fifth, unlike the West Florida region, existing estimates of habitat composition and coverage in the South Atlantic are not based on randomized mapping efforts, but modeling results with a great deal of uncertainty. We attempted to compensate for this by conducting a limited number of habitat mapping surveys to estimate true habitat proportions within various predictive classes; however, significantly more mapping effort would be required to improve the accuracy of

these predictions. Admittedly these assumptions come with a certain degree of uncertainty that cannot be directly quantified by the current study, and the highly variable nature of Greater Amberjack abundance and site-specific occupancy certainly complicates these efforts. However, we feel the abundance estimation framework we developed is reasonable, and with a series of focused studies designed to test or quantify these assumptions, could have broad applicability for other managed reef fishes in the region.

Generating acoustics-based estimates of Greater Amberjack in the South Atlantic region was challenging for two primary reasons. First, acoustic data were unavailable for SERFS sampling sites; therefore, our attempts to produce Greater Amberjack abundance estimates using acoustic data had significant spatial gaps since acoustics were not used in the mid- depth stratum or conducted over unconsolidated sediment in conjunction with the C-BASS survey. And second, many strata sampled with acoustics did not have sufficient S-BRUV data (e.g., deep strata where high current caused BRUVs to move across seafloor) from which ratios of Greater Amberjack to all large-bodied fishes could not be generated to correct acoustics estimates. Because of these limitations, both of which could be addressed in the future with more robust sampling intensity, there were significant disagreements between camera-based estimates and acoustics-based estimates for several strata (for example, in natural reef habitats in Georgia where zero Greater Amberjack were estimated using acoustic-based data while over 300,000 were estimated from the camera-based approach). Despite these challenges, overall estimates of Greater Amberjack abundance were generally in the same order of magnitude for both camera-based and acoustics-based approaches. Improvement can be made toward generation of absolute abundance data with improved habitat mapping data for hard-bottom associated species, understanding of effective area sampled with optical based gears and improved ability to assign species-specific targets to hydroacoustic data.

References

Bacheler, N.M., Fuehring, A.D., Gillum, Z.D., Gregalis, K.C., Pickett E.P., Schobernd, C.M., Schobernd, Z.H. and Teer, B.Z. 2025. Estimating detection probabilities for reef fishes on baited underwater video. Marine Ecology Progress Series. 761:163-177. https://doi.org/10.3354/meps14872

Bacheler, N.M., K.W. Shertzer, J.A. Buckel, P.J. Ruderhausen, and B.J. Runde. 2018. Behavior of gray triggerfish *Balistes capriscus* around baited fish traps determined from fine-scale acoustic tracking. Marine Ecology Progress Series. 606:133-150.

Conley, M.F., M.G. Anderson, N. Steinberg, and A. Barnett, eds. 2017. The South Atlantic Bight Marine Assessment: Species, Habitats and Ecosystems. The Nature Conservancy, Eastern Conservation Science. [https://easterndivision.s3.amazonaws.com/Marine/SABMA/SABMA_Report_11April2018.pdf].

Grasty, S. E. 2014. Use of a towed camera system for estimating reef fish population densities on the West Florida shelf. M.S. thesis, x p. University of South Florida, St. Petersburg, Florida.

Grasty, S. F., Wall, C. C., Gray, J. W., Brizzolara, J., & Murawski, S., 2019. Temporal persistence of Red Grouper holes and analysis of associated fish assemblages from towed camera data in the steamboat lumps marine protected area. Transactions of the American Fisheries Society, 148(3), 652–660. https://doi.org/10.1002/tafs.10154

Lembke, C., S. Grasty, et al., 2017. The Camera-Based Assessment Survey System (C-BASS): A towed camera platform for reef fish abundance surveys and benthic habitat characterization in the Gulf of Mexico. Continental Shelf Research 151: 62-71.

Paxton, A.B and 22 co-authors. 2024. Artificial reef footprint in the United States ocean. Nature Sustainability. Vol. 7, 140 – 147. https://doi.org/10.1038/s41893-023-01258-7

Steward, D.N., Paxton A.B, Bacheler, N.M, Schobernd, C.M, Mille, K., Renchen, J., Harrison, Z., Byrum, J., Martore, R., Brinton, C., Riley, K.L., Taylor, J.C. and Kellison, G.T., 2022. Quantifying spatial extents of artificial versus natural reefs in the seascape. Front. Mar. Sci. 9:980384. doi: 10.3389/fmars.2022.980384

White, A.L., Patterson III, W.F., and Boswell, K.M. 2022. Distribution of acoustic fish backscatter associated with natural and artificial reefs in the Northeastern Gulf of Mexico. Fisheries Research. 248 106199. https://doi.org/10.1016/j.fishres.2021.106199

Table 11.4.1. Summary of all selectable sampling points, randomly selected sites, and randomly selected sites that were sampled in 2023 by the SERFS CVT survey of natural hardbottom sites in the South Atlantic. Although the SERFS survey utilized a simple random site selection protocol (with a global site selection probability), effort summaries were produced following the spatial stratification scheme developed for supplementary sampling efforts based on a combination of region (SERFS sampling does not extend into the South Florida region) and depth. Depth strata were defined as: Inshore (10 – 20m), Shallow (20.1 – 40m), Mid (40.1 – 70m), Deep (70.1 – 150m). Selected and Sampled numbers are from "Random" sites only.

Region	Depth Strata	Natural Points	Natural Selected	Selection Probability	Sites Sampled
North Carolina	*Inshore*	26	9	0.343407	4
	Shallow	664	299	0.343407	162
	Mid	228	106	0.343407	61
	Deep	156	79	0.343407	49
	TOTAL	**1,074**			
South Carolina	*Inshore*	143	23	0.343407	14
	Shallow	955	210	0.343407	104
	Mid	686	175	0.343407	93
	Deep	83	32	0.343407	18
	TOTAL	**1,869**			
Georgia	*Inshore*	64	22	0.343407	6
	Shallow	219	59	0.343407	42
	Mid	143	61	0.343407	53
	Deep	18	7	0.343407	6
	TOTAL	**444**			
North Florida	*Inshore*	10	6	0.343407	5
	Shallow	196	93	0.343407	80
	Mid	247	96	0.343407	90
	Deep	33	17	0.343407	13
	TOTAL	**486**			
Mid Florida	*Inshore*	22	14	0.343407	13
	Shallow	313	127	0.343407	114
	Mid	162	65	0.343407	52

	Deep	-	-	-	-
	TOTAL	**497**			
South Florida* *No sampling*	*Inshore*	-	-	-	-
	Shallow	-	-	-	-
	Mid	-	-	-	-
	Deep	-	-	-	-

Table 11.4.2. Summary of all selectable sampling points, randomly selected sites, and sites that were sampled in 2023 by supplemental S-BRUV sampling on natural hardbottom habitats in the South Atlantic. Summaries were produced following the spatial stratification scheme developed for supplementary sampling efforts based on a combination of region and depth. Depth strata were defined as: Inshore (10 – 20m), Shallow (20.1 – 40m), Mid (40.1 – 70m), Deep (70.1 – 150m). The number of available natural hardbottom sampling points was greater than SERFS universe as additional information was used to populate universe for S-BRUV survey. For the South Florida region, numbers represent microgrids containing natural hardbottom habitats. Numbers in parentheses under Sites Sampled represent the number of sites sampled with hydroacoustics; in some instances, only hydroacoustics were deployed due to high currents that prevented deployment of S-BRUV arrays.

Region	Depth Strata	Natural Points	Natural Selected	Selection Probability	Sites Sampled
North Carolina	*Inshore*	26	3	0.115385	2 (2)
	Shallow	664	13	0.019578	10 (10)
	Mid	228	-		-
	Deep	156	8	0.051282	6 (6)
	TOTAL	**1,074**	**24**		**18 (18)**
South Carolina	*Inshore*	143	3	0.020979	3 (3)
	Shallow	955	18	0.018848	17 (17)
	Mid	688	-		-
	Deep	83	5	0.060241	0 (3)
	TOTAL	**1,869**	**26**		**20 (23)**
Georgia	*Inshore*	64	3	0.046875	3 (3)
	Shallow	219	5	0.022831	3 (3)
	Mid	148	-		-
	Deep	28	4	0.142857	0 (2)
	TOTAL	**459**	**12**		**6 (8)**
North Florida	*Inshore*	22	3	0.136364	2 (2)
	Shallow	1,745	32	0.018338	25 (26)
	Mid	1,105	-		-
	Deep	283	14	0.04947	6 (6)
	TOTAL	**3,155**	**49**		**33 (34)**
Mid Florida	*Inshore*	150	3	0.02	2 (2)

	Shallow	2,555	48	0.018787	30 (34)
	Mid	1,164	-		-
	Deep	698	34	0.048711	0 (17)
	TOTAL	**4,567**	**85**		**32 (53)**
South Florida* *Number of microgrids*	*Inshore*	1,594	10	0.006273	7 (8)
	Shallow	3,795	19	0.005006	17 (16)
	Mid	968	7	0.007231	5 (8)
	Deep	14,078	2	0.000142	1 (1)
	TOTAL	**20,435**	**38**		**30 (33)**
	GRAND TOTAL		**234**		**139 (169)**

Table 11.4.3. Summary of all selectable sampling points, randomly selected sites, and sites that were sampled in 2023 by supplemental S-BRUV sampling on artificial hardbottom habitats in the South Atlantic. Summaries were produced following the spatial stratification scheme developed for supplementary sampling efforts based on a combination of region and depth. Depth strata were defined as: Inshore (10 – 20m), Shallow (20.1 – 40m), Mid (40.1 – 70m), Deep (70.1 – 150m). For South Florida region, numbers represent microgrids containing artificial hardbottom habitats. Numbers in parentheses under Sites Sampled represent the number of sites sampled with hydroacoustics; in some instances, only hydroacoustics were deployed due to high currents that prevented deployment of S-BRUV arrays.

Region	Depth Strata	Artificial Points	Artificial selected	Selection Probability	Sites Sampled
North Carolina	*Inshore*	294	12	0.040816	10 (10)
	Shallow	98	4	0.040816	3 (3)
	Mid	11	-		-
	Deep	11	-		-
	TOTAL	**414**	**16**	0.038647	**13 (13)**
South Carolina	*Inshore*	484	18	0.03719	12 (13)
	Shallow	117	5	0.042735	3 (3)
	Mid	6	-		-
	Deep	1	-		0 (0)
	TOTAL	**608**	**23**	0.037829	**15 (16)**
Georgia	*Inshore*	438	17	0.038813	7 (7)
	Shallow	19	3	0.157895	1 (2)
	Mid	9	-		-
	Deep	0	-		-
	TOTAL	**466**	**20**	0.042918	**8 (9)**
North Florida	*Inshore*	343	13	0.037901	9 (9)
	Shallow	444	18	0.040541	8 (9)
	Mid	14	-		-
	Deep	14	-		-
	TOTAL	**815**	**31**	0.038037	**17 (18)**
Mid Florida	*Inshore*	360	14	0.038889	6 (6)
	Shallow	688	25	0.036337	11 (11)

	Mid	139	3	0.021583	1 (1)
	Deep	13	-		-
	TOTAL	**1,200**	**42**	0.035	**18 (18)**
South Florida* *Number of microgrids*	*Inshore*	73	3	0.041096	2 (2)
	Shallow	136	6	0.044118	4 (4)
	Mid	69	3	0.043478	2 (2)
	Deep	84	4	0.047619	1 (1)
	TOTAL	362	**16**		**9 (9)**
	GRAND TOTAL	**3,865**	**148**		**80 (83)**

Table 12.4.4. Estimates of total area coverage within each subregion (NC: North Carolina; SC: South Carolina; GA: Georgia; NF: North Florida; MF: Mid Florida) and depth strata (Shallow: 20 – 40 m; Mid/Deep: 40 – 150 m) with the South Atlantic survey domain for surveys of unconsolidated habitats.

	Area (km^2)		
	NC/SC	**GA/FL**	
Shallow (20-40 m)	43,160.2	24,973.9	**68,134.1**
MidDeep (41-150 m)	18,797.9	14,243.9	**33,041.8**
TOTAL	**61,958.1**	**39,217.8**	**101,175.9**

Table 11.4.5. Planned (N = 60) and completed (N = 31) C-BASS transects within each subregion (NC: North Carolina; SC: South Carolina; GA: Georgia; NF: North Florida; MF: Mid Florida) and depth strata (Shallow: 20 – 40 m; Mid/Deep: 40 – 150 m) within the South Atlantic survey domain for surveys of unconsolidated habitats.

	NC/SC		**GA/FL**		**TOTAL**	
	Planned	***Actual***	*Planned*	***Actual***	*Planned*	***Actual***
Shallow Transects	26	**13**	15	**8**	42	***21***
MidDeep Transects	10	**5**	8	**5**	18	***10***
Total No. Transects	*37*	***18***	*23*	***13***	60	***31***
n:*N*	0.61	***0.58***	0.39	***0.42***		

Table 11.4.6. Summary of Predicted habitat confidence bins for Florida and Georgia by each pseudo-state strata. Values represent area of predicted habitat bin multiplied by the FWRI mapping weighted ratio to generate estimate of hard bottom (km^2) for each state.

Region	Total Area (km^2)	Medium probability hard bottom (km^2)	High probability hard bottom (km^2)	Highest probability hard bottom (km^2)	Confirmed natural hard bottom (km^2)	Confirmed natural hard bottom associated species (km^2)	Predicted Area of No hard bottom (km^2)	Artificial reefs (km^2)
	FWRI mapping weighted ratios	0.0376	0.0678	0.1305	0.6816	0.1426	0.0245	0.0343
Georgia - Inshore	4,799.426	26.754	7.999	1.333	0	0.255	96.432	0.751
Georgia - Shallow	7,055.134	59.620	180.684	234.010	0	0.273	24.707	0.034
Georgia - Mid	3,038.925	7.917	74.811	224.703	0	0.391	0.002	0.009
Georgia - Deep	534.941	0.788	10.312	47.140	0	0.070	0.004	0
North Florida – Inshore	3,188.480	34.016	10.866	1.462	0.754	0.103	51.531	0.247
North Florida – Shallow	9,129.849	106.439	245.494	265.940	0.902	0.139	15.462	0.238
North Florida – Mid	2,537.660	4.250	73.957	172.776	0	1.119	0.048	0.002
North Florida - Deep	683.138	0.653	11.998	63.279	0	0.513	0.008	0
Mid Florida – Inshore	3542.183	36.722	11.397	5.672	0.894	0.040	57.445	0.263

Mid Florida – Shallow	4,946.091	65.841	139.490	99.742	9.405	0.081	8.514	0.393
Mid Florida – Mid	2,838.704	4.974	84.109	186.746	19.218	0.576	0.001	0.089
Mid Florida - Deep	2,541.535	1.387	34.015	245.270	79.550	0.881	0.015	0
South Florida – Inshore	73.144	0.532	0.252	0.022	6.734	1.861	0.751	0.052
South Florida – Shallow	141.311	1.361	0.914	0.248	35.839	3.306	0.266	0.107
South Florida – Mid	220.815	2.377	3.482	7.451	8.418	0.913	0.654	0.127
South Florida - Deep	433.492	3.793	13.908	3.226	0.051	0.029	2.413	0.137
TOTAL	45,712.7	357.426	903.691	1,559.020	161.764	10.549	258.252	2.447

Table 11.4.7. Summary of estimated areas (km^2) of three habitat types within the GA/FL region.

Region	Total Area (km^2)	Estimated Natural Hardbottom	Estimate Artificial Reef	Estimated uncharacterized
Georgia	15,428.4	998.207	0.793	14,429.43
North Florida	15,539.1	1,061.707	0.487	14,476.93
Mid Florida	13,876.4	1,091.981	0.744	12,783.68
South Florida	868.8	98.803	0.422	769.54
TOTAL (km2)	45,712.7	3,250.70	2.447	42,459.57

Table 11.4.8. Summary of Predicted habitat confidence bins for North and South Carolina by each pseudo-state strata. Values represent area of predicted habitat bin multiplied by the FWRI mapping weighted ratio to generate estimate of hard bottom (km^2) for each state.

Region	Total Area	Medium probability hard bottom	High probability hard bottom	Highest probability hard bottom	Confirmed natural hard bottom	Confirmed natural hard bottom associated species	Artificial reefs	Predicted Area of No hard bottom
	FWRI mapping weighted ratios	0.0376	0.0678	0.1305	0.6816	0.1426	0.0343	0.0245
North Carolina - Inshore	7,442.476	72.899	28.188	8.656	0	0.080	122.675	0.476
North Carolina – Shallow	21,048.894	193.707	410.314	792.749	0	0.624	92.151	0.169
North Carolina – Mid	5,490.694	45.552	74.154	293.747	0	0.365	22.829	0.006
North Carolina - Deep	2,762.939	18.612	34.078	131.201	0	1.478	18.362	0.005
South Carolina – Inshore	8,244.101	72.374	64.297	33.006	0	0.305	125.035	0.425
South Carolina – Shallow	12,582.208	59.400	295.471	829.354	0	0.271	6.960	0.112
South Carolina – Mid	3,935.020	2.719	81.899	343.470	0	2.847	0.064	0.008
South Carolina - Deep	2,510.390	8.055	59.000	180.723	0	0.577	0.908	0
TOTAL	64,016.3	473.317	1,047.402	2,612.906	0	6.546	388.984	1.201

Table 11.4.9. Summary of estimated areas (km^2) of three habitat types within the NC/SC region.

Region	Total Area	Estimated Natural Hardbottom	Estimate Artificial Reef	Estimated uncharacterized
North Carolina	36,745.0	2,362.421	0.655	34,381.93
South Carolina	27,271.3	2,166.718	0.546	25,104.05
TOTAL	64,016.3	4,529.140	1.201	59,485.98

Table 11.4.10. Camera-based estimates (S-BRUV and SERFS trap-video combined) of abundance for *Seriola dumerili* for natural and artificial hardbottom sites in the Georgia/Florida region. Estimates were generated using an estimated effective area sampled of 36,644 m^2 (108 m buffer). Estimates were presented for each pseudo-state and depth stratum. Dots represent no data available for stratum.

	Natural		Artificial	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
Georgia - Inshore	54,242	79,555	3,428	2,530
Georgia - Shallow	72,715	67,985	0	0
Georgia - Mid	50,679	17,993	.	.
Georgia - Deep	156,451	156,451	.	.
North Florida – Inshore	0	0	2,675	2,116
North Florida – Shallow	287,951	104,962	1,181	865
North Florida – Mid	68,187	14,481	.	.
North Florida - Deep	5,511	3,167	.	.
Mid Florida – Inshore	0	0	76	75
Mid Florida – Shallow	51,901	27,789	1,304	830
Mid Florida – Mid	107,631	52,467	0	0
Mid Florida - Deep	.	.	.	.
South Florida – Inshore	0	0	28	29
South Florida – Shallow	0	0	0	0
South Florida – Mid	0	0	0	0
South Florida - Deep	0	0	0	0
Total	855,267	224,751.3	8,692	3,510.2

Table 11.4.11. Camera-based estimates (S-BRUV and SERFS trap-video combined) of abundance for *Seriola dumerili* for natural and artificial hardbottom sites in the North Carolina/South Carolina region. Estimates were generated using an estimated effective area sampled of 36,644 m^2 (108 m buffer). Estimates were presented for each pseudo-state and depth stratum. Dots represent no data available for stratum.

Spatial Strata	Natural		Artificial	
	Estimate	Standard Error	Estimate	Standard Error
North Carolina - Inshore	37,902	41,290	38	38
North Carolina – Shallow	1,201,665	845,729	996	602
North Carolina – Mid	35,967	19,926	.	.
North Carolina - Deep	17,295	17,166	.	.
South Carolina – Inshore	0	0	1,131	750
South Carolina – Shallow	715,626	624,407	580	446
South Carolina – Mid	149,154	37,782	.	.
South Carolina - Deep	13,591	6,129	.	.
TOTAL	2,171,199	1,053,092.0	2,745	1,060.4

Table 11.4.12. Confusion matrix for North Carolina/South Carolina region comparing agreement between S-BRUV and acoustic surveys in relation to whether Greater Amberjack were observed on video and potential Greater Amberjack were detected via analysis of paired acoustics data for natural and artificial hardbottom surveys.

Natural Hardbottom Surveys

		Acoustic Survey	
		Absent	Present
S-BRUV Survey	Absent	13	16
	Present	1	8

Artificial Hardbottom Surveys

		Acoustic Survey	
		Absent	Present
S-BRUV Survey	Absent	12	8
	Present	3	5

Table 11.4.13. Confusion matrix for Georgia/Florida region comparing agreement between S-BRUV and acoustic surveys in relation to whether Greater Amberjack were observed on video and potential Greater Amberjack were detected via analysis of paired acoustics data for natural and artificial hardbottom surveys.

Natural Hardbottom Surveys

		Acoustic Survey	
		Absent	Present
S-BRUV Survey	Absent	50	28
	Present	7	13

Artificial Hardbottom Surveys

		Acoustic Survey	
		Absent	Present
S-BRUV Survey	Absent	14	23
	Present	4	10

Table 11.4.14. Acoustic-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites in the Georgia/Florida region. Estimates were generated using an estimated effective area sampled of 36,644 m^2 (108 m buffer). Estimates were presented for each pseudo-state and depth stratum.

	Artificial		Natural	
Spatial Strata	Estimate	Standard Error	Estimate	Standard Error
Georgia - Inshore	113	106.6	0	0
Georgia - Shallow	0	0	0	0
Georgia - Mid	.	.	.	.
Georgia - Deep	.	.	.	.
North Florida – Inshore	158	157.6	0	0
North Florida – Shallow	745	801.7	827,974	342,473.5
North Florida – Mid	.	.	.	.
North Florida - Deep	.	.	100,165	97,268.4
Mid Florida – Inshore	0	0	0	0
Mid Florida – Shallow	233	129.6	51,410	31,835.7
Mid Florida – Mid	0	0	.	.
Mid Florida - Deep	.	.	.	.
South Florida – Inshore	0	0	0	0
South Florida – Shallow	0	0	0	0
South Florida – Mid	0	0	0	0
South Florida - Deep	0	0	0	0
Total	1,249	834.1	979,549	357,439.2

Table 11.4.15. Acoustic-based estimates of abundance for *Seriola dumerili* for natural and artificial hardbottom sites in the North Carolina/South Carolina region. Estimates were generated using an estimated effective area sampled of 36,644 m^2 (108 m buffer). Estimates were presented for each pseudo-state and depth stratum.

Spatial Strata	Artificial		Natural	
	Estimate	Standard Error	Estimate	Standard Error
North Carolina - Inshore	755	557.7	0	0
North Carolina – Shallow	10,182	2,591.8	1,979,472	1,229,534.0
North Carolina – Mid	.	.	.	.
North Carolina - Deep	.	.	49,612	26,914.6
South Carolina – Inshore	0	0	0	0
South Carolina – Shallow	4,106	776.2	22,801	9,158.1
South Carolina – Mid	.	.	.	.
South Carolina - Deep	.	.	.	.
TOTAL	15,043	2,762.4	2,051,885	1,229,862.6

Figure 11.4.1. The South Atlantic sampling frame. The South Atlantic was subdivided into six regions (North Carolina, South Carolina, Georgia, North Florida, Mid Florida, South Florida) and four depth strata (Inshore: 10 – 20m; Shallow: 20.1 – 40m; Mid: 40.1 – 70m; Deep: 70.1 – 150m) for supplemental sampling effort allocation.

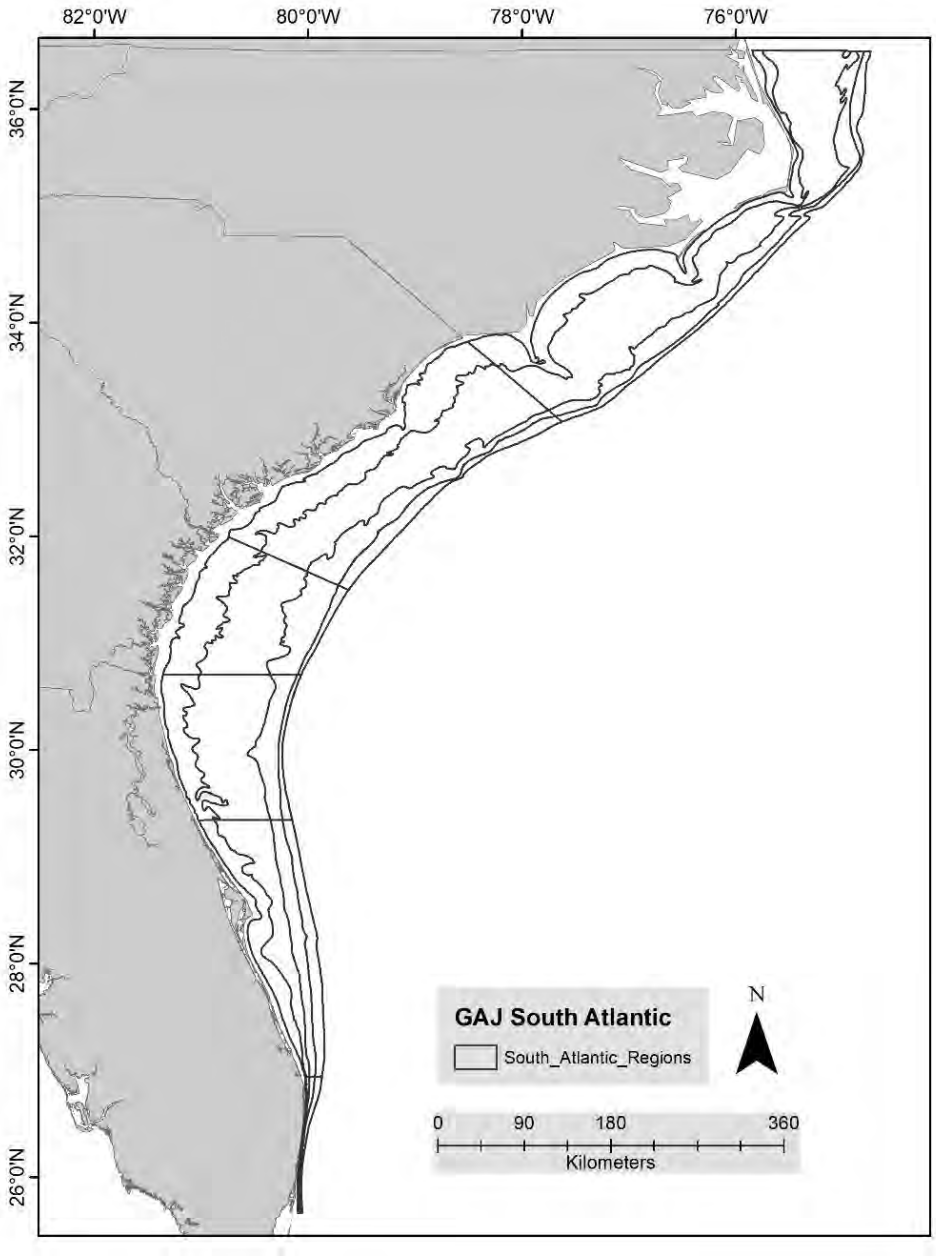


Figure 11.4.2. Map highlighting all available sites in the SERFS sampling frame (n=4,370; left) and all randomly selected sites for sampling in 2023 (right). For the right panel, green circles identify sites that were sampled, and black circles identify sites that were unsampled.

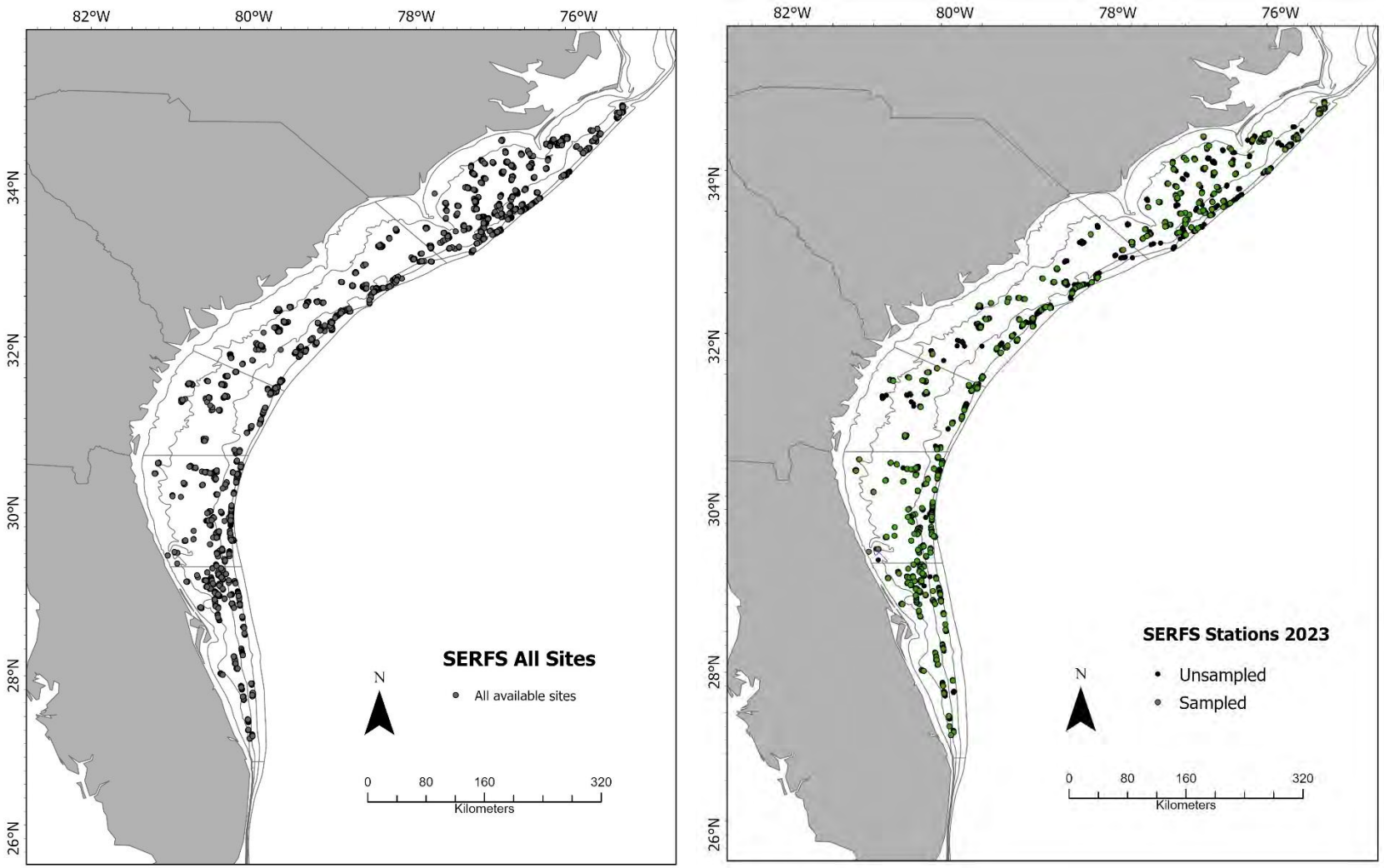


Figure 11.4.3. South Florida sampling region with habitat strata from Florida Unified Reef map.

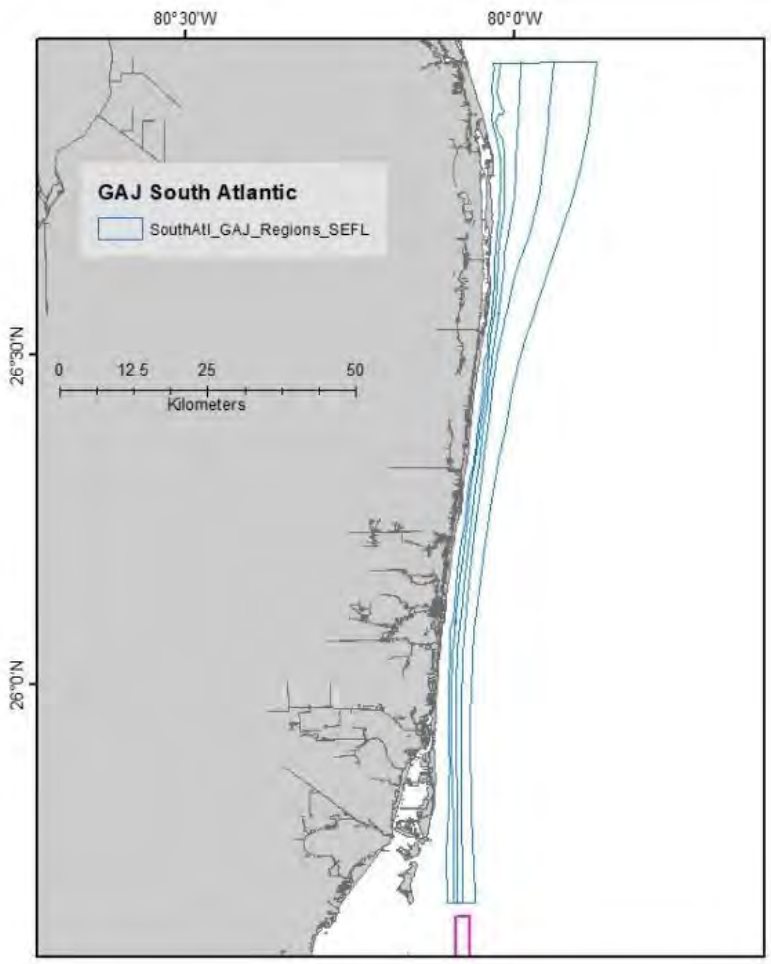


Figure 11.4.4. Northern portion of the South Florida region which was gridded with 0.1 x 0.1nm cells. Each cell was intersected with hard bottom habitat data from the Florida Unified Reef map.



Figure 11.4.5. Map highlighting all available natural reef sites in the S-BRUV sampling frame (n=11,124; left) and all randomly selected sites for sampling in 2023 (right). For the right panel, blue circles identify sites that were sampled, and red circles identify sites that were unsampled.

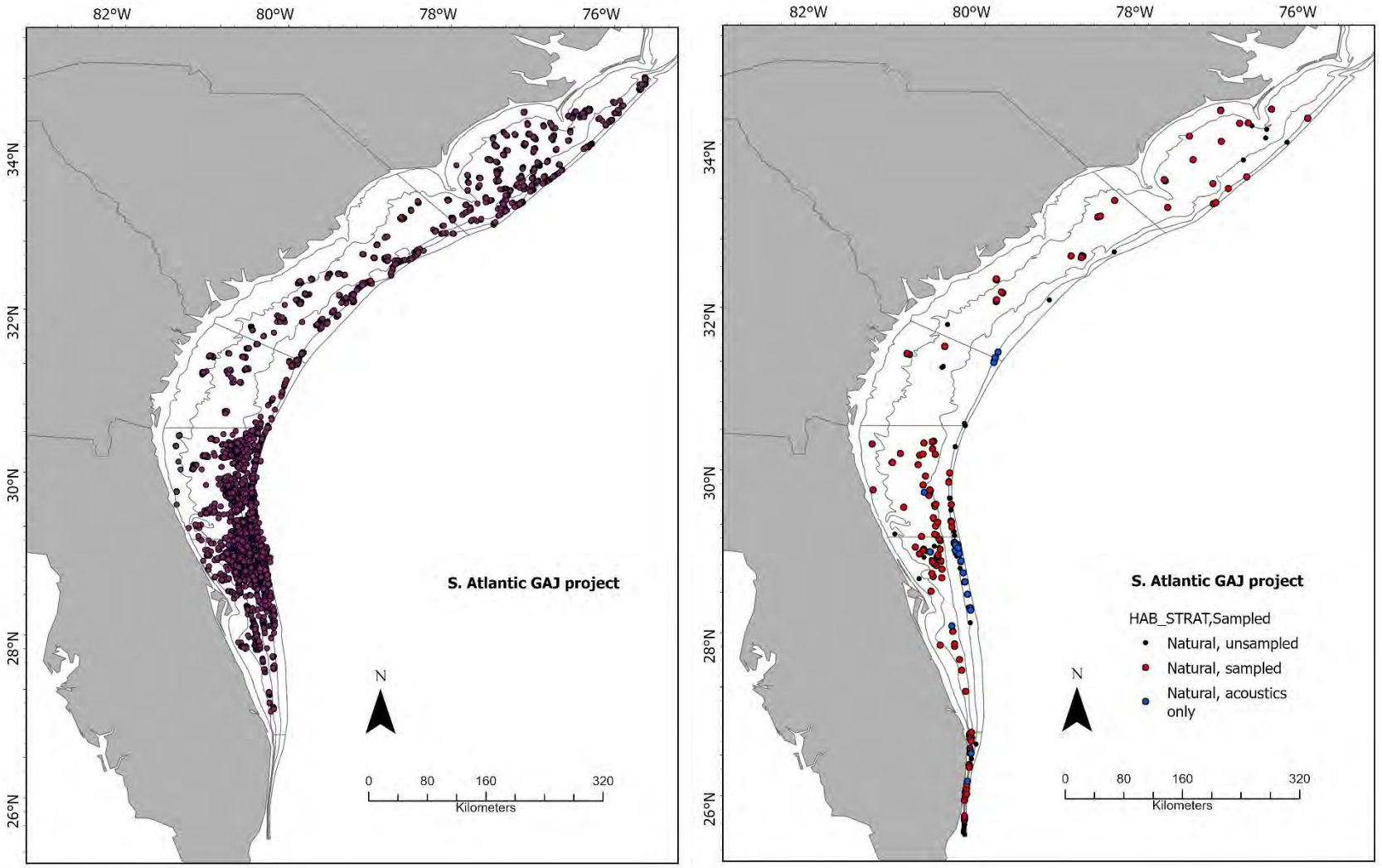


Figure 11.4.6. Map highlighting all available artificial reef sites in the S-BRUV sampling frame (n=3,865; left) and all randomly selected sites for sampling in 2023 (right). For the right panel, blue circles identify sites that were sampled, and red circles identify sites that were unsampled.

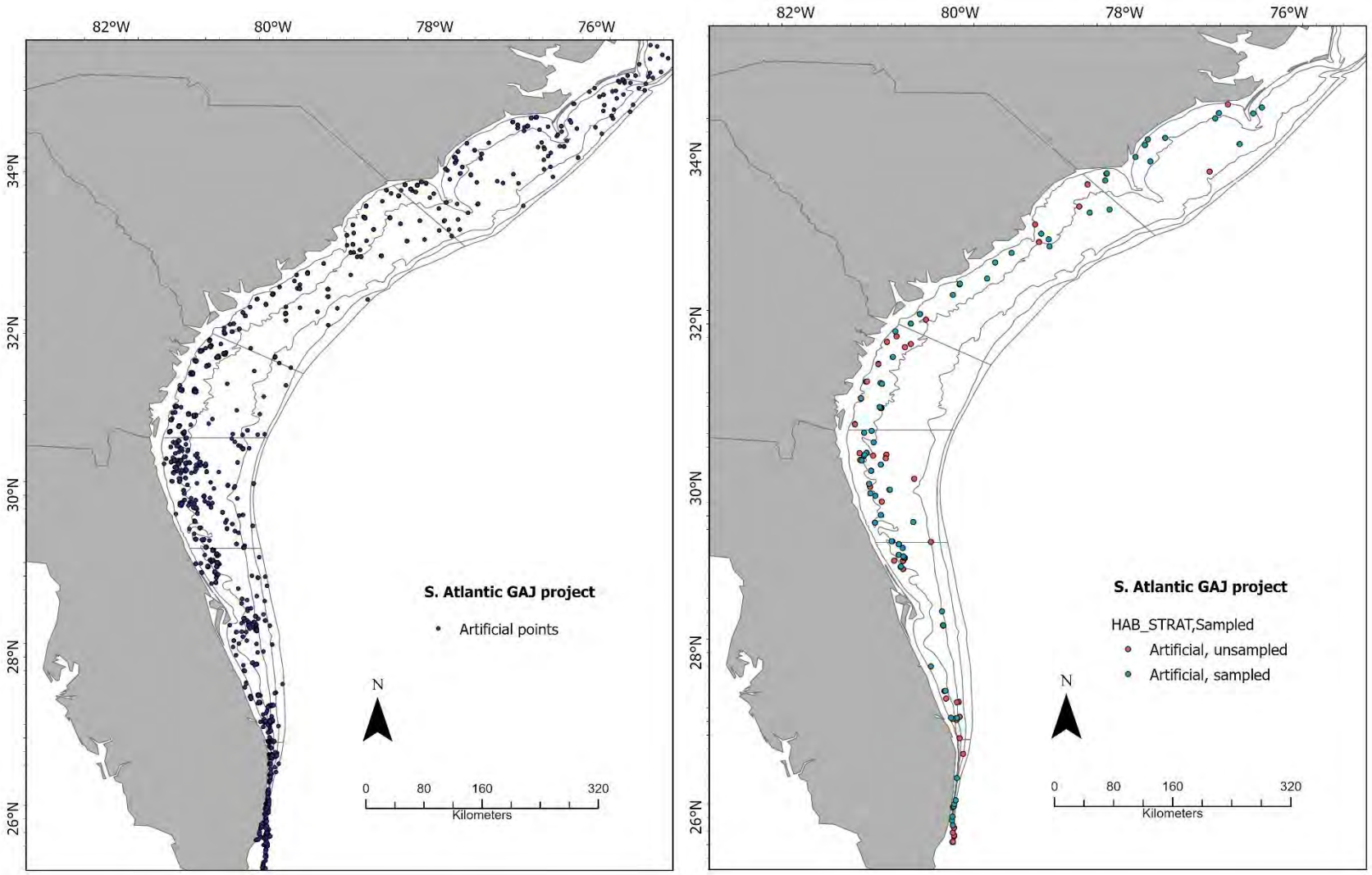


Figure 11.4.7. Images of the three sampling gears deployed in the South Atlantic: a) S-BRUV; b) chevron-video trap used by the SERFS survey; c) C-BASS

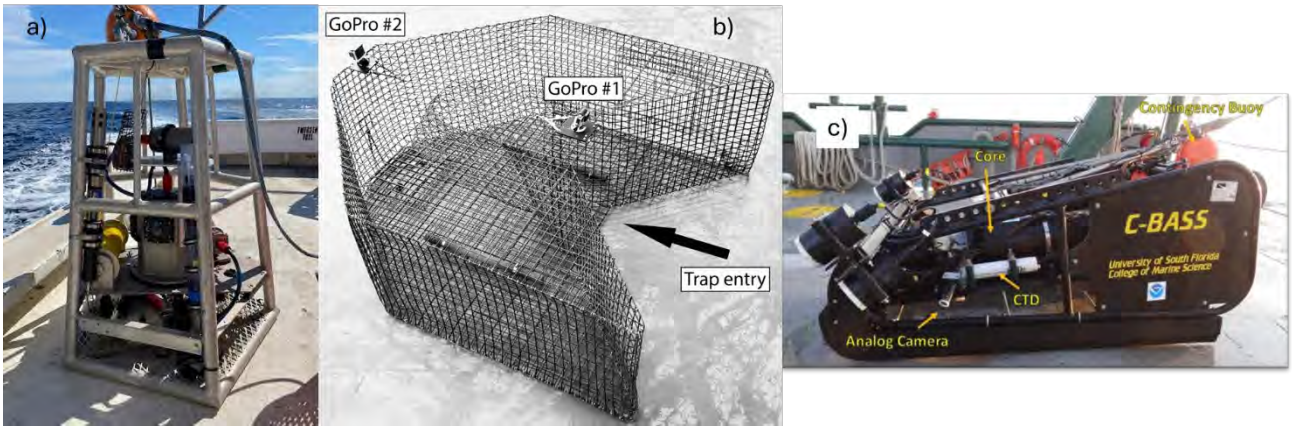


Figure 11.4.8. Example from Southeast Conservation Adaptation Strategy Atlas; Atlantic Coral and Hardbottom predictive model from Mid Florida to Georgia coastal waters. Colors show varying confidence intervals for hard bottom presence. The model was bounded by the GAJ project footprint for the South Atlantic regions.

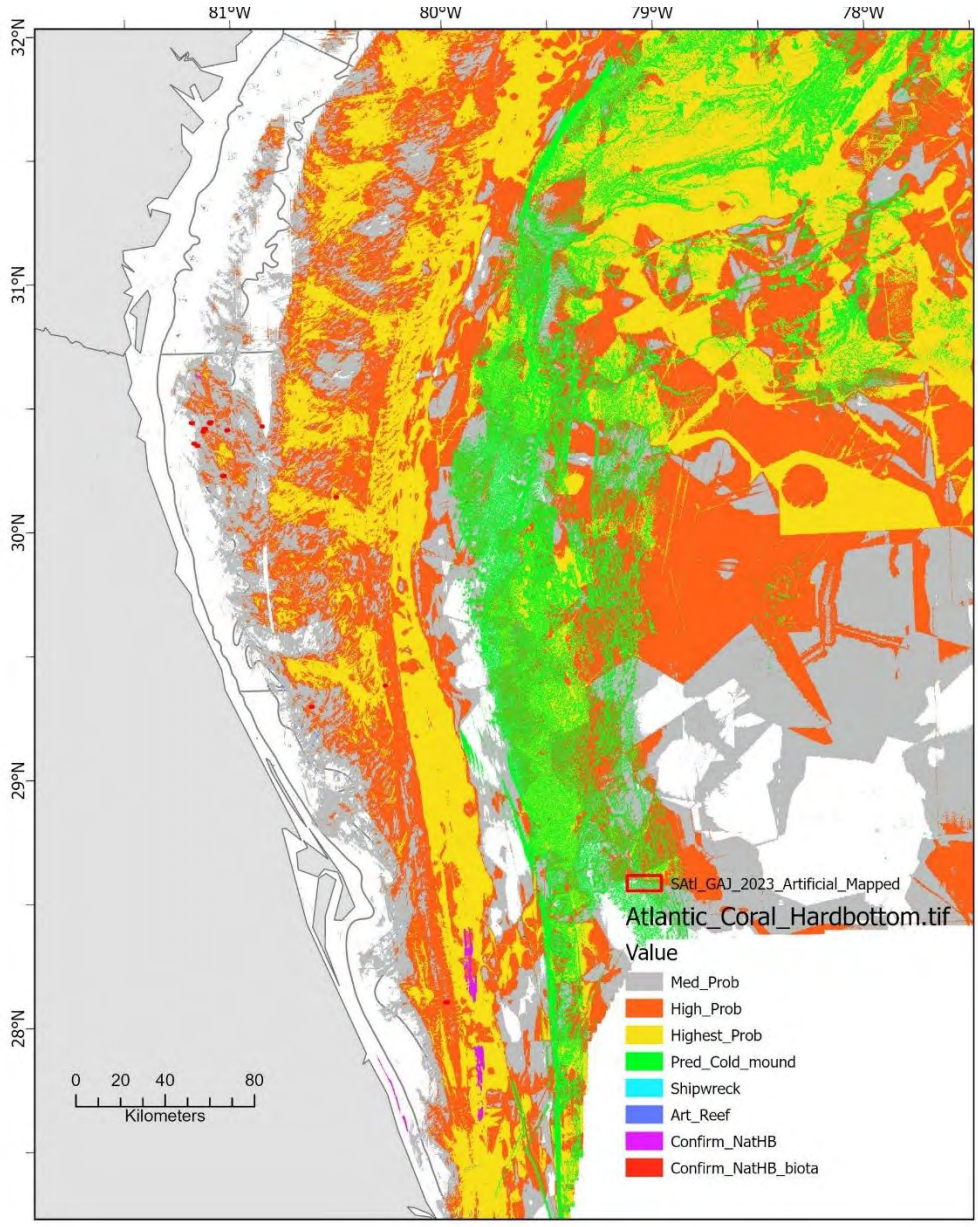


Figure 11.4.9. Example of Predicted habitat map with two side scan surveys shown. In left figure, side scan imagery was digitized into habitat classes and shown as polygons within the survey. The right figure shows raster data on predicted habitat data.

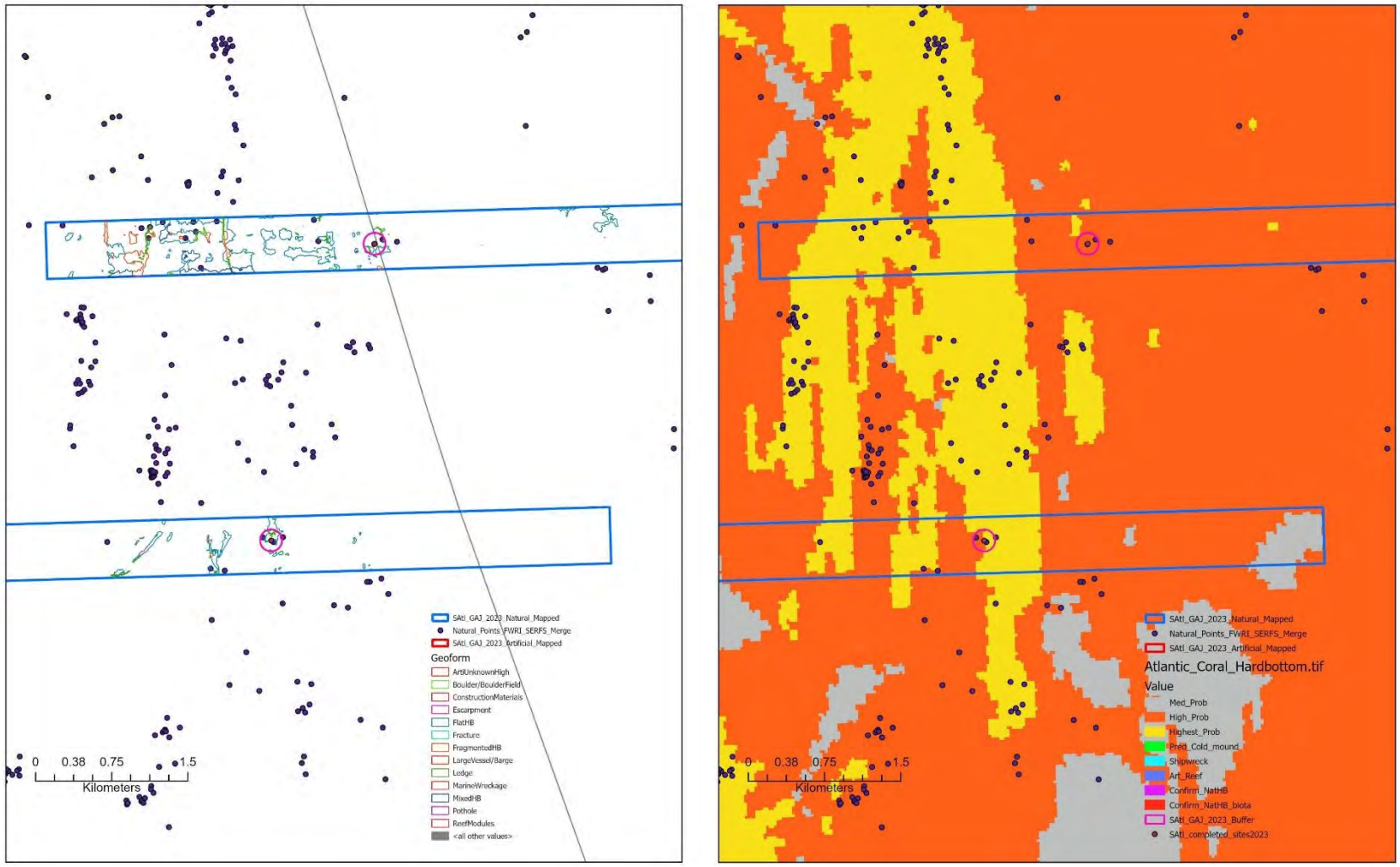


Figure 11.4.10. Comparative cumulative frequency plot of single-camera MaxN counts from the SERFS survey and spherical-camera MaxN counts from the S-BRUV survey on natural hardbottom habitats from 10 – 40 m. Conversion of single-camera to spherical-camera abundance currency was conducted by converting single-camera abundances to the corresponding spherical-camera quantile.

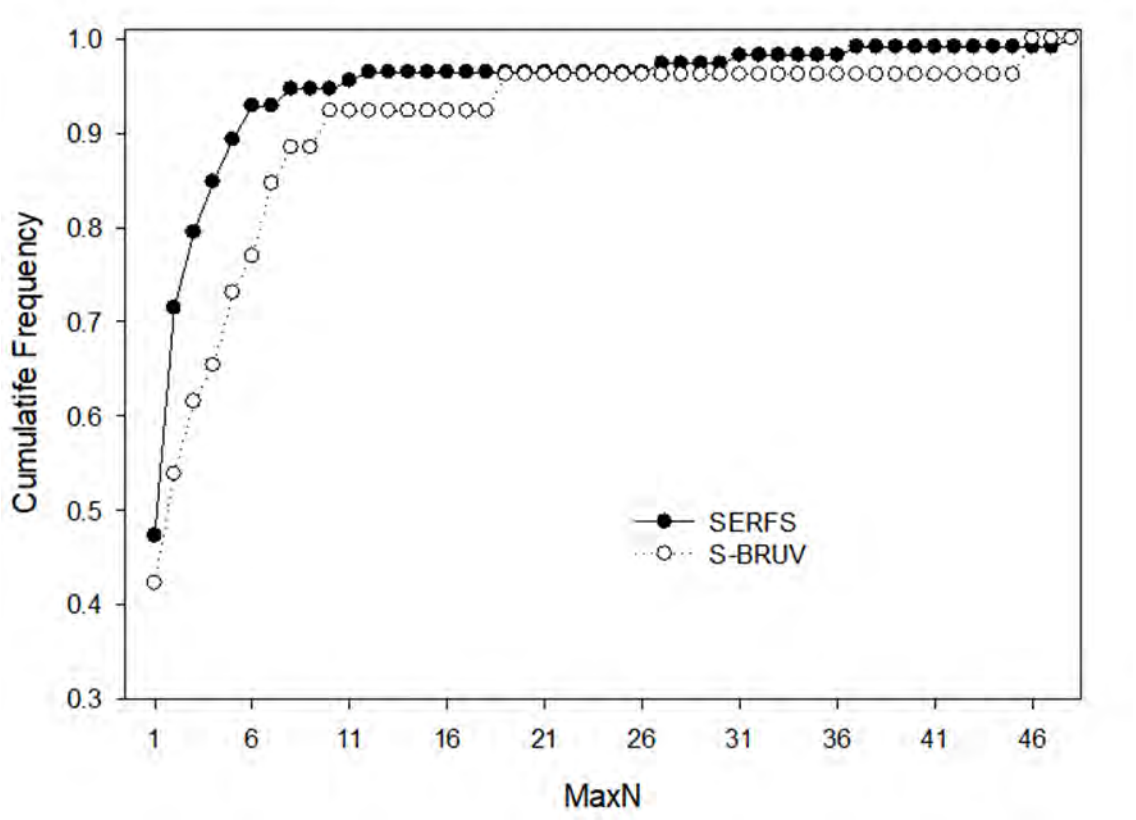


Figure 11.4.11. Regional frequency distribution of MaxN values from the SERFS video survey of natural hardbottom habitats in the South Atlantic. Values are reported as converted spherical camera currency.

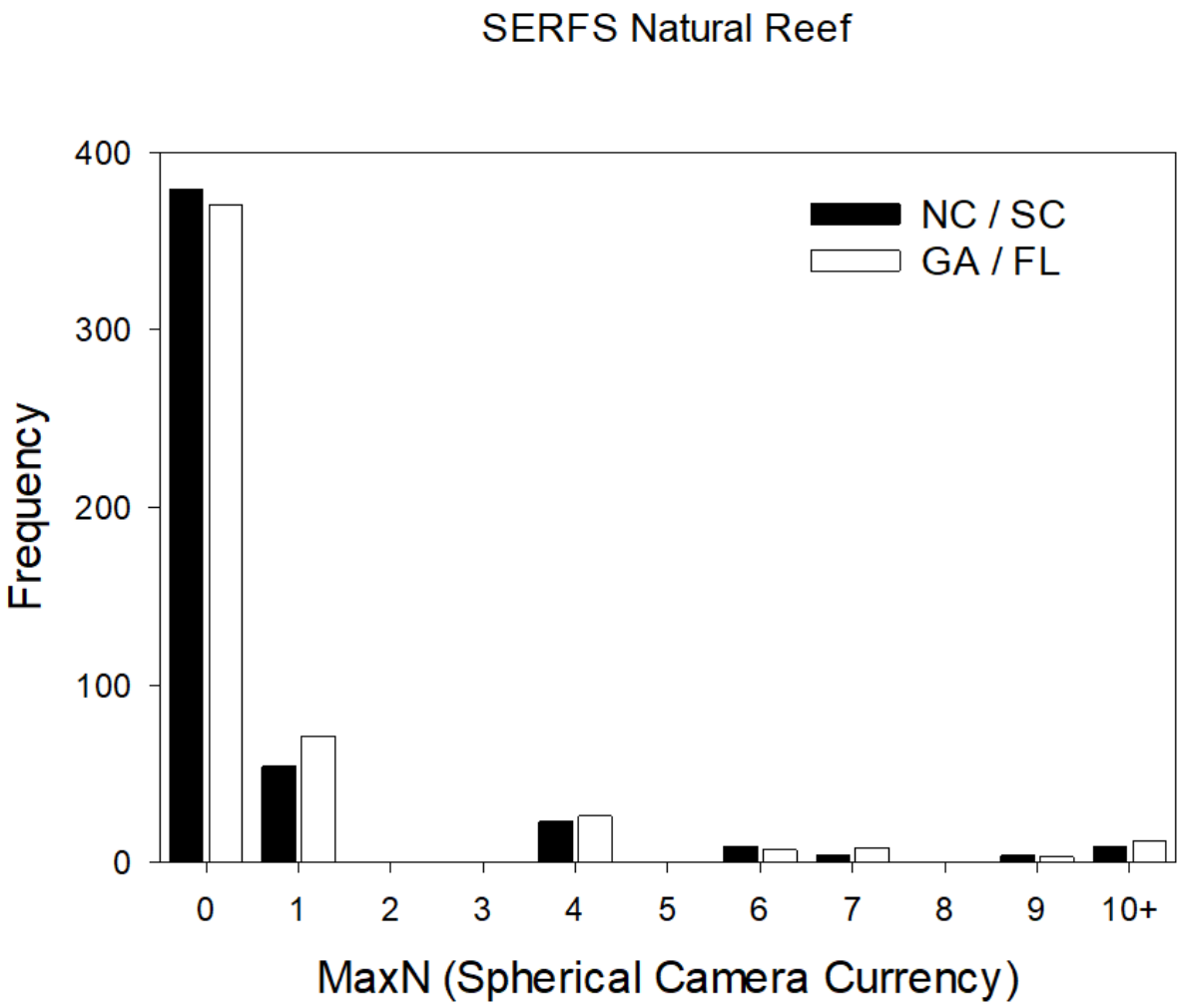


Figure 11.4.12. Regional frequency distribution of the effective reef area sampled during the SERFS video survey of natural hardbottom habitats in the South Atlantic.

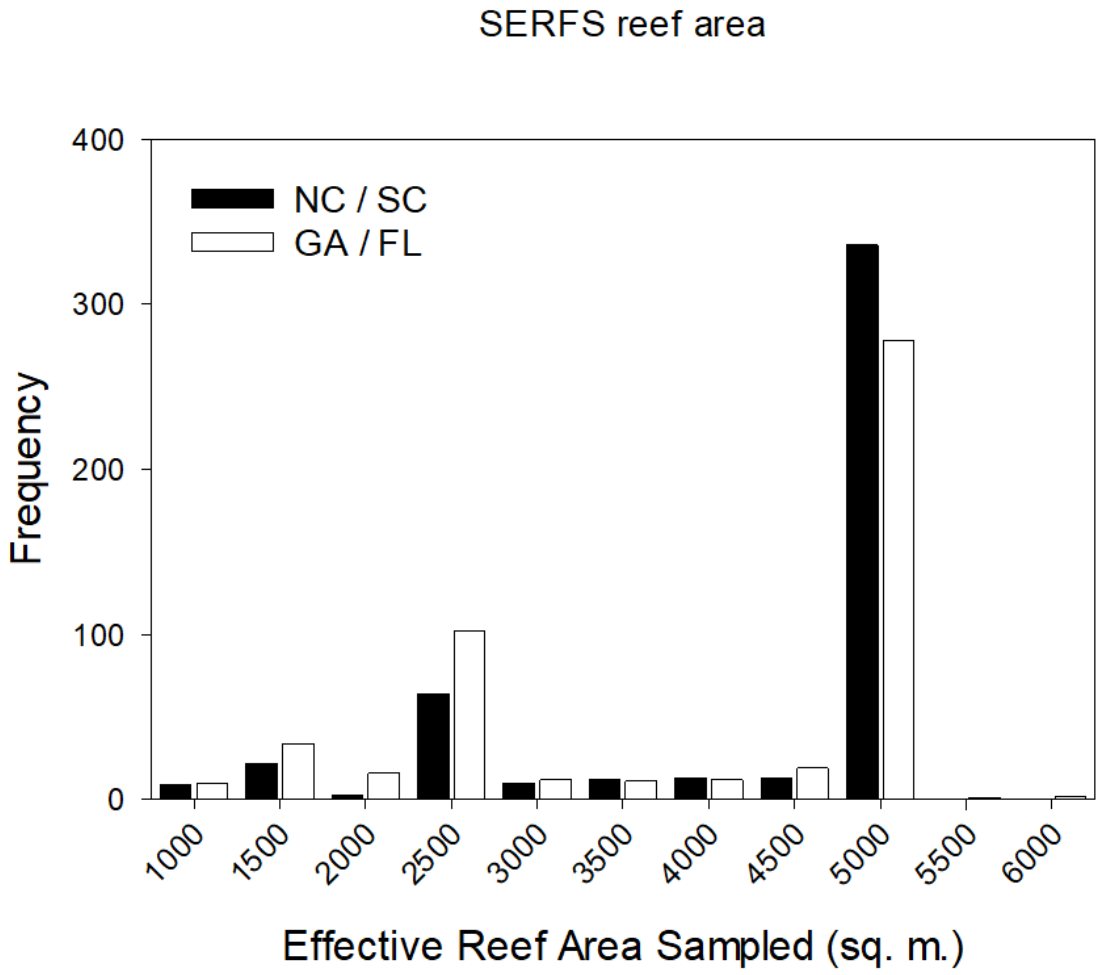


Figure 11.4.13. Mean CPUE (individuals per site) of Greater Amberjack by area (NC: North Carolina; SC: South Carolina) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the SERFS survey of natural hardbottom habitats. Values represent converted spherical camera abundance currency.

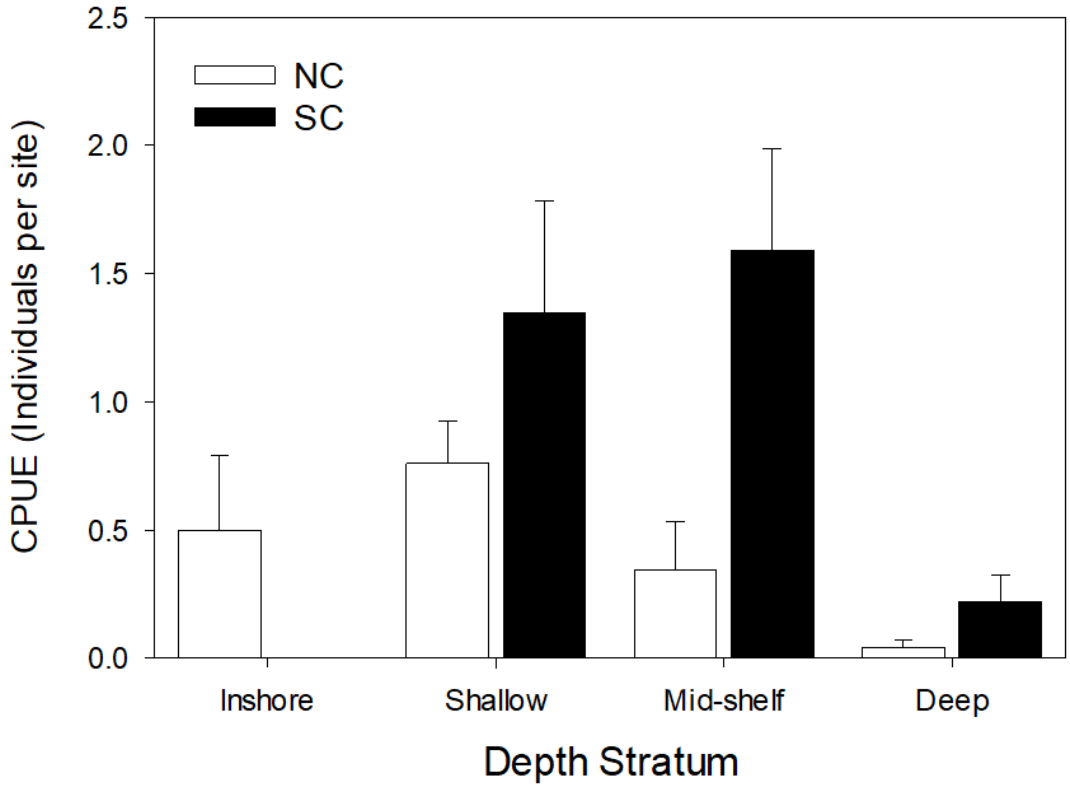


Figure 11.4.14. Mean CPUE (individuals per site) of Greater Amberjack by area (GA: Georgia; NF: North Florida; MF: Mid Florida) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the SERFS survey of natural hardbottom habitats. Values represent converted spherical camera abundance currency. Note that SERFS sampling did not extend into South Florida, and no SERFS sampling was conducted in the Deep stratum in Mid Florida.

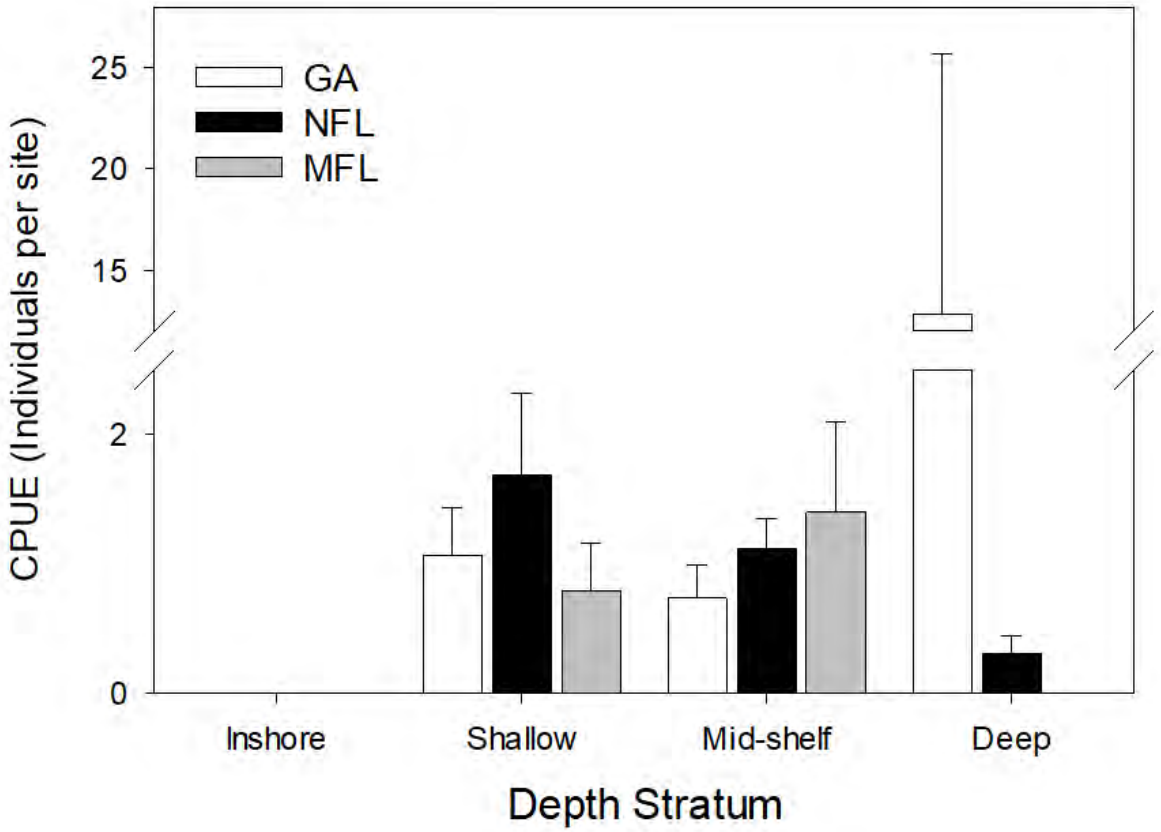


Figure 11.4.15. Regional frequency distribution of MaxN values from the S-BRUV video survey of natural hardbottom habitats in the South Atlantic.

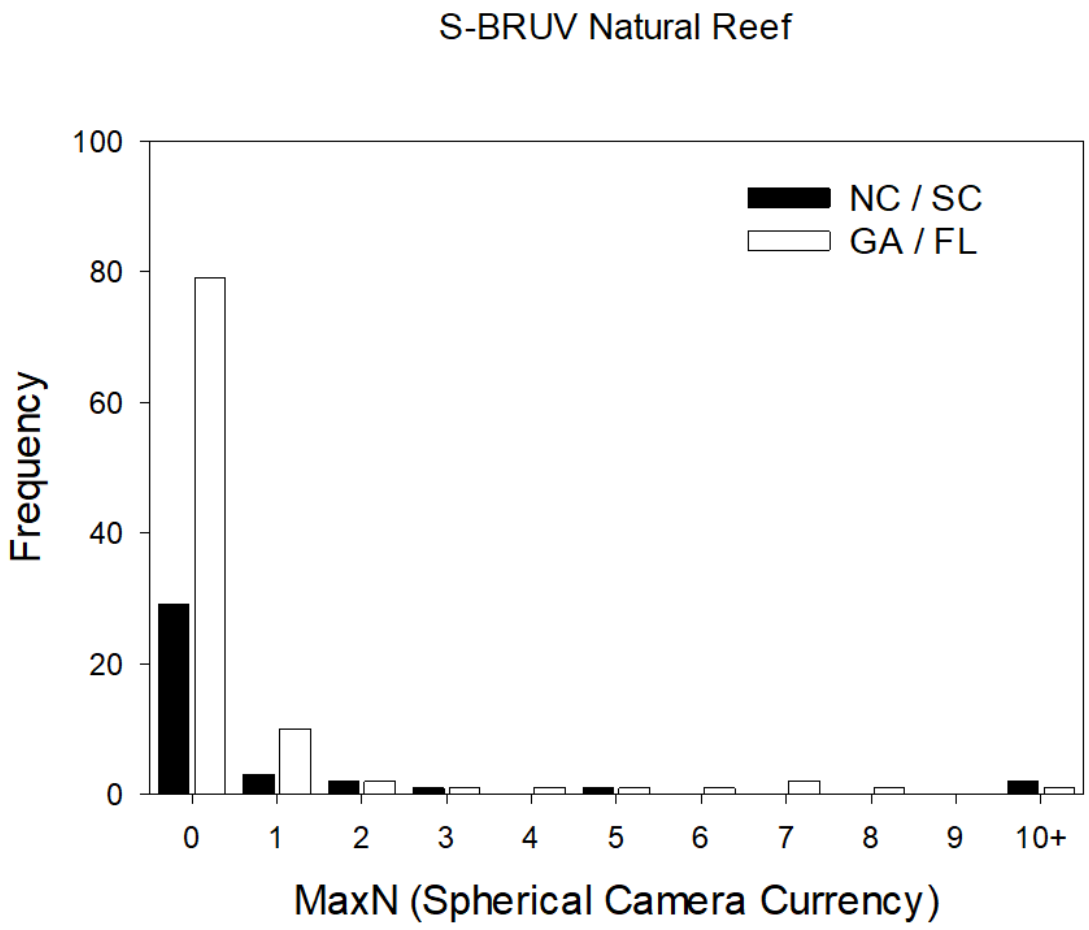


Figure 11.4.16. Regional frequency distribution of the effective reef area sampled during the S-BRUV video survey of natural hardbottom habitats in the South Atlantic.

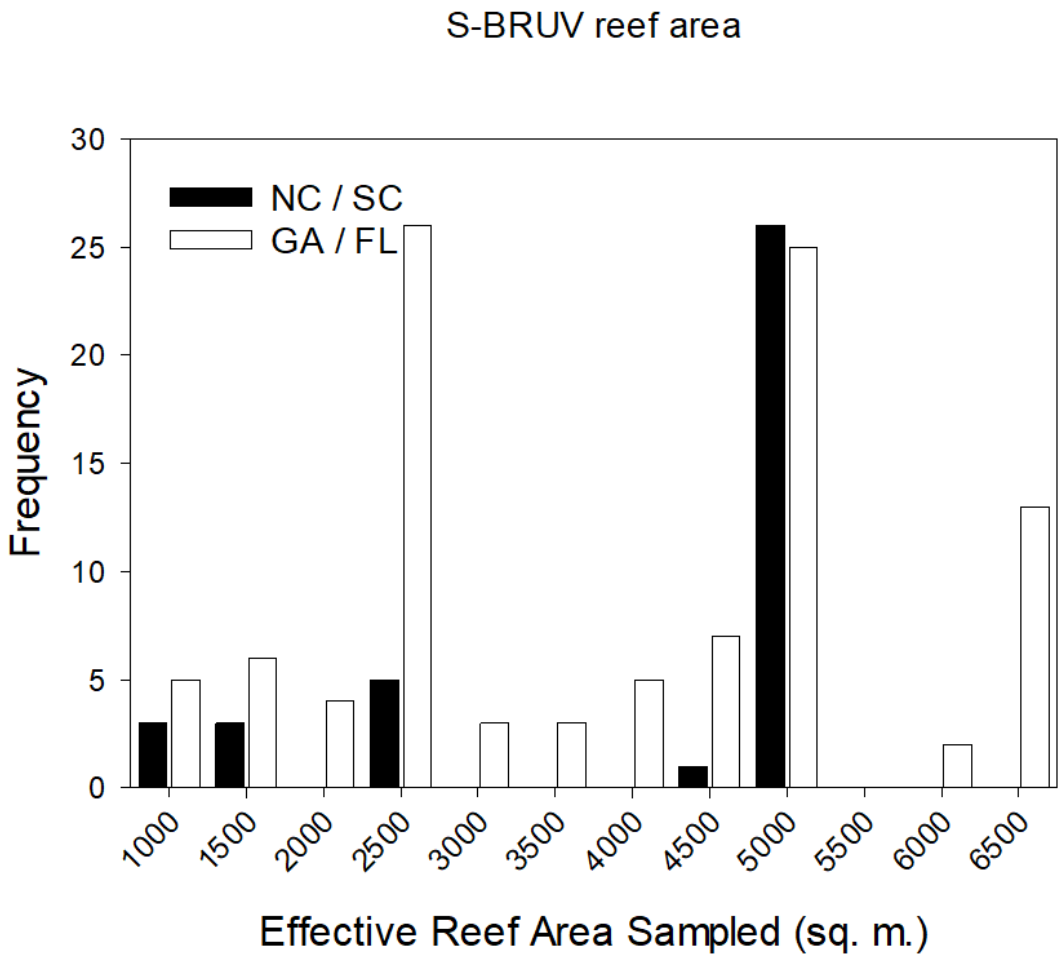


Figure 11.4.17. Mean CPUE (individuals per site) of Greater Amberjack by area (NC: North Carolina; SC: South Carolina) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the S-BRUV survey of natural hardbottom habitats. Note that no S-BRUV sampling occurred within the Mid-shelf depth stratum (both NC and SC), and the Deep stratum in SC.

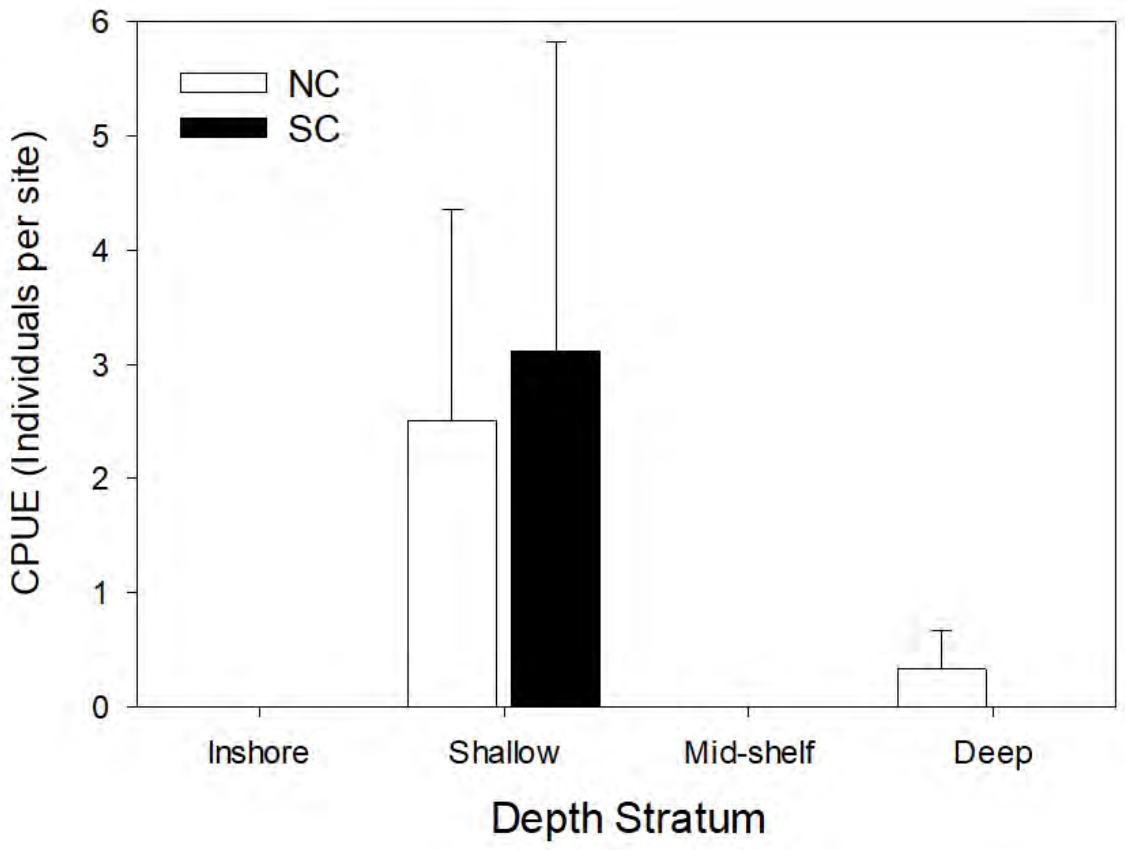


Figure 11.4.18. Mean CPUE (individuals per site) of Greater Amberjack by area (GA: Georgia; NFL: North Florida; MFL: Mid Florida; SFL: South Florida) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the S-BRUV survey of natural hardbottom habitats. Values represent converted spherical camera abundance currency. Note that no S-BRUV sampling occurred within the Mid-shelf depth stratum (GA, NFL, MFL) or Deep depth stratum (GA, MFL) for some areas.

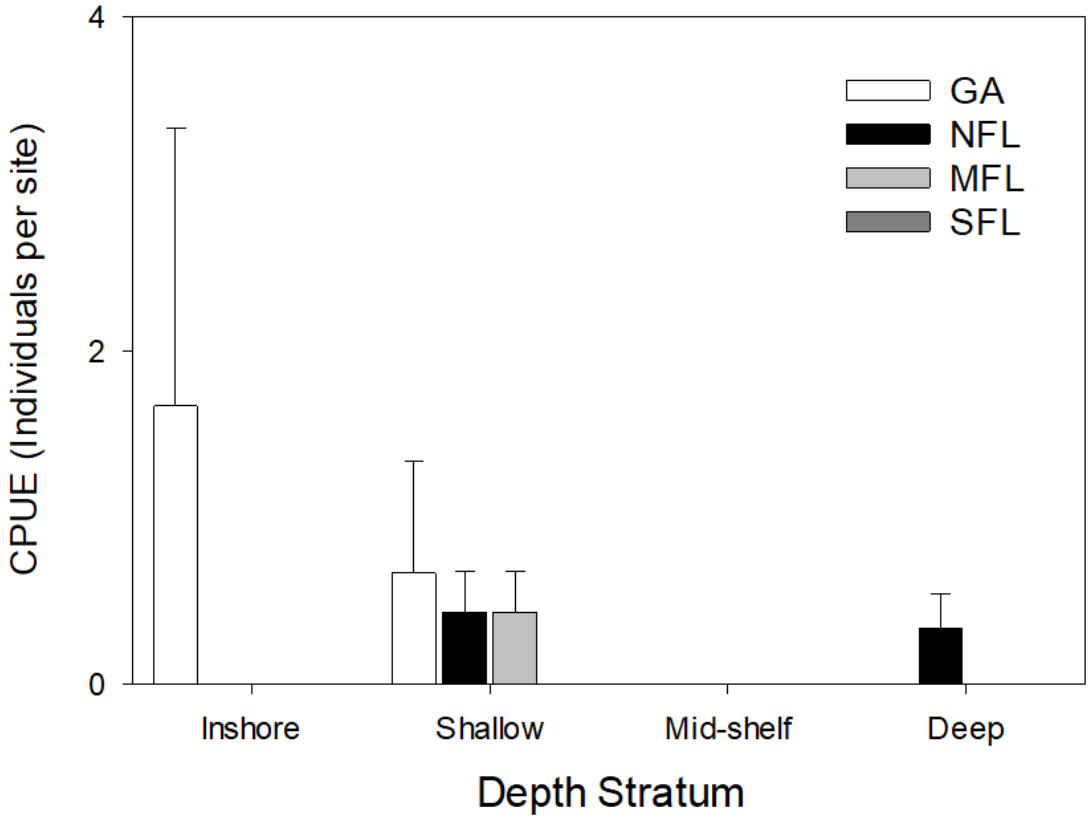


Figure 11.4.19. Regional frequency distribution of MaxN values from the S-BRUV video survey of artificial hardbottom habitats in the South Atlantic.

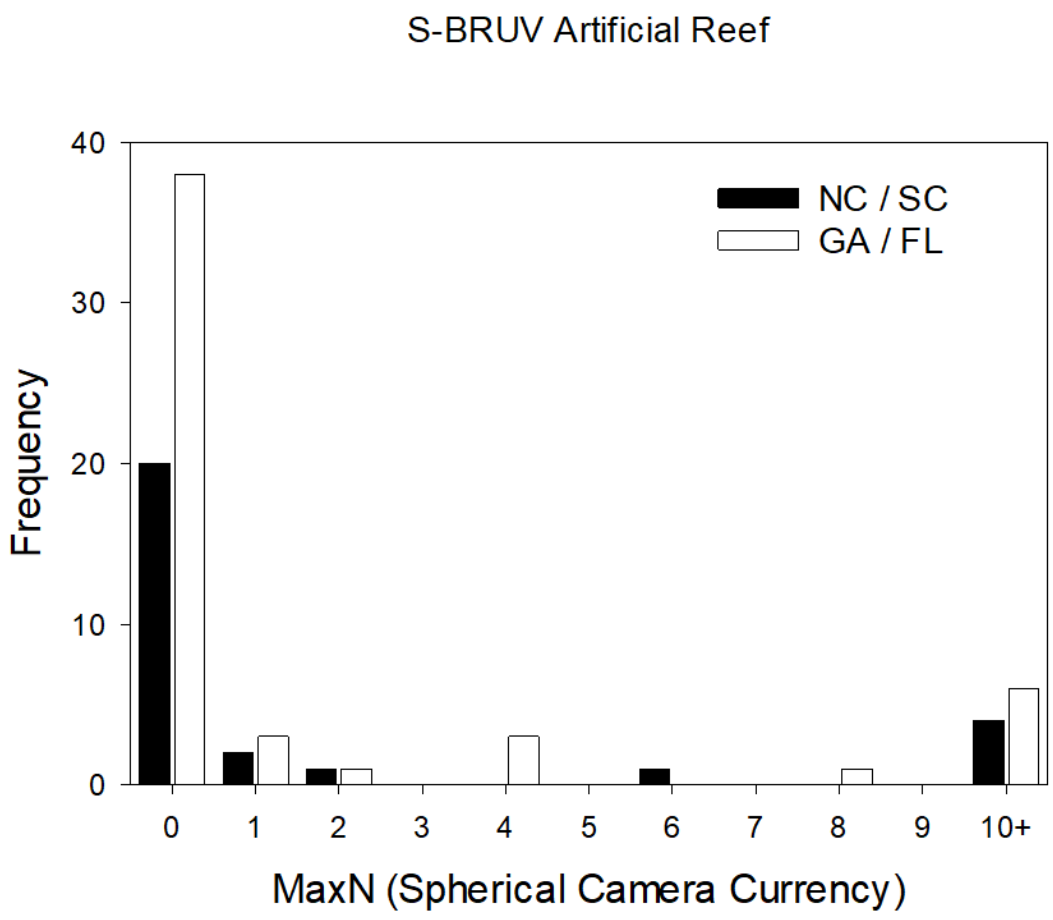


Figure 11.4.20. Regional frequency distribution of the effective reef area sampled during the S-BRUV video survey of artificial hardbottom habitats in the South Atlantic.

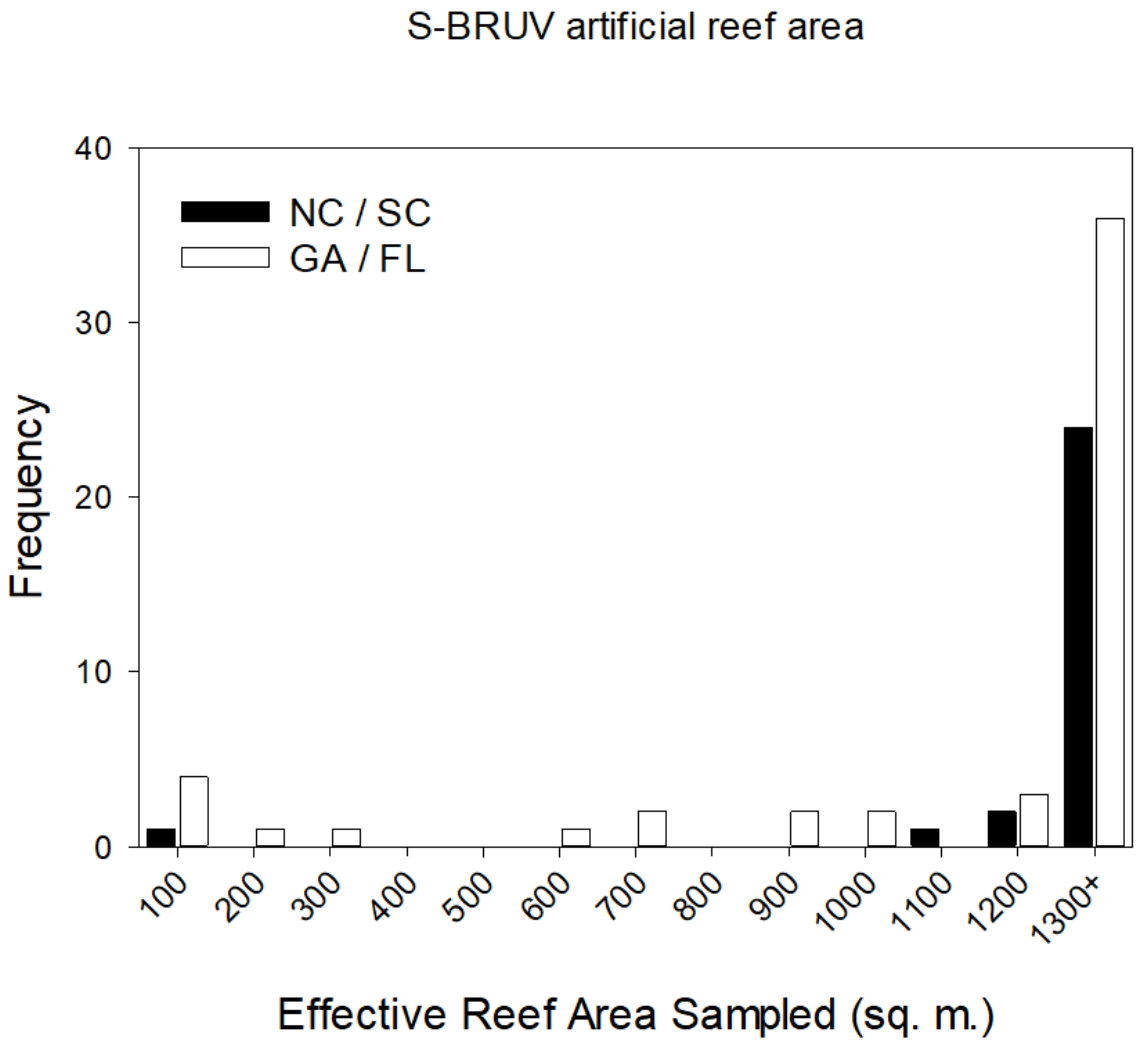


Figure 11.4.21. Mean CPUE (individuals per site) of Greater Amberjack by area (NC: North Carolina; SC: South Carolina) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the S-BRUV survey of artificial hardbottom habitats. Note that no S-BRUV sampling occurred within the Mid-shelf or Deep strata for artificial reefs.

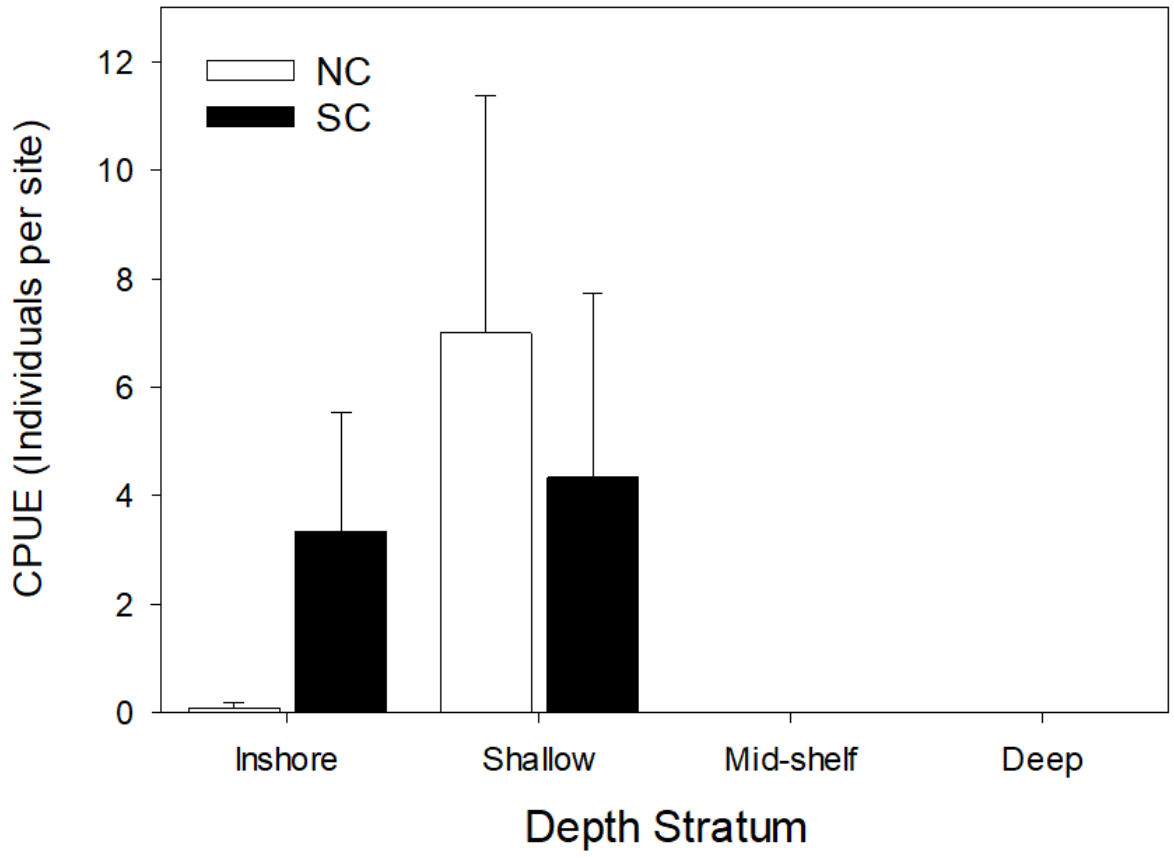


Figure 11.4.22. Mean CPUE (individuals per site) of Greater Amberjack by area (GA: Georgia; NF: North Florida; MF: Mid Florida; SF: South Florida) and depth strata (Inshore: 10 – 20 m; Shallow: 20 – 40 m; Mid-shelf: 40 – 70 m; Deep: 70 – 150 m) from the S-BRUV survey of artificial hardbottom habitats. Aside from MFL and SFL, no sampling was conducted in the Mid-shelf and Deep depth strata.

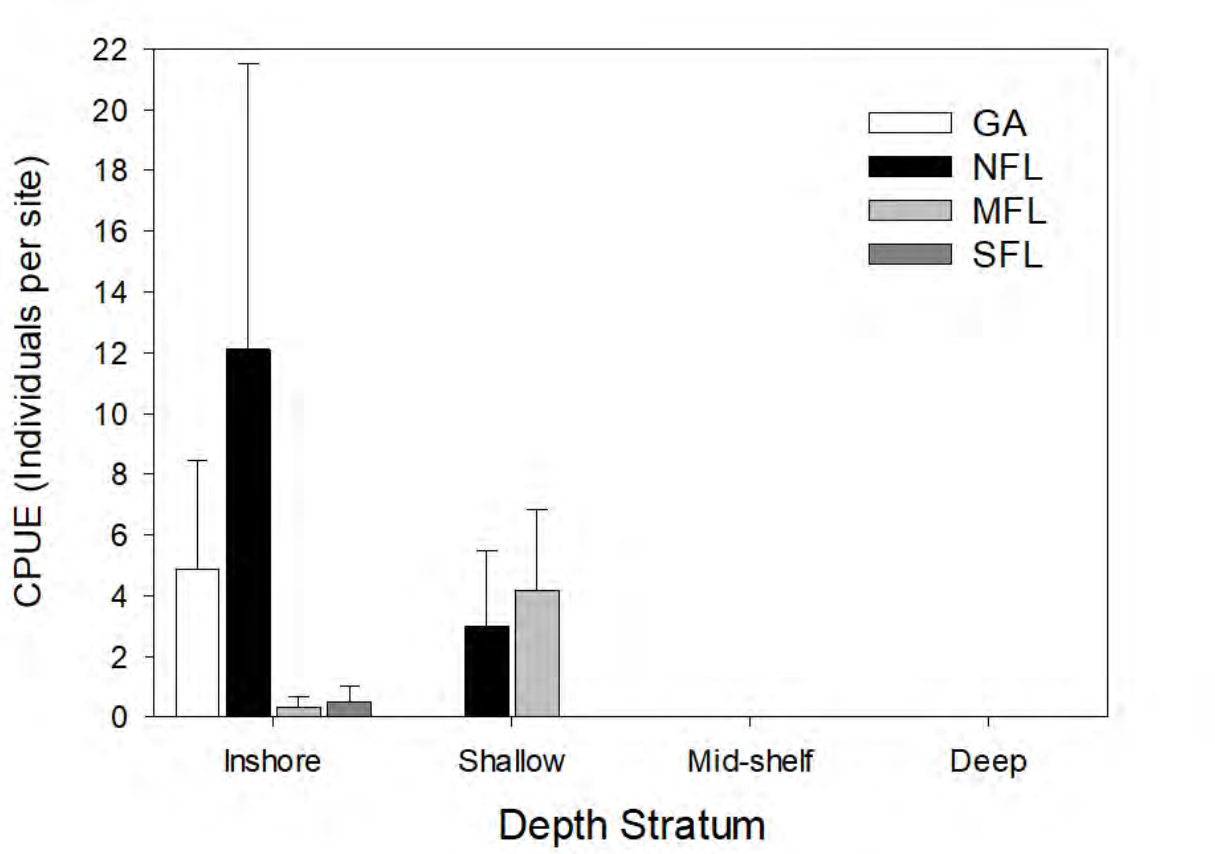


Figure 11.4.23. Map of all towed C-BASS sampling transects conducted targeting unconsolidated sediment habitats in the South Atlantic. The location pointed to by the red arrow is the only site where Greater Amberjack were observed on unconsolidated sediment habitats (N = 1).

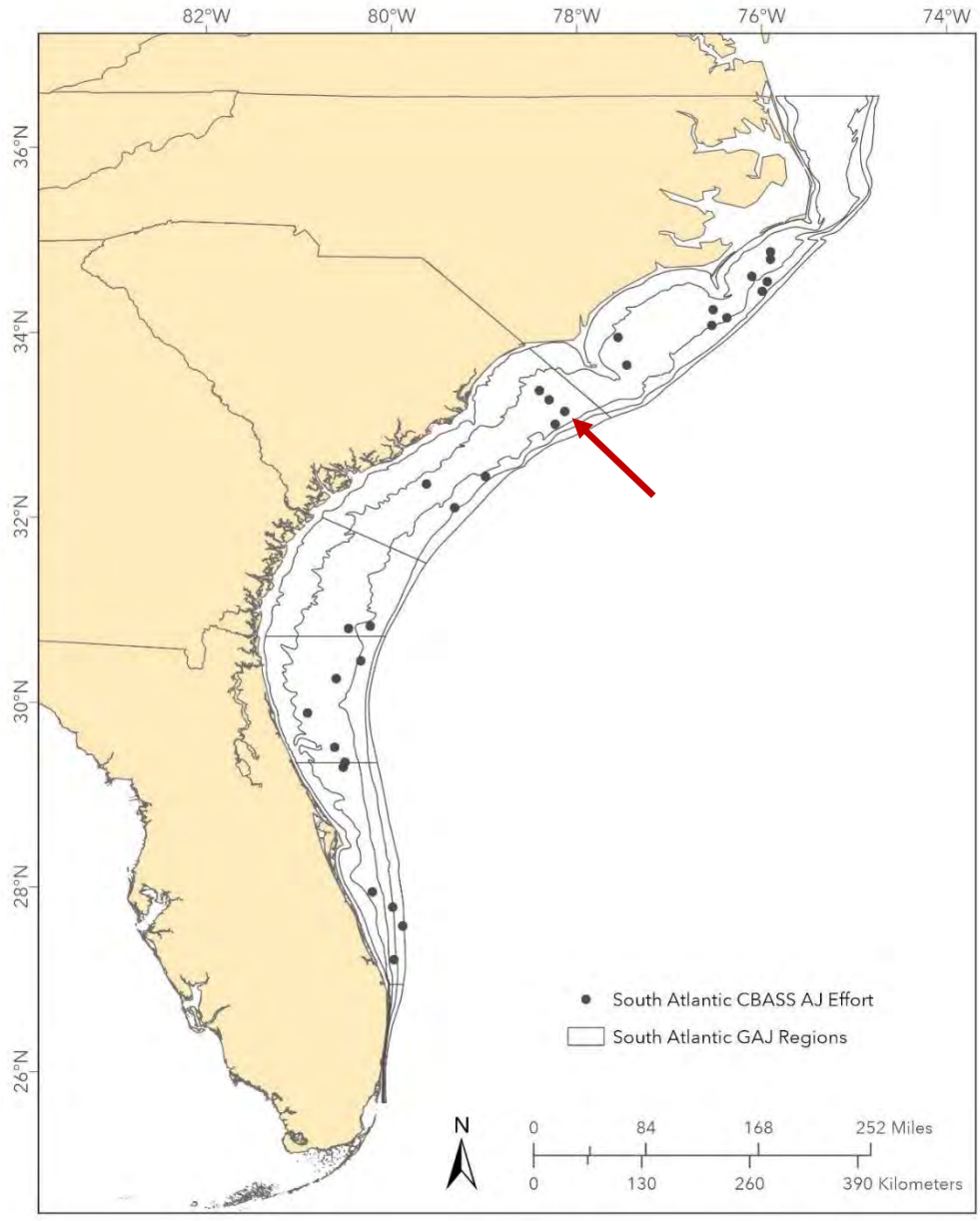
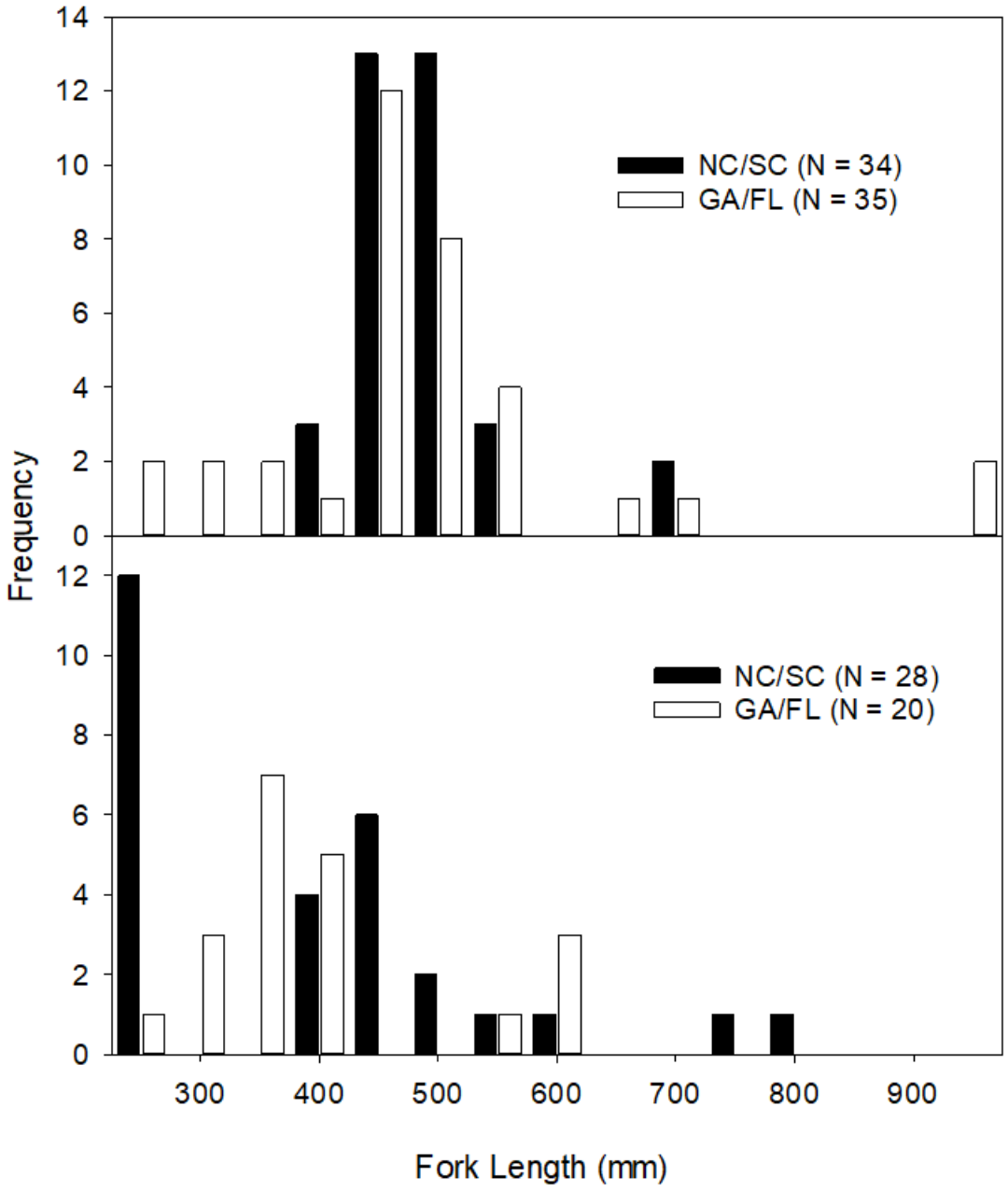


Figure 11.4.24. Regional frequency plots of stereo measurement data from FWRI video surveys on natural and artificial reefs using S-BRUVS in the South Atlantic.

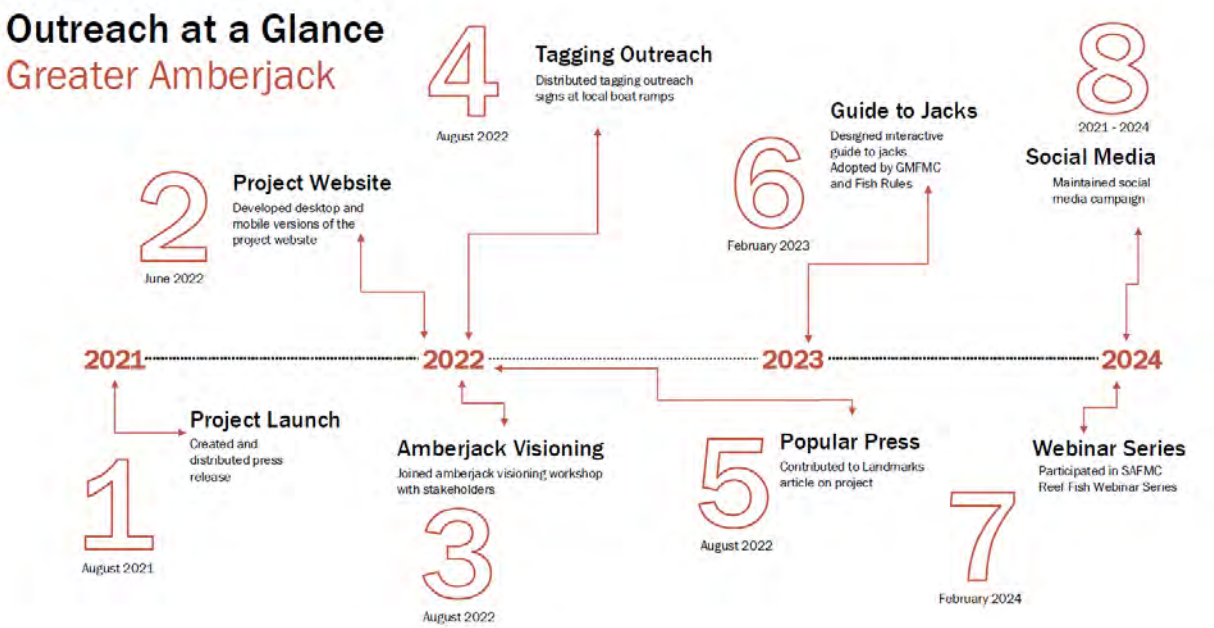


12. Public Outreach

J. Marcus Drymon

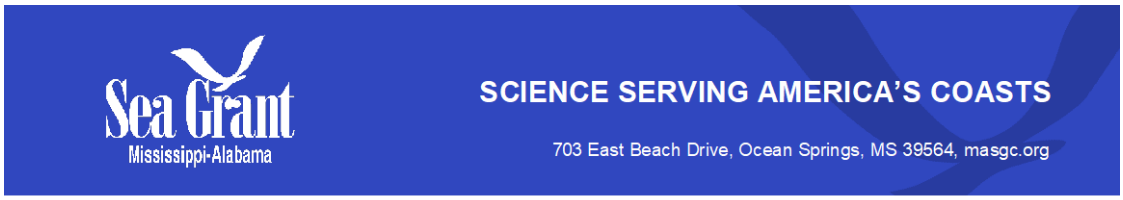
SUMMARY

The public outreach portion of the Greater Amberjack Count was implemented by the Reef Fish Extension Program, a separately funded, concurrent effort established to support outreach and extension for regional reef fish projects (e.g., the Greater Amberjack Count). Outreach and extension efforts related to the Greater Amberjack Count are summarized in the graphic below. Examples of all eight components of outreach and extension for the Greater Amberjack Count are included.



12.1. Project Launch

The initial press release for this project was launched on August 25, 2021. The press release was distributed through NOAA Fisheries and the National Sea Grant Office, as well as through an extensive email distribution list collated and maintained by the Reef Fish Extension Program.



FOR RELEASE ON AUG. 25, 2021

Contacts: Melissa Schneider, Communications Coordinator, Mississippi-Alabama Sea Grant Consortium, melissa.schneider@usm.edu, 228-818-8838; Media inquiries: Lance Crawford, Director of Media Relations, University of South Alabama, lancecrawford@southalabama.edu, 251-460-6211 (office), 610-716-1218 (cell)

Team selected to estimate abundance of Greater Amberjack in South Atlantic, Gulf of Mexico

(OCEAN SPRINGS, Miss.) – A team of scientists, selected through a peer-review process organized by the Mississippi-Alabama Sea Grant Consortium, will conduct an independent study to estimate the number of Greater Amberjack in the U.S. waters of the South Atlantic and Gulf of Mexico regions. The team will also investigate the fish's movement and how it is distributed by habitat, including artificial, natural and uncharacterized habitats.

The research team, made up of 18 scientists from 13 institutions, was awarded $9 million in federal funds for the two-year project. With matching funds from the institutions, the project will total $11.7 million.

In fiscal year 2020, Congress appropriated $5 million for the National Sea Grant College Program and $5 million for the National Marine Fisheries Service (NMFS) for Sea Grant to implement a research program independent of NMFS to support the Greater Amberjack stock assessment process.

"The Greater Amberjack abundance study in the South Atlantic and Gulf of Mexico is one way Sea Grant is helping to ensure a safe and sustainable supply of seafood products for current and future generations," said National Sea Grant College Program Director Jonathan Pennock. "This project will provide valuable information for the management of Greater Amberjack as well as insight useful to future fisheries research and engagement."

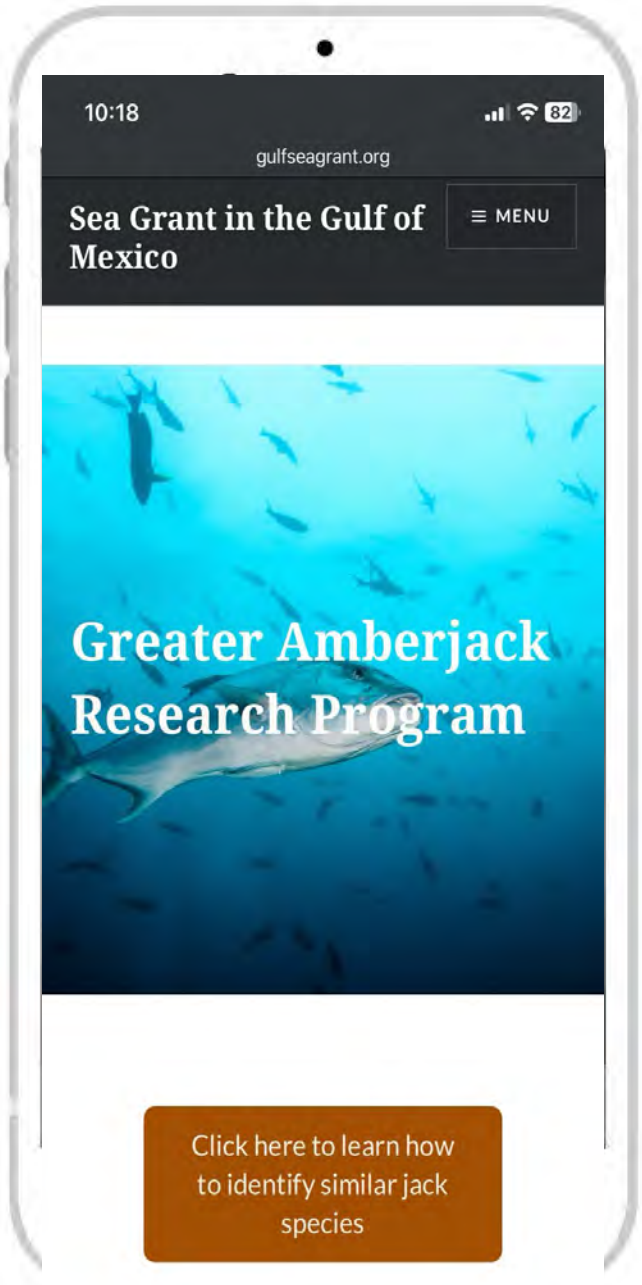
Interdisciplinary team spans regions

The selected interdisciplinary research team includes geneticists, modelers, statisticians, ecologists and oceanographers from along the Atlantic and Gulf coasts. Project leader Sean Powers, chair of the Marine Sciences Department at the University of South Alabama and a senior marine scientist with Dauphin Island Sea Lab, said many of the project researchers have worked together for more than 20 years and have helped advance the science that makes these large-scale studies possible.

Page **1** of **4**

12.2. Project Website

Mobile and desktop versions of the project website were developed to share information about the Greater Amberjack Count. The website also functioned as a repository for products developed in support of the Greater Amberjack Count. To date, the website has received more than 2,500 unique visits.



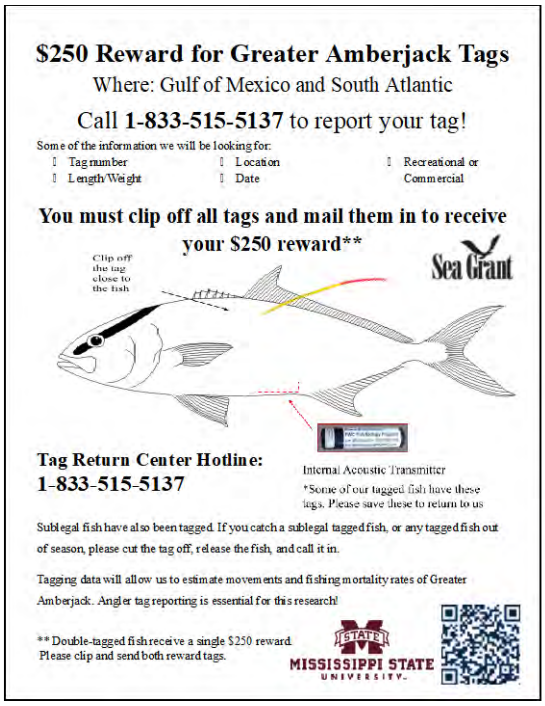
12.3. Amberjack Visioning

In August 2022, a team led by researchers at the University of Florida hosted a Greater Amberjack visioning workshop. The goals of this workshop were to engage with stakeholders to identify and prioritize regional research needs. This workshop included stakeholders from the Gulf of Mexico and the Southeastern United States. Members from the Greater Amberjack Count project team gave presentations on the research approach and the outreach and extension plan.



12.4. Tagging Outreach

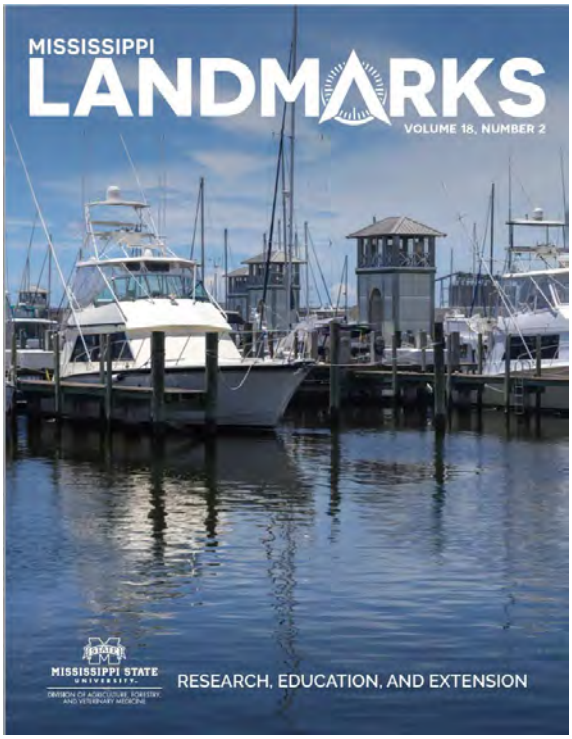
Throughout the fall of 2022, concerted efforts were made to establish and increase public awareness of the tagging component of the Greater Amberjack Count. State-specific signs about tag returns were created and distributed to each state (Texas, Louisiana, Mississippi, Alabama, Florida, Georgia, South Carolina, and North Carolina); these signs included the reward amount ($250), the placement of internal and external tags, the phone number to call, and a QR code linked to more information. Shown below is the sign for Mississippi and the locations where the Mississippi sign was placed.



Location	City	Type
Bay St. Louis Harbor	Bay St. Louis	Boat Launch
Cedar Point Boat Launch	Bay St. Louis	Boat Launch
Shaggys Marina	Pass Christian	Boat Launch
Pass Christian Harbor	Pass Christian	Boat Launch
Pass Christian Yacht Club	Pass Christian	Boat Launch
Long Beach Harbor	Long Beach	Boat Launch
Jones Park Marina	Gulfport	Boat Launch
Jones Park Bait Shop	Gulfport	Shop
West Marine	Gulfport	Shop
Biloxi Port Division	Biloxi	Boat Launch
Gorenflo's Bait and Tackle	Biloxi	Shop
Biloxi Small Craft Harbor	Biloxi	Boat Launch
Ocean Springs Harbor	Ocean Springs	Boat Launch
Ocean Springs Marine Mart	Ocean Springs	Shop
Lake Mars Boat Launch	Ocean springs	Boat Launch
Lighthouse Park Boat Launch	Pascagoula	Boat Launch
Pascagoula Little River Park Boat Launch	Pascagoula	Boat Launch
Pascagoula Point Park Boat Launch	Pascagoula	Boat Launch
Back Bay Boat Launch	Biloxi	Boat Launch
Academy Sports	Biloxi	Shop

12.5. Popular Press

To further increase the reach of the Greater Amberjack Count, we engaged popular press outlets (e.g., Mississippi Landmarks). The images and text below are from an article published in October 2022.

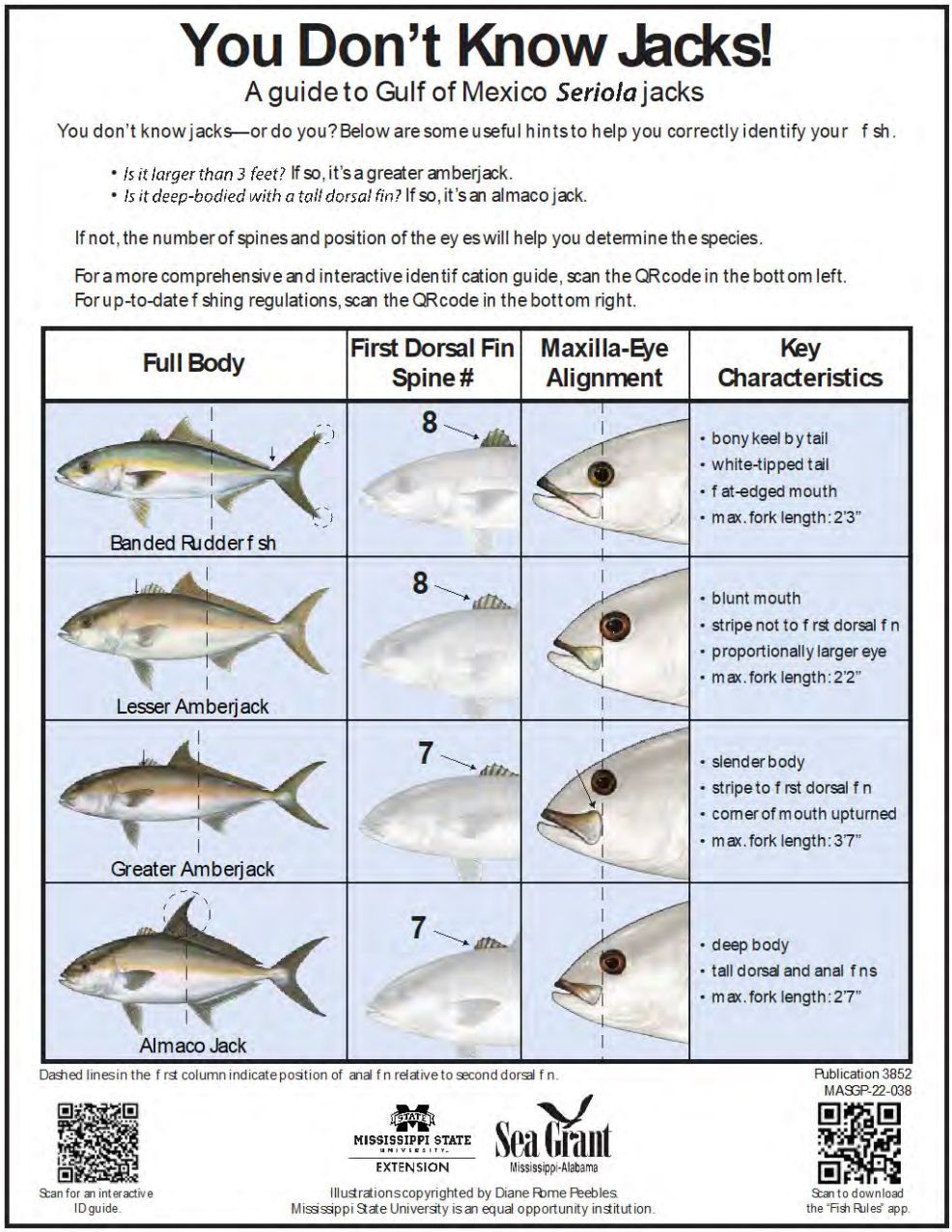






12.6. Guide to Jacks

At the inception of the project, the Reef Fish Extension team discussed the need for an easy-to-use, detailed guide for anglers to differentiate between Greater Amberjack and three congeners (Lesser Amberjack, Banded Rudderfish, and Almaco Jack). The team determined that a two-pronged formatting approach was needed to reach the widest possible audience. Accordingly, we developed a single-paged visual guide (fact sheet), but included a QR code to a digital, interactive version of the guide where interested anglers could access more details. This guide was ultimately adopted by the Gulf of Mexico Fishery Management Council and subsequently incorporated into the popular fishing regulations app "Fish Rules." Both formats of this guide are shown below.



You Don't Know Jacks!

A guide to Gulf of Mexico *Seriola* jacks

You don't know jacks—or do you? Below are some useful hints to help you correctly identify your fish.

- *Is it larger than 3 feet?* If so, it's a greater amberjack.
- *Is it deep-bodied with a tall dorsal fin?* If so, it's an almaco jack.

If not, the number of spines and position of the eyes will help you determine the species.

For a more comprehensive and interactive identification guide, scan the QR code in the bottom left.
For up-to-date fishing regulations, scan the QR code in the bottom right.

Full Body	First Dorsal Fin Spine #	Maxilla-Eye Alignment	Key Characteristics
Banded Rudderfish	8		• bony keel by tail • white-tipped tail • flat-edged mouth • max. fork length: 2'3"
Lesser Amberjack	8		• blunt mouth • stripe not to first dorsal fin • proportionally larger eye • max. fork length: 2'2"
Greater Amberjack	7		• slender body • stripe to first dorsal fin • corner of mouth upturned • max. fork length: 3'7"
Almaco Jack	7		• deep body • tall dorsal and anal fins • max. fork length: 2'7"

Dashed lines in the first column indicate position of anal fin relative to second dorsal fin.

Publication 3852
MASGP-22-038

Scan for an interactive ID guide.

MISSISSIPPI STATE UNIVERSITY EXTENSION

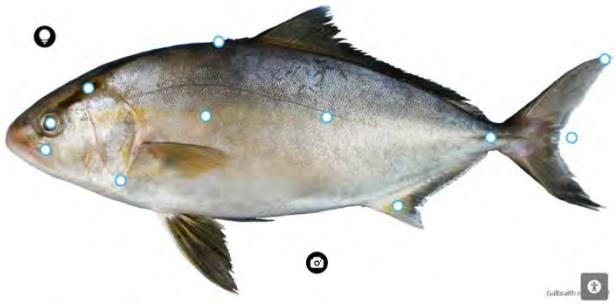
Sea Grant Mississippi-Alabama

Illustrations copyrighted by Diane Rome Peebles.
Mississippi State University is an equal opportunity institution.

Scan to download the "Fish Rules" app.

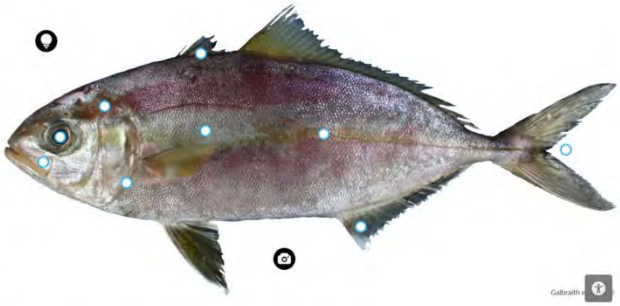
12.6. Guide to jacks (cont.)

Banded rudderfish (*Seriola zonata*)



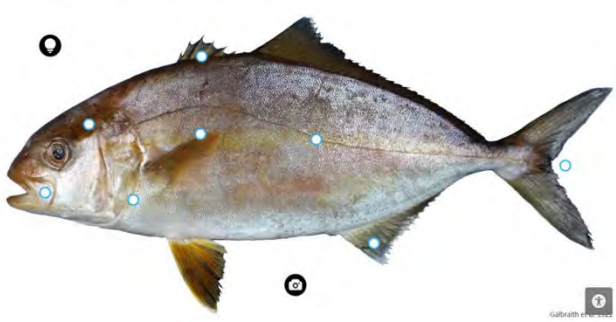
Click here to view the accessible version of this interactive content

Lesser amberjack (*Seriola fasciata*)



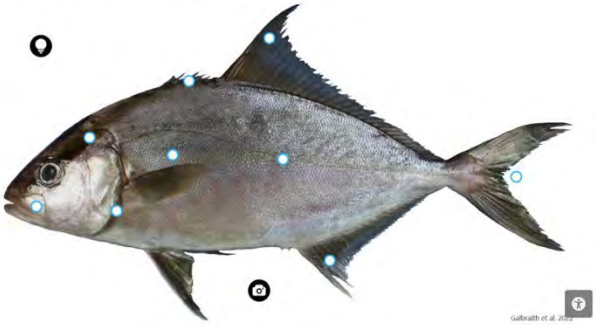
Click here to view the accessible version of this interactive content

Greater amberjack (*Seriola dumerili*)

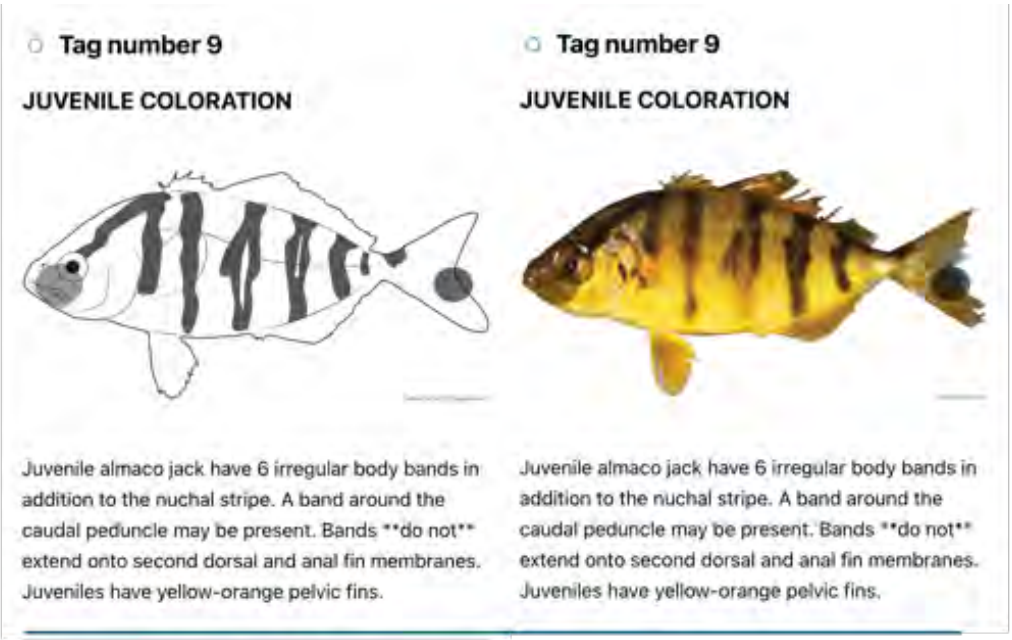


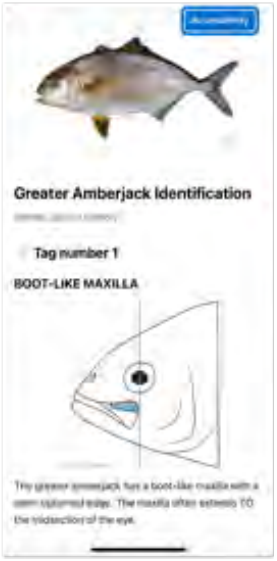
Click here to view the accessible version of this interactive content

Almaco jack (*Seriola rivoliana*)



Click here to view the accessible version of this interactive content


Tag number 9
JUVENILE COLORATION
Juvenile almaco jack have 6 irregular body bands in addition to the nuchal stripe. A band around the caudal peduncle may be present. Bands **do not** extend onto second dorsal and anal fin membranes. Juveniles have yellow-orange pelvic fins.
Tag number 9
JUVENILE COLORATION
Juvenile almaco jack have 6 irregular body bands in addition to the nuchal stripe. A band around the caudal peduncle may be present. Bands **do not** extend onto second dorsal and anal fin membranes. Juveniles have yellow-orange pelvic fins.


Greater Amberjack Identification
Tag number 1
BOOT-LIKE MAXILLA


Accessibility
Tag number 2
ANAL FIN LENGTH
The greater amberjack has an anal fin that is approximately **2/3** the length of the second dorsal fin.
Tag number 3
The greater amberjack has a nuchal stripe that extends from the snout **into** the first dorsal fin.


Tag number 4
FORK LENGTH
Tag number 5
ADDITIONAL PHOTOS

12.6. Guide to jacks (cont.)

Accessibility

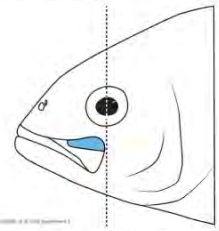


Greater Amberjack Identification

Seriola Jacks (Author)

Tag number 1

BOOT-LIKE MAXILLA



The greater amberjack has a boot-like maxilla with a semi-upturned edge. The maxilla often extends TO the midsection of the eye.

Accessibility

Tag number 2

ANAL FIN LENGTH



The greater amberjack has an anal fin that is approximately **2/3** the length of the second dorsal fin.

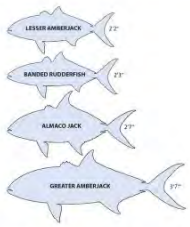
Tag number 3

The greater amberjack has a nuchal stripe that extends from the snout **into** the first dorsal fin.

Accessibility

Tag number 4

FORK LENGTH



The greater amberjack is the largest of the four species, with a maximum fork length of 3 ft 7 in.

Tag number 5

ADDITIONAL PHOTOS



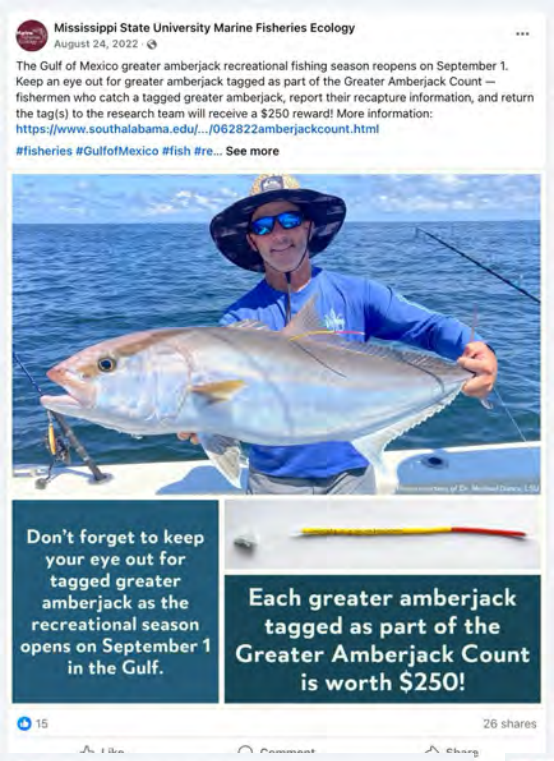
12.7. Webinar Series

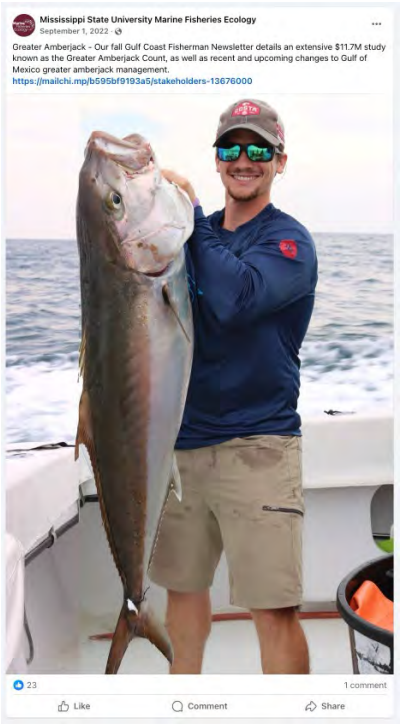
In February 2024, researchers from the Greater Amberjack Count presented as part of the Reef Fish Webinar series hosted by the Reef Fish Extension Fellow, who is housed at the South Atlantic Fishery Management Council (SAFMC). The seminar was well attended, and the recording was viewed 81 times. Below is the social media post pertaining to the webinar, which is archived on the SAFMC's website.

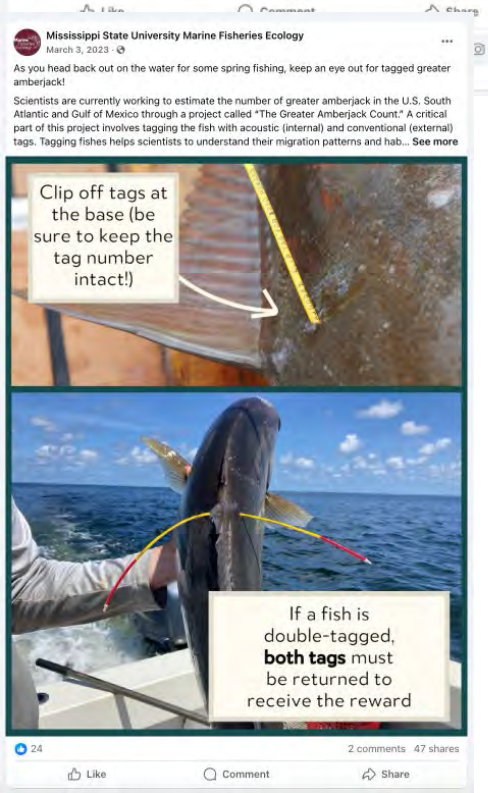


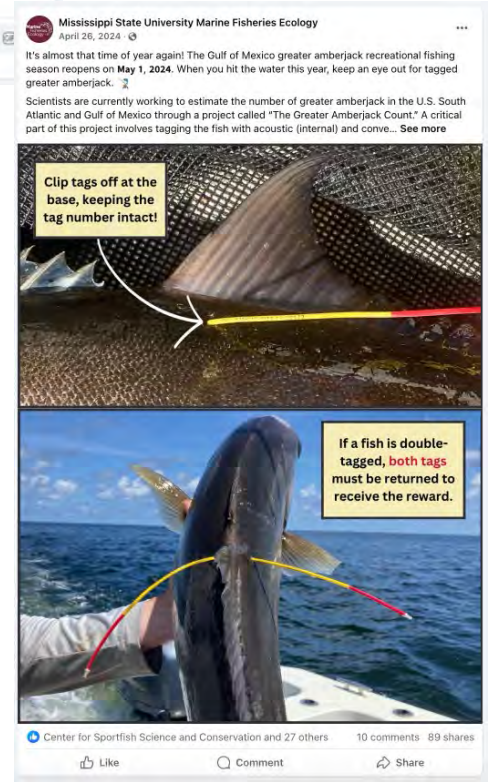
12.8. Social Media

A dedicated social media campaign was executed throughout the project. Examples of social media posts are shown below.









13. Co-produced science-angler involvement

Savannah H. Swinea, Sean P. Powers, Mark A. Albins, J. Marcus Drymon, Steven B. Scyphers

Objectives

With an aim to harness diverse knowledge systems to understand Greater Amberjack, we deployed a social survey targeting South Atlantic and Gulf of Mexico recreational fishermen. We sought to understand the knowledge, attitudes, and intended behavior of Greater Amberjack recreational fishermen in order to not only characterize the social dimensions of the Greater Amberjack fishery, but to solicit information concerning Greater Amberjack ecology maintained within fishermen knowledge systems. The goals of the survey were fourfold: (1) characterize fishermen perceptions of temporal trends in Greater Amberjack abundance; (2) contextualize Greater Amberjack perceptions against better-studied reef fish species; (3) characterize perceptions of Greater Amberjack habitat preferences; and (4) map reported Greater Amberjack recreational fishing behavior and evaluate drivers of this reported behavior using an in-survey experiment.

Methods

The human subjects research in our study was approved by Northeastern University's Institutional Review Board (IRB #13-07-16), and informed consent was acquired from all participants.

Survey instrument and experimental design

The survey instrument was divided into three major sections. The first section examined fishing characteristics, behaviors, and attitudes, especially for reef fish species. Respondents indicated the state in which they had fished the most in the last 12 months and fishing region (Gulf of Mexico or South Atlantic) was coded from these states. Respondents also reported fishing experience in years and offshore fishing days in the last 12 months. Respondents were then asked to rate the importance of fishing for reef fish in the context of all their fishing. Respondents who chose an option besides *Not at all important* were shown a list of 36 species in the reef fish complex in the South Atlantic and Gulf of Mexico regions and were prompted to indicate important, very important, and a single-most important species from that list. Respondents indicated reasons for fishing for their single-most important reef fish species that were developed from early project visioning phases to set research priorities (Lorenzen et al. 2021).

The second section of the survey captured Greater Amberjack perceptions. We coded a dummy variable that characterized respondents as Greater Amberjack fishermen if they indicated that (1) Greater Amberjack was an important reef fish species, (2) they had ever caught a Greater Amberjack, or (3) they had ever targeted Greater Amberjack. Survey items measuring Greater Amberjack perceptions were only displayed to Greater Amberjack fishermen identified using this coding scheme. These items included trends in abundance as well as when abundance started changing, habitat preference for Greater Amberjack fishing, and expected catch-per-unit-effort for sites with their preferred habitat. To facilitate direct comparisons between survey items concerning a respondent's most important reef fish species and Greater Amberjack, an identical question in the first and second section of the survey addressed reasons for fishing.

The third and final section of the survey involved a participatory mapping exercise to assess potential implications of collaborative tagging projects to influence reported fishing behavior. To develop the participatory mapping survey questions, we produced map layouts in ArcGIS Pro (ESRI 2024) to display the

coastal and open-ocean areas of all South Atlantic and Gulf of Mexico states. We adapted a NOAA Fisheries grid-based system to display 0.5x0.5 degree grid cells enclosed by the 200-nautical mile Exclusive Economic Zone over which the United States maintains territorial fishing rights. Grid cell size was optimized based on three factors: maximizing applicability across scales; ensuring accuracy where issues of privacy can be at stake; and minimizing precision error given the digital platform on which respondents were selecting individual grid cells. This modification allowed us to collect spatial information on a finer scale, but also facilitates aggregation and integration with existing mapping methods for this area. Each map layout contained layers displaying detailed state coastlines, state boundaries, and depth contours starting at 50 feet and extending out every 100 feet in the 100-500 foot depth region. To encourage respondents to quickly become oriented within the map layout, we also displayed relevant coastal cities including those with large populations (at least 10,000 residents), as well as NOAA fishing communities.

Respondents were filtered into the participatory mapping exercise using the same coding scheme that was described to capture Greater Amberjack perceptions in the previous section of the survey. Respondents were randomly assigned to one of two treatments: in the control treatment, respondents were prompted to highlight one or more grid cells in their state's fishing territory where they would go to target and catch a Greater Amberjack. The tagging treatment was framed as a hypothetical scenario in which the respondent was participating in a collaborative project with scientists to tag Greater Amberjack and were also asked to highlight one or more grid cells. This exercise was designed to assess if reported spatial fishing effort differed depending on if the exercise was framed simply to describe fishing areas for Greater Amberjack, or as an opportunity to co-produce knowledge with scientists.

Data collection, quality assurance, and quality control

Data were collected in Qualtrics Research Panels, an online service that has rapidly gained popularity as a cost-effective and robust panel sampling platform (Zack et al. 2019). Quotas of 200 respondents from each state were set to achieve equal coverage across the study area. The only exception was Florida for which the state was divided into Atlantic and Gulf coasts and separate quotas were assigned to each coast.

Given drawbacks associated with nonprobability sampling methods, we implemented multiple strategies before and after survey dissemination to ensure response quality. Only respondents who indicated that they would "provide their best answers" in a commitment to quality to survey item, as well as reported at least one saltwater fishing trip within the past 12 months, were allowed to proceed into the survey. After data collection, responses were reviewed and qualified as definitely bad if they had any of the following characteristics: "straight-lining" (consistent selections of the same answer), "speeding" or "inattention" (completion times too fast or slow in comparison to the median time), duplicate locations or other variables, or invalid IP addresses (Zhang and Conrad 2014). We then examined the remaining responses and qualified them as possibly bad if they contained irrelevant or similar open-ended responses. This resulted in 17% of responses being flagged as definitely bad, 4% as possibly bad, and 79% as not suspicious. All bad responses were removed and replaced by Qualtrics, and new responses underwent the same review process as was detailed above. Data collection commenced and was concluded in June 2022 with n = 1,819 responses.

Analyses

All statistical analyses were performed in R Version 4.4.0 (R Core Team 2024) or the Statistical Package for the Social Sciences (SPSS) Version 28 (IBM Corp. 2021). Differences concerning perceptions of Greater Amberjack habitat preferences were tested using Fisher's exact tests across region and state. Kruskal-Wallis tests were used

to evaluate differences in perceptions of abundance trends by region and state and expected catch-per-unit-effort by habitat preference. To understand how motivations differed for Greater Amberjack in comparison to other reef fish species, respondents were filtered such that they were identified as Greater Amberjack fishermen and chose Red Snapper as their most important reef fish species. Reasons for fishing were compared across these two species using non-parametric McNemar tests which are suitable for paired nominal data.

Participatory mapping outcomes were analyzed using two analytical approaches: simulations and regression. The participatory mapping data were summed across maps such that each grid cell was assigned a total number of selections by the respondents for that map. The grid cells were clipped to fall within 20m and 150m isobaths, which were determined by the Greater Amberjack Research Program team to enclose a reasonable biological sampling area within which Greater Amberjack are most likely to reside. The average number of clicks per respondent was calculated for each treatment and map within this clipped area, and this value informed a random draw from a binominal distribution to simulate the number of cells clicked by each respondent. This set of values was applied to each map through Monte Carlo simulations to calculate the probabilities of a single arbitrary cell getting clicked by that respondent sample under the null hypothesis of random cell choice. These probabilities were used to calculate 80%, 90%, and 95% quantiles of the distribution under the null hypothesis, and the true number of clicks was evaluated across these quantiles. For regression analyses, the proportion of clicks by a respondent was calculated by dividing the number of cells a respondent clicked by the total number of cells on a map. The interacting influence of treatment and map were regressed over this standardized proportion of clicks using a beta regression (betareg() function in R following Cribari-Neto and Zeileis 2010). The contribution of terms to model fit was evaluated using likelihood ratio tests and the interaction term was subsequently removed (lmtest package in R, Zeileis et al. 2002). Estimated marginal means were generated post hoc to facilitate pairwise comparisons (emmeans package in R, Lenth et al. 2024).

Results

Panel sample fishing characteristics

On average, respondents had 14.3 years of fishing experience and 8.6 years of reef fishing experience. In the context of all saltwater fishing, offshore fishing for reef fishes was considered extremely important by 20.2%, very important by 25.1%, moderately important by 30.1%, slightly important by 12.1%, and not at all important by 12.5%. Among all species in the reef fish complex for the South Atlantic and Gulf of Mexico, Red Snapper was most frequently chosen as important (50.5% of respondents), and Greater Amberjack was chosen by 17.5% of respondents. Red Snapper was also most frequently chosen to be the single most important offshore reef fish species (35.7% of respondents), where Greater Amberjack was the single most important species for 3.0% of respondents (N = 48).

Greater Amberjack knowledge

61.4% (N = 1,117) of respondents were coded as Greater Amberjack fishermen and 38.6% (N = 702) of respondents were coded as other fishermen. Among those coded as Greater Amberjack fishermen, perceptions of trends in Greater Amberjack abundance were mixed; respondents most frequently indicated that they were not sure if abundance was changing (30.1% of respondents). Perceptions of change in abundance did not vary significantly across region or a respondent's most fished state. Respondents who perceived that Greater Amberjack abundance was increasing most frequently reported that this change began from 2010-2020 (20.0% of respondents), and respondents who perceived a decrease in abundance most frequently reported that they were not sure when this change started (23.1% of respondents).

Greater Amberjack habitat and fishing effort

Perceptions of the best habitat for Greater Amberjack fishing when condensed into artificial structure, natural structure, and natural unstructured bottom were not different by region, but were different by state (Figure 13.1, $p < 0.001$). Respondents most frequently chose artificial structure as the best habitat in Alabama, Louisiana, Mississippi, South Carolina, and Texas. Natural structure was most frequently chosen for the Atlantic coast of Florida, Georgia, and North Carolina. Artificial and natural structure were chosen equally as frequently for the Gulf coast of Florida. The number of Greater Amberjack respondents expected to catch in one hour of fishing on their favorite fishing habitat did not differ across those habitat categories.

Fishing motivations: Greater Amberjack and Red Snapper

Reasons for fishing for these two species also diverged (Figure 13.2). The following reasons for fishing were associated significantly more often with Greater Amberjack than Red Snapper: "Puts up a good fight" ($p = 0.039$) and "Catch-and-release" ($p = 0.006$). Different reasons for fishing were associated significantly more often with Red Snapper than Greater Amberjack: "Good tasting" ($p < 0.001$) and "Taking home a big keeper" ($p = 0.004$). Two reasons for fishing were not significantly associated with either fish species: "Fills a gap between other fishing seasons" ($p = 0.054$) and "Tournament fish/state record" ($p = 0.441$).

Participatory mapping

When evaluated across 80%, 90%, and 95% quantiles, reported fishing effort achieved variable levels of significance (Figs. 14.3 and 14.4). Both the map to which respondents were exposed as well as their treatment group were significant predictors of the proportion of cells clicked in the participatory mapping exercise (Table 13.1). Examination of pairwise comparisons between maps demonstrated that the response for the central GOM (containing Louisiana, Mississippi, and Alabama) was responsible for the majority of differences observed; in other words, the central GOM response was significantly different from four out of the five remaining maps (Table 13.2).

Discussion

This study represented an initial characterization of the human dimensions of an understudied fishery in two economically and culturally vibrant fishing regions of the United States. Greater Amberjack fishermen held mixed perceptions of biological trends and habitat preferences that mirror the lack of technical knowledge that has been developed for this species. Perceptions of trends in Greater Amberjack size abundance and when changes may have occurred in this trend did not converge. We posit that mixed perceptions of change were a result of the 9 years of reef fishing experience reported by these fishermen, over which time the stock statuses of Greater Amberjack have not significantly changed according to fisheries-dependent assessments. In addition, habitat preferences for Greater Amberjack fishing were significantly different across states, reflecting the diversity of habitat types in the study area. Unexpectedly, reported catch-per-unit-effort did not differ across preferred Greater Amberjack habitat, which contradicts existing reports of Greater Amberjack preferring to school near structure.

We observed species-specific differences in perceptions among reef fishermen. Fishing motivations for Red Snapper reflected a greater focus on catch, whereas motivations for Greater Amberjack were more associated with the overall fishing experience rather than the catch itself. Interestingly, the fishing motivation associated with seasonal gaps did not differ across the two species, despite evidence that Greater Amberjack catch can increase up to sevenfold during the closed Red Snapper season on offshore charter boats (Garner and Patterson

2015). These significant differences in motivations across species underscore the need for caution when generalizing recreational fishermen perceptions across reef complexes. Importantly, these findings suggest that ecosystem-based management approaches must account for species-specific nuances, as fishermen's perceptions can vary dramatically, even within established management units.

We developed a novel and statistically rigorous participatory mapping experiment that evaluated spatially-explicit hot and cold spots for reported fishing effort for Greater Amberjack. Ground truthing these data against field-based abundance surveys in the Greater Amberjack Research Program would provide an uncommon opportunity to synergize social and ecological fisheries data sources. The participatory mapping exercise also demonstrated the importance of both geography and treatment for reported fishing effort. For geography, the central Gulf of Mexico map (containing Alabama, Louisiana, and Mississippi) had a significantly larger proportion of cells clicked in comparison to most other maps, so this area may receive a disproportionate amount of Greater Amberjack fishing effort. For treatment, when exposed to the same maps, respondents in the tagging treatment clicked a larger proportion of available cells than respondents in the control treatment. We postulate that this difference in reported fishing effort is derived from perceived benefits from involvement in tagging collaborations with scientists. These benefits can be manifested through multiple pathways including monetarily (i.e. rewards from catching tagged fish) as well as with management implications (i.e. increased engagement and buy-in). The method we developed not only advances knowledge co-production techniques within the participatory mapping domain but also supports the growing pursuit of co-management opportunities in these regions, which remain underutilized in fisheries management processes.

Tables

Table 13.1. Beta regression estimates with significance levels for the proportion of cells clicked by respondents in the participatory mapping exercise.

	Estimate	Std. Error	P-value
Treatment			
Mapping Treatment	0		
Control Treatment	-0.1032	0.0449	0.0216*
Map			
GOM-W	0		
GOM-C	-0.4366	0.0760	< 0.001***
GOM-E	-0.3266	0.0814	< 0.001***
SA-S	-0.0945	0.0858	0.2710
SA-C	-0.1192	0.0791	0.1317
SA-N	-0.1027	0.0938	0.2738

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table 13.2. Estimated marginal means for the map variable from the participatory mapping beta regression.

Contrast	Estimate	Std. Error	P-value
GOM-W – GOM-C	0.0151	0.0029	< 0.001***
GOM-W – GOM-E	0.0119	0.0031	0.0015**
GOM-W – SA-S	0.0038	0.0035	0.8822
GOM-W – SA-C	0.0048	0.0032	0.6720
GOM-W – SA-N	0.0041	0.0038	0.8815
GOM-C – GOM-E	-0.0033	0.0021	0.6263
GOM-C – SA-S	-0.0113	0.0026	0.0003**
GOM-C – SA-C	-0.0104	0.0023	0.0001**
GOM-C – SA-N	-0.0110	0.0030	0.0033**
GOM-E – SA-S	-0.0081	0.0029	0.0539
GOM-E – SA-C	-0.0071	0.0025	0.0534
GOM-E – SA-N	-0.0078	0.0032	0.1466
SA-S – SA-C	0.0009	0.0030	0.9996
SA-S – SA-N	0.0003	0.0036	1.0000
SA-C – SA-N	-0.0006	0.0033	1.0000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Figures

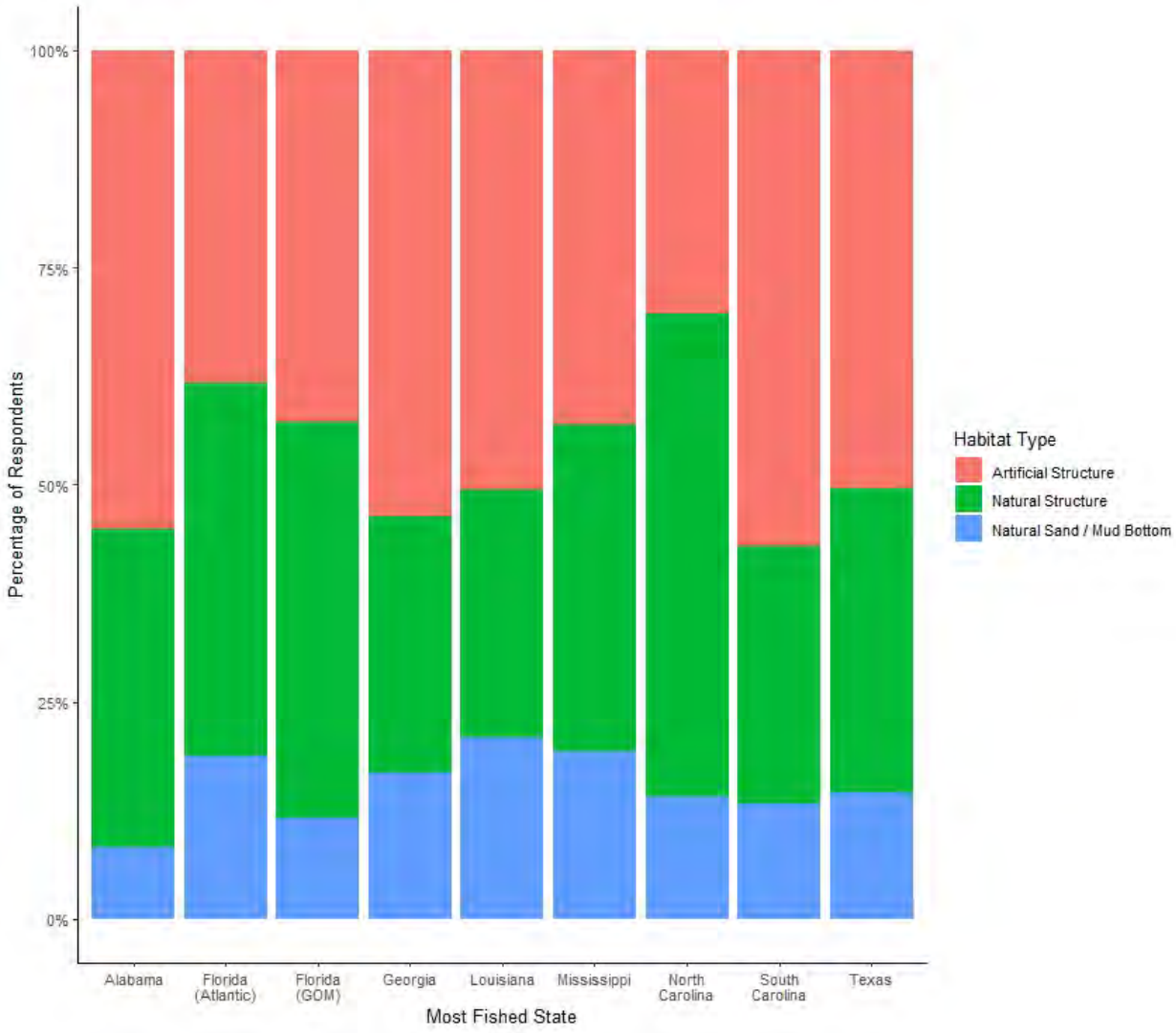


Figure 13.1. Habitat preference for Greater Amberjack fishing across artificial structure, natural structure (e.g. live patches, reefs), or natural sand / mud bottom in a respondent's most fished state. These habitat categories were coded such that artificial structure included artificial reefs (pyramids, chicken coops), bridge rubble, offshore oil/gas rigs, shipwrecks, and sunken rigs.

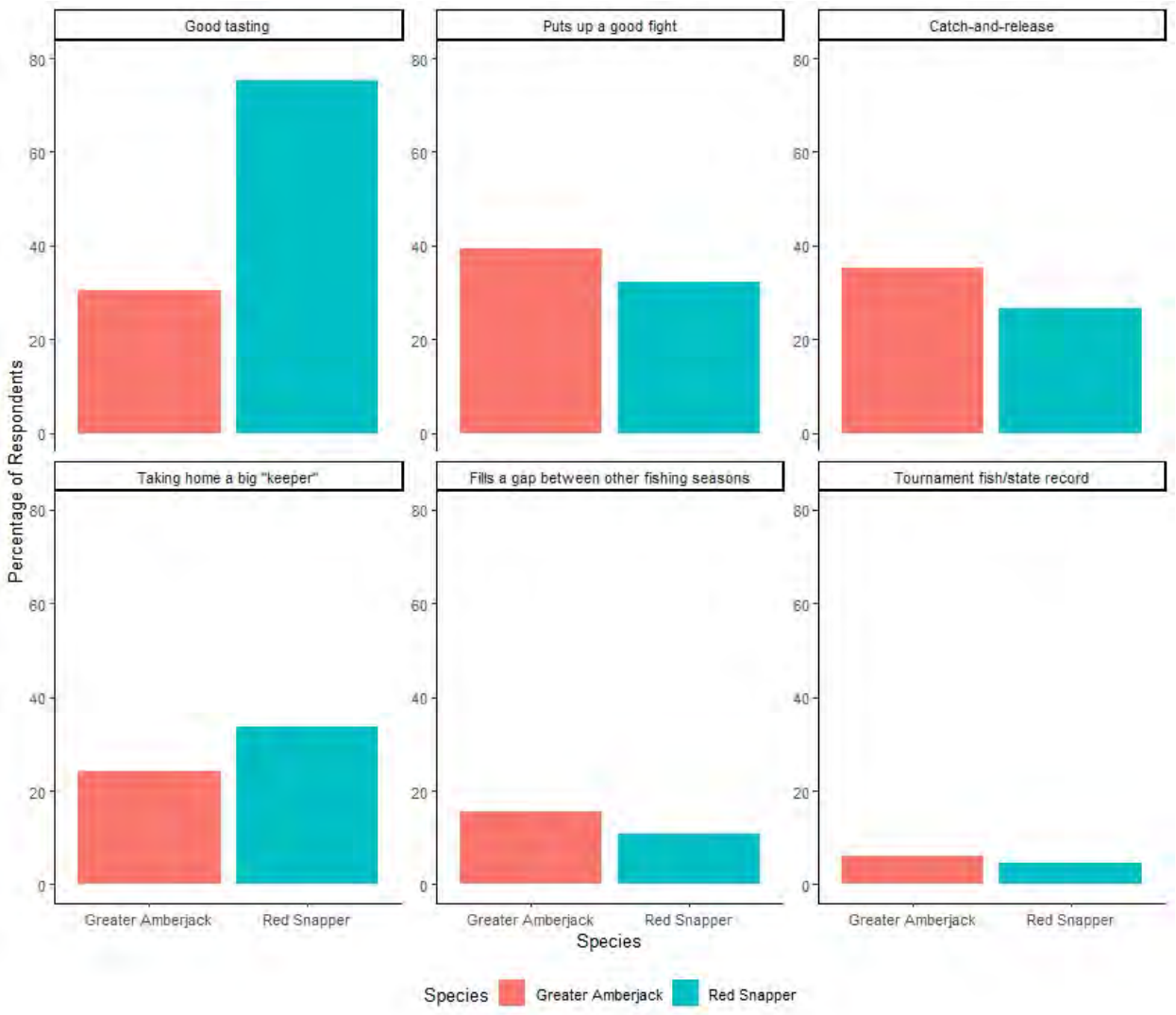
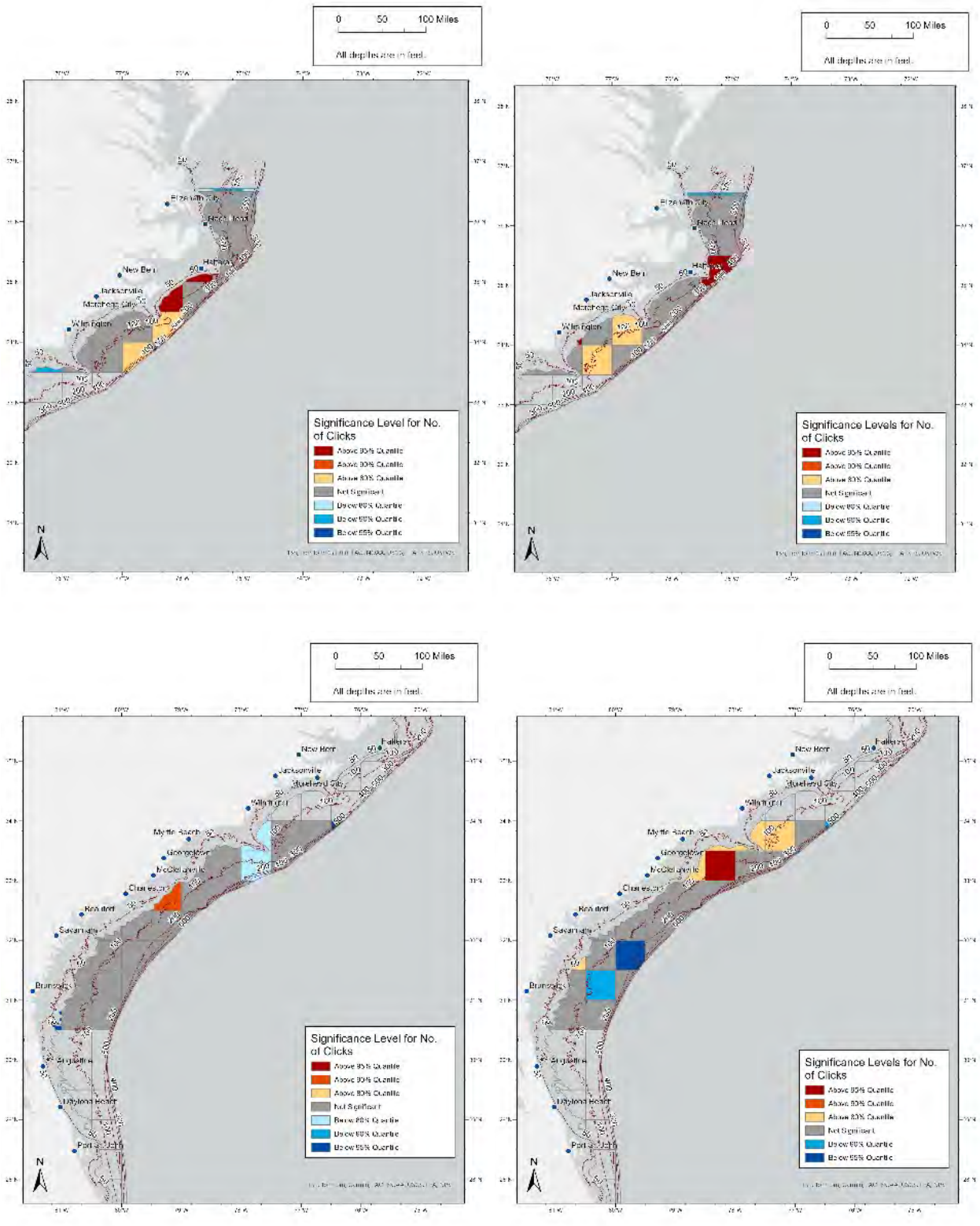


Figure 13.2. Differences in the percentage of respondents who indicated reasons for their fishing across Greater Amberjack and Red Snapper species.


0 50 100 Miles
All depths are in feet.
Significance Level for No. of Clicks
Above 95% Quantile
Above 90% Quantile
Above 80% Quantile
Not Significant
Below 80% Quantile
Below 90% Quantile
Below 95% Quantile
Significance Levels for No. of Clicks
N

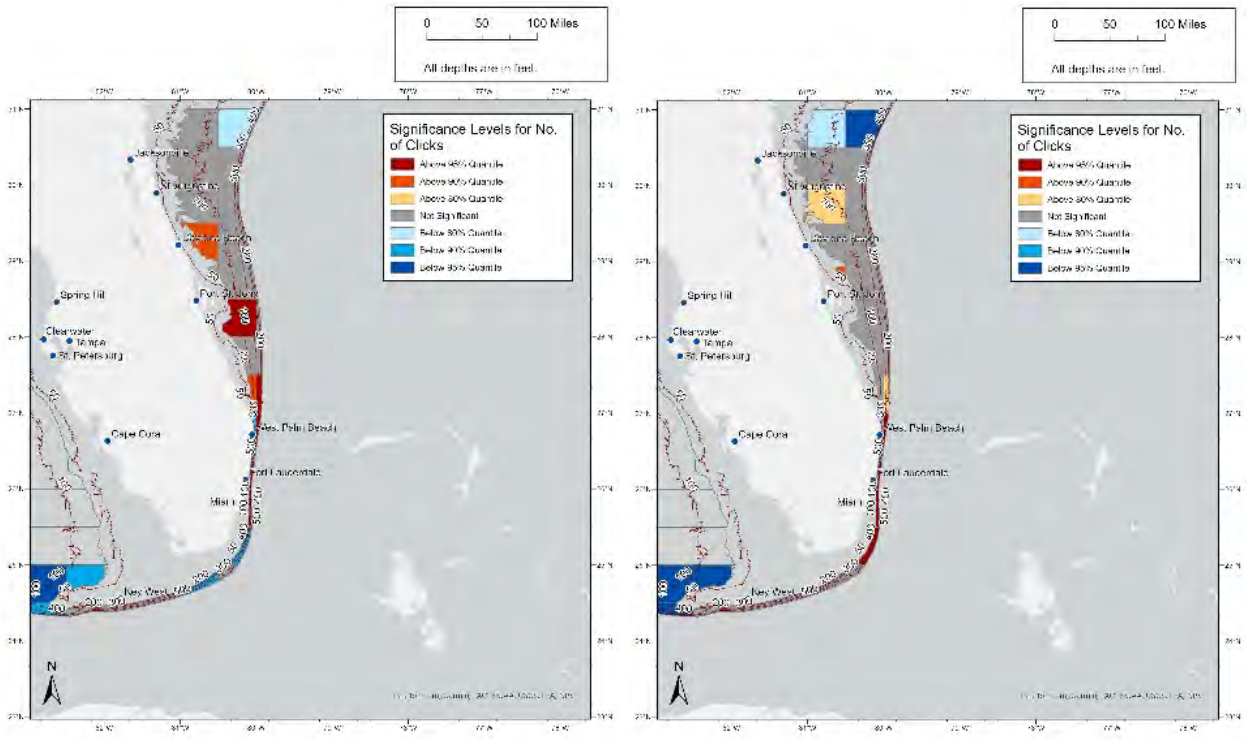
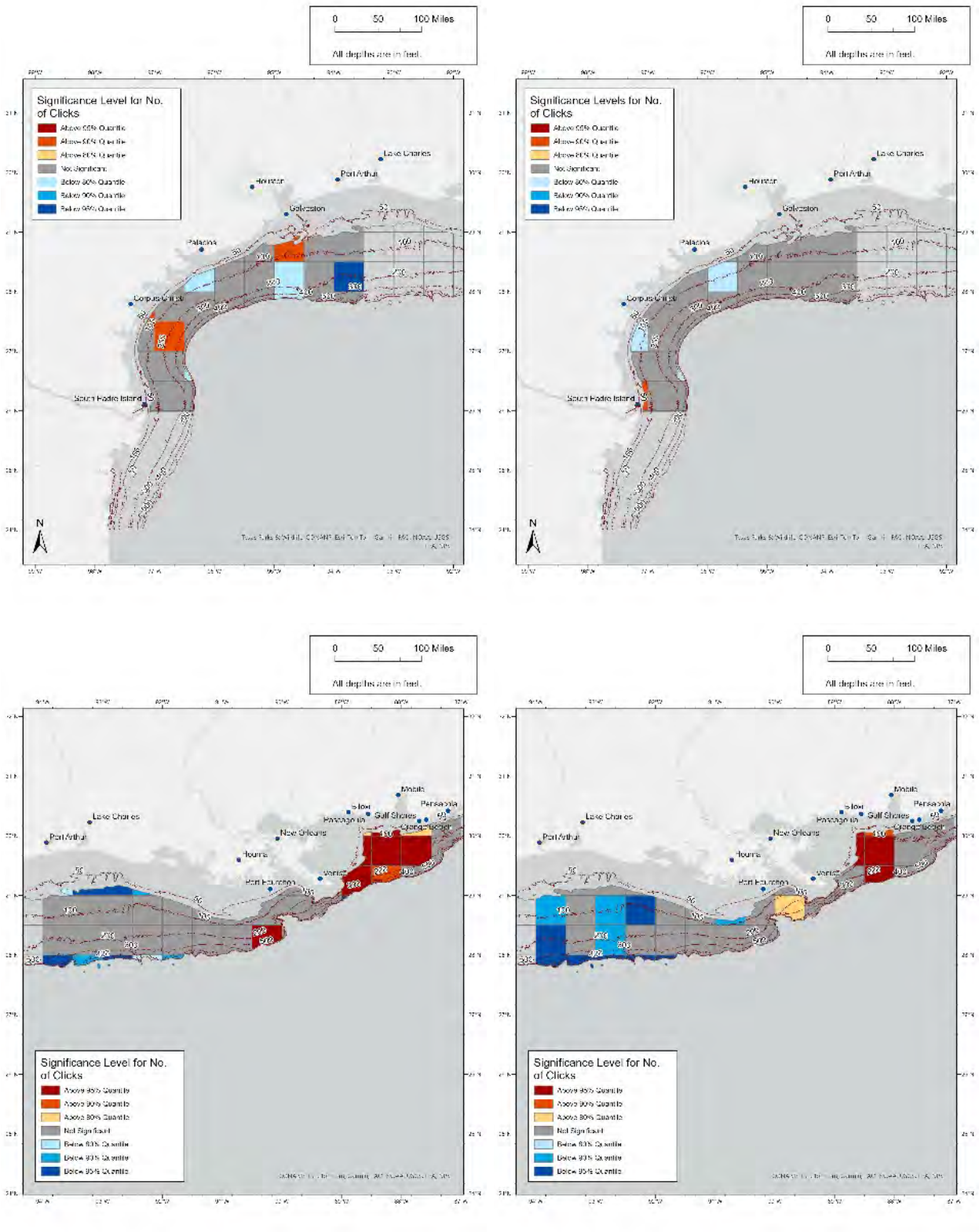


Figure 13.3. Participatory mapping results for the South Atlantic region. The first set of maps contain North Carolina, the second set of maps contain South Carolina and Georgia, and the third set of maps contain the Atlantic coast of Florida. Maps on the left represent results from the tagging treatment, and maps on the right represent maps from the control treatment. “Hot” and “cold” spots are symbolized to represent cells that fall within or outside of quantiles produced from data simulations.


0 50 100 Miles
All depths are in feet.
Significance Level for No. of Clicks
Above 95% Quantile
Above 90% Quantile
Above 80% Quantile
Not Significant
Below 80% Quantile
Below 90% Quantile
Below 95% Quantile
Lake Charles
Houston
Port Arthur
Galveston
Palacios
Corpus Christi
South Padre Island
N
0 50 100 Miles
All depths are in feet.
Significance Levels for No. of Clicks
Above 95% Quantile
Above 90% Quantile
Above 80% Quantile
Not Significant
Below 80% Quantile
Below 90% Quantile
Below 95% Quantile
Lake Charles
Houston
Port Arthur
Galveston
Palacios
Corpus Christi
South Padre Island
N
0 50 100 Miles
All depths are in feet.
Mobile
Pensacola
Lake Charles
Port Arthur
New Orleans
Houma
Gulf Shores
Orange Beach
Significance Level for No. of Clicks
Above 95% Quantile
Above 90% Quantile
Above 80% Quantile
Not Significant
Below 80% Quantile
Below 90% Quantile
Below 95% Quantile
0 50 100 Miles
All depths are in feet.
Mobile
Pensacola
Lake Charles
Port Arthur
New Orleans
Houma
Gulf Shores
Orange Beach
Significance Level for No. of Clicks
Above 95% Quantile
Above 90% Quantile
Above 80% Quantile
Not Significant
Below 80% Quantile
Below 90% Quantile
Below 95% Quantile

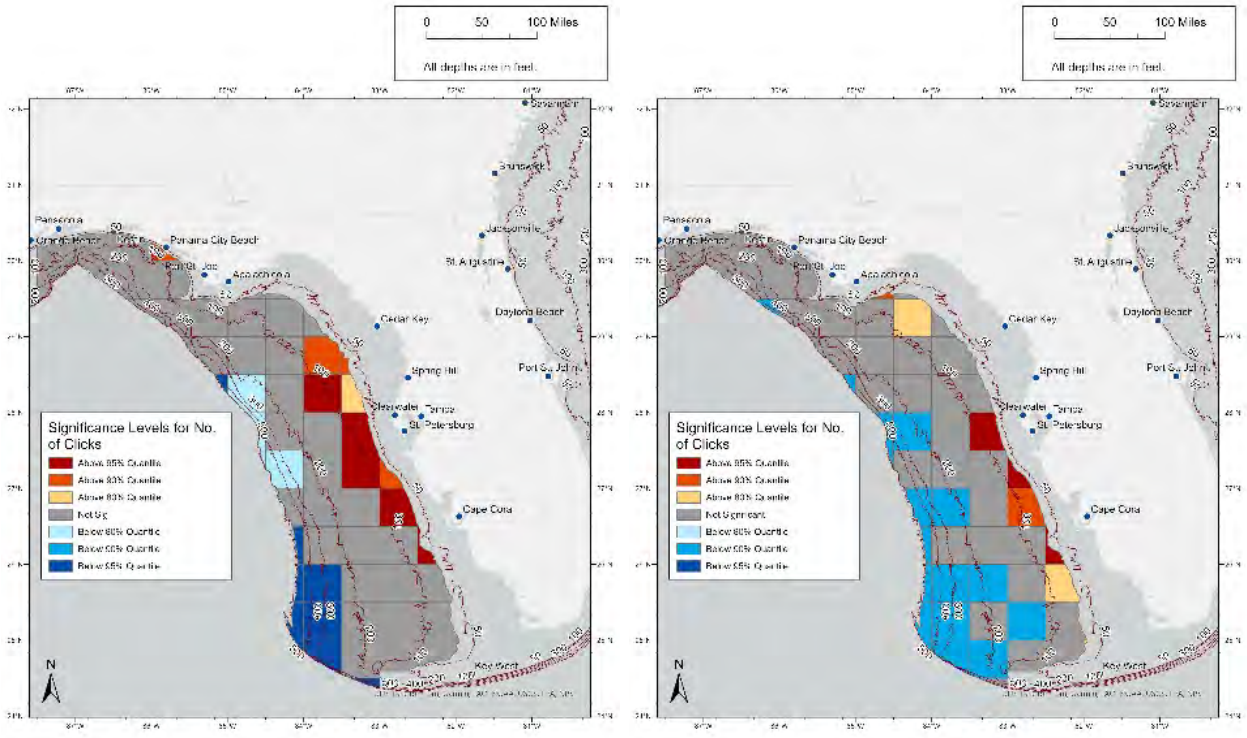


Figure 13.4. Participatory mapping results for the Gulf of Mexico region. The first set of maps contain Texas, the second set of maps contain Louisiana, Mississippi, and Alabama, and the third set of maps contain the Gulf coast of Florida. Maps on the left represent results from the tagging treatment, and maps on the right represent maps from the control treatment. “Hot” and “cold” spots are symbolized to represent cells that fall within or outside of quantiles produced from data simulations.

References

Cribari-Neto, F. and Zeileis, A. 2010. Beta Regression in R. Journal of Statistical Software, 34: 1–24. doi:10.18637/jss.v034.i02

Environmental Systems Research Institute (ESRI). 2024. ArcGIS Pro: Release 3.0.0. Environmental Systems Research Institute, Inc., Redlands, CA.

Garner, S. B. and Patterson, W. F. 2015. Direct observations of fishing effort, catch, and discard rates of charter boats targeting reef fishes in the northern Gulf of Mexico. Fishery Bulletin 113: 157-166. https://doi.org/10.7755/FB.113.2.4

IBM Corp. 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., and Singmann, H. 2024. Estimated marginal means, aka least-squares means. https://CRAN.R-project.org/package=emmeans

Lorenzen, K., Camp, E., Hazell, J., Crandall, C., Montes, N., Hervas, S., Lehmensiek, M., and Chong, L. 2021. 2020-2023 Gulf of Mexico and South Atlantic Greater Amberjack Research Program Visioning Component Report. https://www.flseagrant.org/wp-content/uploads/2023/05/GAJ_Visioning-Report.pdf

R Core Team. 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Zack, E. S., Kennedy, J., and Long, J. S. 2019. Can nonprobability samples be used for social science research? A cautionary tale. Survey Research Methods 13: 215–227.

Zeileis, A. and Hothorn, T. 2002. Diagnostic Checking in Regression Relationships. R News 2: 7–10. https://CRAN.R-project.org/doc/Rnews/

Zhang, C., and Conrad, F. 2014. Speeding in web surveys: the tendency to answer very fast and its association with straight-lining. Survey Research Methods 8: 127–135.

14. Appendices

14.1. Appendix A: Post-release Mortality

Acoustic telemetry-based estimates of Greater Amberjack Post Release Mortality

Post-release mortality (PRM) was calculated using data from Greater Amberjack equipped with acoustic tags in North Carolina, Alabama, and Louisiana. In both North Carolina and Alabama, tags were deployed at sites monitored by acoustic positioning systems (APS) that allowed for precise measurement (<10 m) of Amberjack space use. In Louisiana and Alabama, fish were tagged using specialized transmitters containing a pressure sensor that allows for measurement of depth use by tagged individuals. From these data, we were able to assign fates to fish post-release and use these fates to estimate PRM. In regions with APS individual fates were assigned based on space use within the array (North Carolina, n = 35; Alabama, n = 72), which was supplemented by depth use data from individuals equipped with pressure sensing tags (Alabama, n = 28). In Louisiana, individual fates were assigned based on depth use patterns post release (n = 38).

Based on detection data, five possible fates were established based on condition post-release. Fates assigned were as follows: Alive – remained in acoustic arrays and demonstrated swimming patterns consistent with amberjack (i.e. moving throughout the array, variable depth use), or emigrated from the array but was later detected at other sites, Predation – sudden erratic behavior, broad swimming patterns that do not associate with structure or move frequently between habitats, Tag Loss – acoustic positions indicate the tag is stationary within the array, depth data indicate the tag is lying motionless on the seafloor, Emigrant – the fish quickly leaves the site post-release and is never again detected, Unknown – a tagged fish that is never detected either at the release array or on any other receivers.

In the present context, we assume that tag loss cases are either the result of post release predation or mortality, as tags were surgically implanted with the wound sutured shut, making immediate tag shedding unlikely. In addition, all cases of Emigrant in Louisiana were characterized by immediate departure from the site with only minutes to a few hours of data for these fish. None of these emigrant fish were detected following this behavior, suggesting they were either chased off site by a predator and consumed or perished and drifted out of our array.

For each state we calculated a low and a high estimate of PRM based on the percentage of fish assigned a mortality fate. For the low estimate, we included only Tag Loss and Predation fates in the calculation, whereas for the high estimate we also included Emigrant fates. Unknown fates were excluded from both calculations. Fate classifications and estimates of PRM are presented in Table 14.1.1. See figures for examples detection data used for fate classifications.

Table 14.1.1: Fate classification for Greater Amberjack included in post release mortality estimates as well as low and high estimates of post release mortality by state.

State	*Alive*	*Predation*	*Tag Loss*	*Emigrant*	*Unknown*	*PRM Low*	*PRM High*
North Carolina	26	1	3	5	0	11.43%	25.71%
Alabama	59	0	12	1	6	16.67%	18.06%
Louisiana	26	1	0	7	4	2.94%	23.53%
Overall	**111**	**2**	**15**	**13**	**10**	**12.06%**	**21.28%**

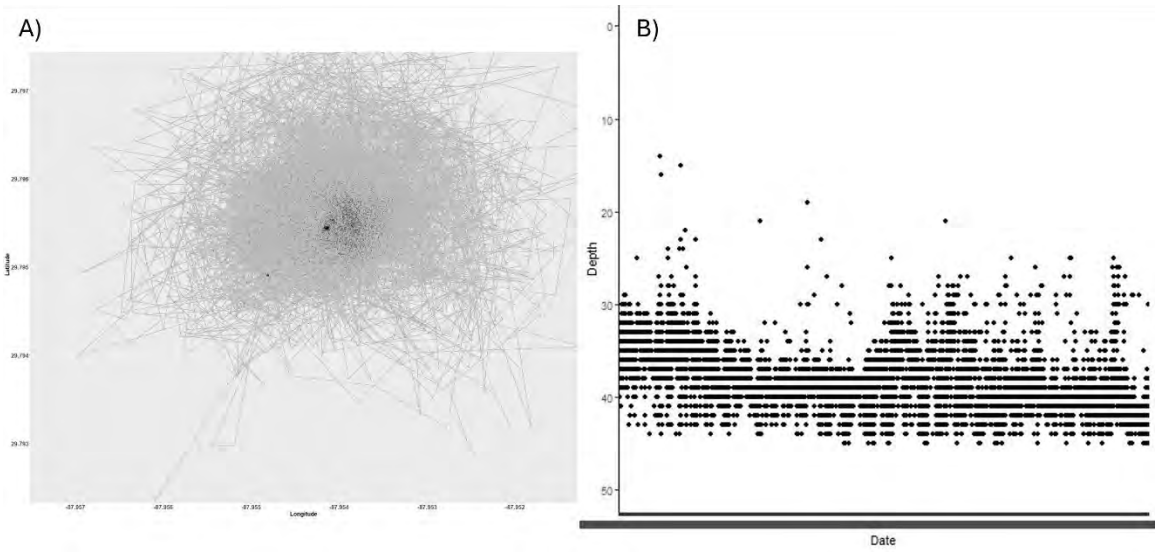


Figure 14.1.1: A) APS positional data for a Greater Amberjack tagged in Alabama that was classified as Alive. B) Depth use data from the same individual.

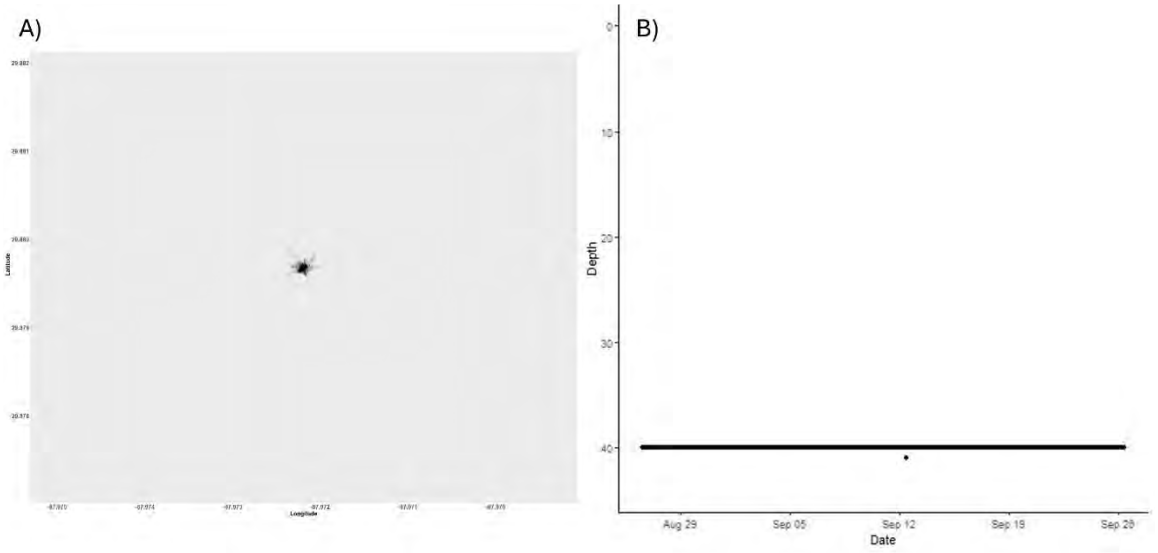


Figure 14.1.2: A) APS positional data for a Greater Amberjack tagged in Alabama that was classified as a Tag Loss. B) Depth use data from the same individual.

14.2. Appendix B: Diet

Life Stage-Specific Diet Composition of Greater Amberjack (*Seriola dumerili*)

Mischa K. Schultz[1], Michael A. Dance[1], Patrick Graham[1], Aryan Ravari[1], Creed Branham[1], Brett Falterman[2], and Stephen R. Midway[1],*

[1]Department of Oceanography and Coastal Science, Louisiana State University, Baton Rouge, LA. Mschu31@lsu.edu

[2]Fisheries Research Support LLC, Mandeville, LA 70448

INTRODUCTION

Knowledge of feeding habits is crucial for understanding how organisms acquire energy, directly influencing their survival and the structure of ecosystems they inhabit. Understanding the feeding ecology of an organism helps identify predator-prey interactions and elucidate food web structure. Stomach content analysis (SCA) can capture short-term dietary information (Baker et al., 2024; Dance et al., 2018; Wells et al., 2008) and when used in conjunction with genomics can provide a more thorough diet description (Berry et al., 2015; Harms-Tuohy et al., 2016; Leray et al., 2015). Several studies have successfully applied this complementary approach to characterize the diets of predatory marine fishes.

Greater amberjack (*Seriola dumerili*) are predatory reef fishes that undergo an ontogenetic habitat shift from sargassum-associated nursery habitat to a benthic adult stage. Across their range, individuals can occupy both pelagic and benthic habitats, suggesting an opportunistic, generalist foraging strategy. Life stage- and habitat-specific changes in feeding ecology would have meaningful implications for ecosystem-based fisheries modeling and food web dynamics. Despite this, published information on greater amberjack trophic ecology in the Gulf of America remains limited, hindering our ability to accurately characterize reef food web dynamics across the region.

Here, we use complementary analyses, SCA and metabarcoding analysis, to investigate the size-based differences in the feeding ecology of the greater amberjack in the northern Gulf of America (hereafter Gulf), providing the most taxonomically precise prey classification of their diet to date.

METHODS

Data collection

Greater amberjack (> 863 mm FL) in the northern Gulf of America were sampled from recreational charter landings via dockside from August 2022 – 2024. Subadult (< 863 mm FL) greater amberjack were collected off reef habitat, sargassum mats and other flotsam in the summer months during fishery independent surveys. Greater amberjack were measured to the nearest 1.0 mm (FL) and sexed (when possible) before removing the stomach. In addition, weight (kg), habitat type, water depth, distance from shore, and bait type were recorded for each specimen. Samples were immediately transported on ice to Louisiana State University, where stomachs were stored at -20°C.

Visual assessment of stomach contents (SCA)

In the lab, frozen stomachs were thawed, and the contents were removed for identification and numeration. Each prey item found in the greater amberjack stomach was sorted, weighed, and identified to the lowest

possible taxonomic level. Two functional groups were used to describe the main taxonomic groups (Teleostei, Invertebrata), which allows for the inclusion of unidentifiable prey items. Any prey items that were identified as bait were removed from further analyses. Greater amberjack diet was quantified using numerical percentages (%N) and wet weight percentage (%W) of each prey contribution, as well as frequency of occurrence (%FO). An index of relative importance (IRI) was computed that incorporates both numerical and weight-based metrics: IRI= (%N+%W) × %FO and %IRI= (IRI/ ΣIRI) × 100. Stomachs were analysed for a vacuity index (VI) or empty stomach index (cite more vacuity studies) using the following equation:

$$\%\text{VI} = \frac{\text{Total number of empty stomachs noted}}{\text{Total number of stomach analyzed}} \text{ x } 100$$

and was based on the following scales: %VI values ranging from 0 – 20 (voracious), 20 – 40 (slight voraciousness), 40 – 60 (moderate feeding), 60 – 80 (relatively low feeding), and 80 – 100 (low feeding).

Metabarcoding (MA)

[Placeholder for final methods from CD genomics]

Stomach content samples will be subjected to DNA metabarcoding to characterize prey communities at the molecular level. Total DNA will be extracted from homogenized stomach content material and the 18S rRNA gene (zooplankton, fish larvae, gelatinous organisms) will be amplified via PCR. Amplicons will be sequenced on at CD Genomics (Shirley, NY).

Data analysis

Stomach content results were analyzed according to life stage (adult > 500 mm and juveniles < 500 mm). Life stage classes were chosen according to approximate reef settlement stage for greater amberjack (Schultz et al., 2025). The %W of dominant prey groups was used as the dependent variable for statistical analyses of SCA, as it is a useful proxy for estimating the nutritional contribution of prey groups (Bowen 1996). A permutational multivariate analysis of variance (PERMANOVA) was used to assess whether diets (%W) varied significantly (p <0.05) between juveniles and adults. Finally, a similarity percentage (SIMPER) test was used to identify the contribution of individual prey groups to the differences in diet between juveniles and adults.

RESULTS

Diet composition based on SCA

A total of 77 stomachs (35 juveniles, 42 adults) have been analyzed. Vacuity index (VI) was calculated as 35.1% for all samples (n = 27), 22.9% (slight voraciousness; n = 8) for juveniles, and 45.2% (moderate feeding; n = 19) for adults; though this difference was not statistically significant (Fisher's exact test; p = 0.055). Teleosts were the dominant prey group in greater amberjack diet for adults (IRI = 44.3%), whereas crustaceans were the dominate prey group for juveniles (IRI = 43.5%; Table 1, Figure 1). Fish from the family Balistidae were the most dominant teleost prey species for adults (IRI = 51.4%). Diet composition differed significantly between adult and juvenile greater amberjack (PERMANOVA: $F_{1,42}$ = 9.94, R^2 = 0.191, p = 0.001), with life stage explaining approximately 19% of the variation in Bray-Curtis dissimilarity. SIMPER analysis identified four prey categories contributing cumulatively to 74.7% of the dissimilarity between life stages (Teloest, Portunidae, Balistidae, Decapoda). Teleosts were the primary contributor to dissimilarity (25.0%), occurring predominantly in adult diets (mean relative abundance: 0.536 vs. 0.158 in juveniles; p = 0.002). Conversely, Portunidae (17.2% contribution; juveniles: 0.370 vs. adults: 0.000; p = 0.001) and Decapoda (15.1% contribution; juveniles: 0.302 vs.

adults: 0.000; p = 0.009) were exclusive to juvenile diets, as was Balistidae, which occurred only in adult diets (17.3% contribution; adults: 0.367 vs. juveniles: 0.000; p = 0.002). Taxonomic heat trees illustrated the hierarchical structure of greater amberjack diets, where node size, edge width, and color intensity were scaled proportionally to %IRI for each prey taxon (Figure 3).

Metabarcoding results

[Placeholder until results come in from CD genomics]

CONCLUSIONS

Greater amberjack undergo a clear ontogenetic diet shift, with juveniles consuming predominantly crustaceans associated with sargassum (Portunidae, Hippolytidae), while adults shifted towards piscivory. Diet composition differed significantly between adult and juvenile greater amberjack, indicating distinct dietary patterns between life stages. Adults exhibited higher stomach vacuity (45%) compared to juveniles (23%), suggesting reduced feeding frequency or longer digestion periods in larger individuals. High stomach vacuity and highly digested prey in adults led to low resolution in the adult diet profile. Genomics results will be necessary for characterizing finer resolution diet of greater amberjack. These results establish foundational trophic ecology data for greater amberjack in the northern Gulf of Mexico.

TABLES AND FIGURES

Table 1: Stomach contents from greater amberjack (*Seriola dumerili*) caught off the coast of Louisiana, USA reported to the lowest identified taxa as %IRI.

Prey		*Juvenile (n = 35)*	*Adult (n = 42)*
Teleostei	Balistidae	–	51.4
	Carangidae	–	2.98
	Elops saurus	–	1.13
	Unidentified Teleost	9.92	44.30
Malacostraca	Decapoda	28.50	–
	Brachyura	3.50	–
	Hippolytidae	11.4	–
	Portunidae	43.5	–
	Portunus sayi	0.15	–
	Stomatopoda	–	0.07
	Phronimidae	0.13	–
	Phronsinidae	0.13	–
Other arthropods	Insecta	0.04	–
	Zygoptera	0.13	–
Porifera	Unidentified sponge	–	0.07
Anthropogenic	Plastic	1.61	0.08

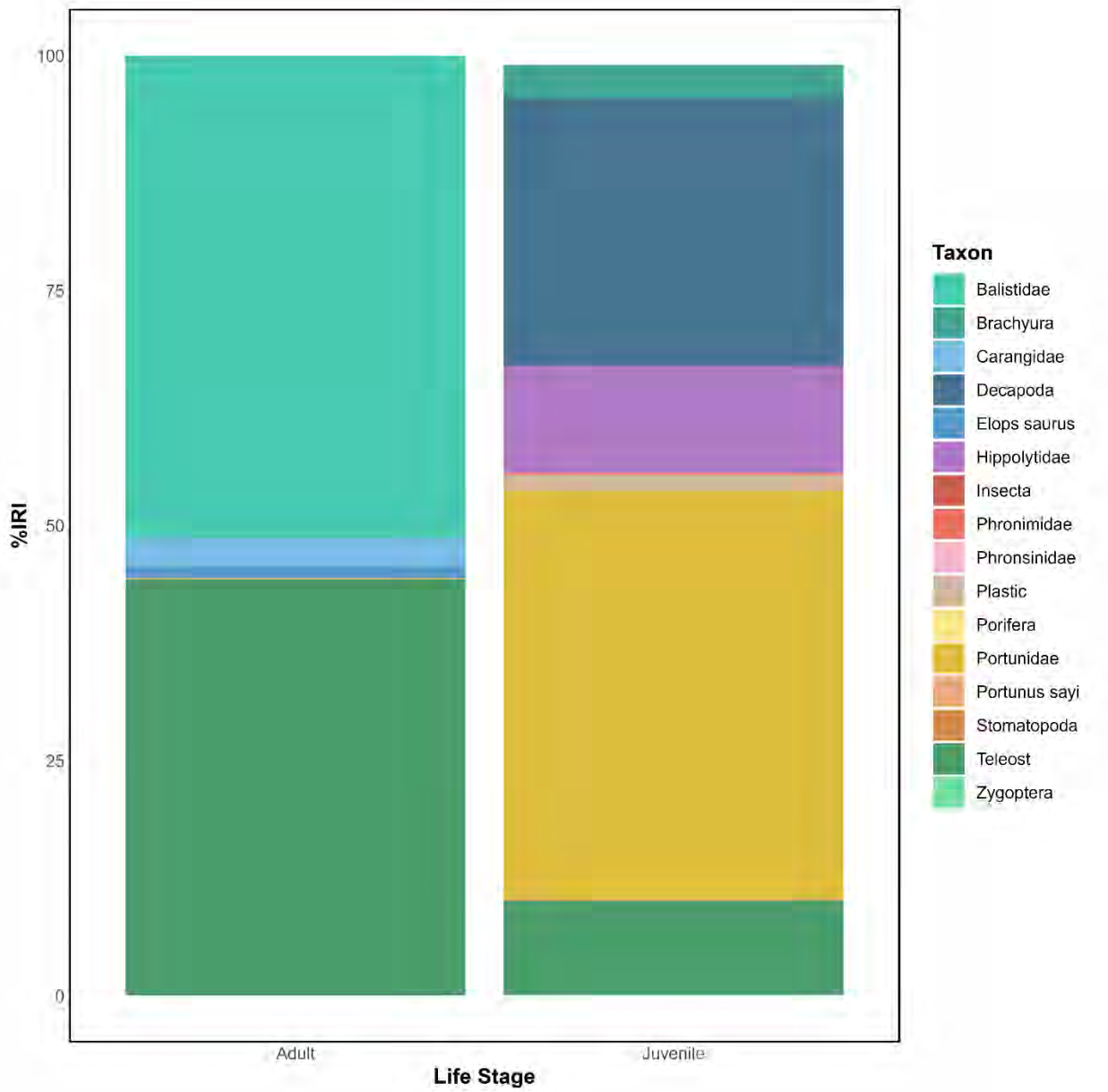


Figure 1: % IRI contributions of each major prey group to both adult and juvenile greater amberjack (*Seriola dumerili*) from Louisiana, USA.

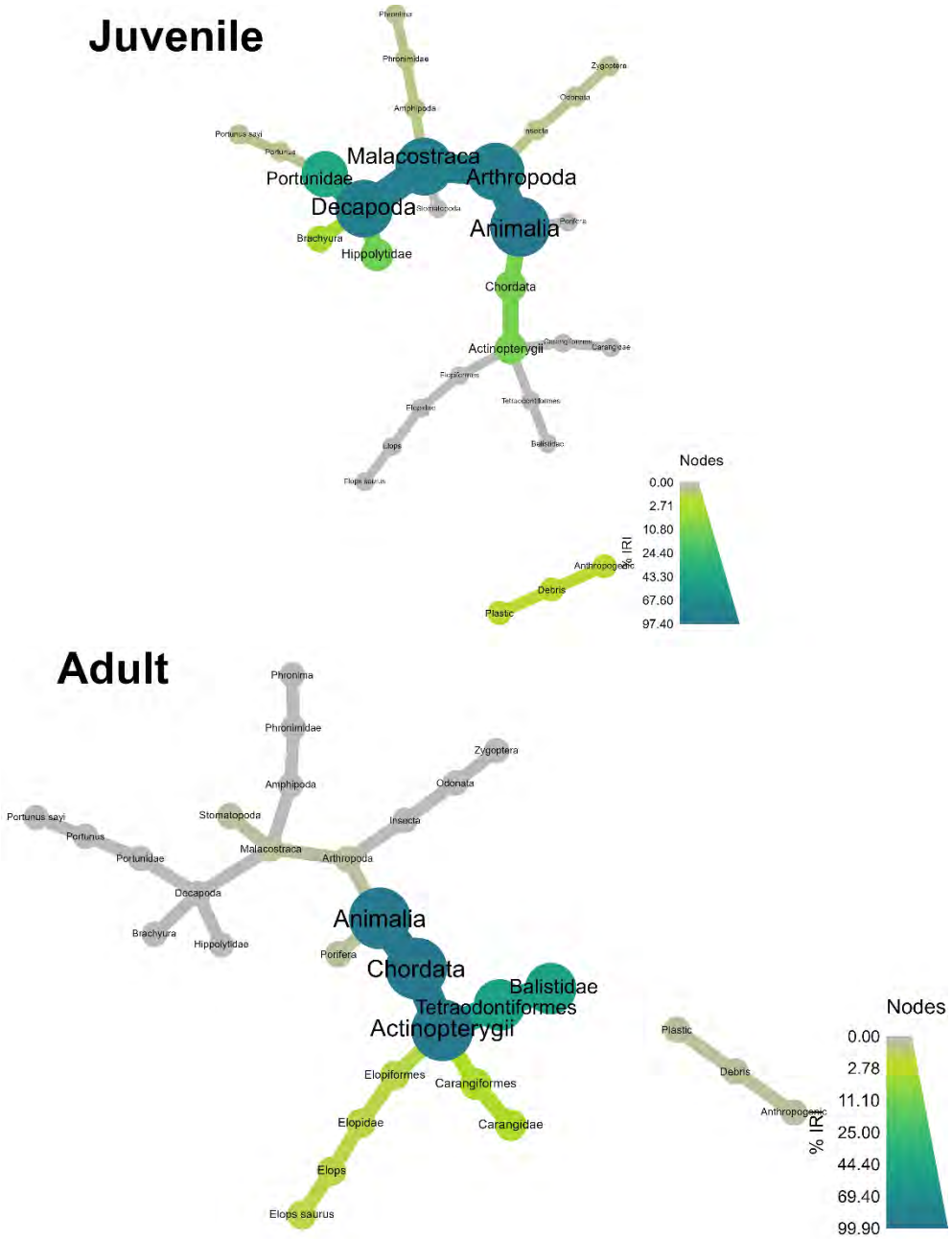


Figure 2: A heat tree showing the importance of prey items (identified to the lowest possible taxonomy) to greater amberjack (*Seriola dumerili*) diets based on %IRI. The central node of each heat tree represents the total diet (Animalia) where color, node size, and text size are scaled to prey item importance.

REFERENCES

Baker, R., Crawford, H. M., & Sheaves, M. (2024). Stomach content analysis. In *Quantifying Diets of Wildlife and Fish: Practical and Applied Methods* (pp. 0). CSIRO Publishing. https://doi.org/10.1071/9781486315017.BK08027_CH3

Berry, O., Bulman, C., Bunce, M., Coghlan, M., Murray, D. C., & Ward, R. D. (2015). Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. *Marine Ecology Progress Series*, *540*, 167-181. https://www.int-res.com/abstracts/meps/v540/meps11524

Dance, K. M., Rooker, J. R., Shipley, J. B., Dance, M. A., & Wells, R. J. D. (2018). Feeding ecology of fishes associated with artificial reefs in the northwest Gulf of Mexico. *PLoS ONE*, *13*(10), e0203873. https://doi.org/10.1371/journal.pone.0203873

Harms-Tuohy, C. A., Schizas, N. V., & Appeldoorn, R. S. (2016). Use of DNA metabarcoding for stomach content analysis in the invasive lionfish Pterois volitans in Puerto Rico. *Marine Ecology Progress Series*, *558*, 181-191. https://www.int-res.com/abstracts/meps/v558/meps11738

Leray, M., Meyer, C. P., & Mills, S. C. (2015). Metabarcoding dietary analysis of coral dwelling predatory fish demonstrates the minor contribution of coral mutualists to their highly partitioned, generalist diet. *PeerJ*, *3*, e1047. https://doi.org/10.7717/peerj.1047

Schultz, M. K., Dance, M. A., Powers, S., & Midway, S. R. (2025). Global Review of Greater Amberjack (*Seriola dumerili*) Life History. *Reviews in Fisheries Science & Aquaculture*, 1-22. https://doi.org/10.1080/23308249.2025.2595023

Wells, R. J. D., Cowan Jr, J. H., & Fry, B. (2008). Feeding ecology of red snapper Lutjanus campechanus in the northern Gulf of Mexico. *Marine Ecology Progress Series*, *361*, 213-225.

14.3. Appendix C: Otolith microchemistry

Population structure and regional connectivity of greater amberjack (*Seriola dumerili*) in the U.S. Gulf

Mischa K. Schultz[1], Jeffrey D. Plumlee[2], Emily F. Hura[2], Achim D. Herrmann[3], Michael A. Dance[1], and Stephen R. Midway[1],*

[1]Department of Oceanography and Coastal Science, Louisiana State University, Baton Rouge, LA.

[2]School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, Louisiana, USA

3Department of Geology and Geophysics, Louisiana State University, Baton Rouge, LA 70803, USA

INTRODUCTION

The concept of what constitutes a population is an essential discipline of fisheries science, yet general population structure scenarios have been debated. In fisheries management, the term stock is often used as the operational unit of a population — defined by the geographic or jurisdictional boundaries within which harvest is regulated (Cadrin, 2020). Most stock assessments assume that a stock is a single discrete population (Hilborn & Walters, 1992), and does not incorporate more complex spatially-structured populations (metapopulations) within that stock. Many marine populations appear to have a metapopulation structure that is either driven by early life dispersal or adult migration (David et al., 2022; James et al., 2002; Kent Smedbol & Wroblewski, 2002; Kritzer & Sale, 2006; Seijo, 2007), and the importance of considering metapopulation structure in the management of marine fisheries is increasingly recognized (Hart & Cadrin, 2004; Heath et al., 2008).

Here, we apply otolith chemistry to investigate the population structure and connectivity of greater amberjack (*Seriola dumerili*) in the Gulf of America (hereafter referred to as Gulf). The aim of this research was to answer the questions (1) Does greater amberjack natal and adult habitat differ within the Gulf? (2) Are greater amberjack captured in the Gulf associated with different natal habitats? And (3) Are adult populations of greater amberjack regionally distinct? A better understanding of greater amberjack population structure and connectivity are needed for effective management of the species in the Gulf.

METHODS

Sample collection and preparation

Otoliths were collected by NOAA Fisheries observers and independent sampling by Louisiana State University in 2020 and 2022 in U.S. waters of the Gulf. Archived otoliths from greater amberjack were provided by the National Oceanic and Atmospheric Administration (NOAA) Fisheries' Southeast Regional Office in Panama City, Florida. Otoliths were sampled from greater amberjack ranging from 835–1370 mm fork length (FL). Samples were organized into three regions: northwest Gulf (Texas and Louisiana; hereafter West), northeast Gulf (Alabama, Mississippi, and along the Florida coast until Tampa Bay; hereafter East) and southeast Gulf (Tampa Bay, FL to the Florida Keys; hereafter Southeast). Otoliths were set in Struers epoxy resin and sectioned along a transverse plane. Otolith sections were mounted in a mixture of Epofix resin in stubs (Figure 1) and polished using a series of 250-, 600-, and 800-grit sandpapers until the otolith primordium (core) was exposed and growth increments were visible. Ages for each otolith were determined by enumerating growth increments (Schultz et al, *Chap. 8*).

Trace elements

Trace element composition of greater amberjack otoliths were assessed using a laser ablation inductively coupled plasma mass spectrometer (LA-ICP-MS). A series of 70 µm diameter spots from the core to the margin were ablated on each otolith, with spacing of 150 µm between centroids of successive spots (Sanchez et al., 2020). Organization of ablation spots targeted 3 distinct regions of the otolith: core (young-of-year period) and edge (recent growth period). Following Sanchez et al. (2020), the core consisted of 5 spots arranged in an M-shape contained within the core at with a central point at the base of the sulcal groove (Figure 1). Edge samples consisted of at least 3 points positioned laterally along the distal edge of the otolith.

National Institute of Standards and Technology (NIST) silicate glass standards were used to calibrate the instrument before each sample, followed by the internal carbonate standards: National Homogeneity Foraminifera Standard (NHFS) and Japanese Carbonate Technical (JCT) standard. To assess potential drift within the LA-ICP-MS, NIST, NHFS, and JCT standards, with ablation setting identical to otolith samples, were ablated after each individual otolith. Spots were ablated at 8% laser output and 18 Hz intensity, and each ablation had a 40 sec washout delay. Trace element concentrations were measured for ^{7}Lithium (Li), ^{24}Magnesium (Mg), ^{55}Manganese (Mn), ^{65}Copper (Cu), ^{85}Rubidium (Rb), ^{88}Strontium (Sr), and ^{137}Barium (Ba).

Trace element data reduction was performed using iolite 4 (Release 4.10.4, standalone version) with the 3D Trace Elements data reduction scheme and multiple reference materials. Calcium (^{43}Ca) was used as the internal standard, with certified concentrations of 38.53 wt% for NFHS, 38.95 wt% for JCT, and 8.50 wt% for NIST 612. NIST SRM 612 served as the primary calibration standard using certified reference values. Baseline intervals were fit using a weighted linear fit, and output channels for NIST 612, JCT, and NFHS were smoothed using a spline autosmooth function. Limits of detection were calculated following the method of Longerich et al. (1996). Secondary reference materials (NFHS and JCT) were analyzed throughout the session to monitor accuracy and precision. Concentrations were in parts per million (ppm) determined by the internal carbonate standard (^{43}Ca).

Data analysis

Before analysis, element values were pooled by region, and all samples within ±5 SD of the mean were considered outliers and removed from any subsequent analysis. Otolith trace element ratios (Li, Ba, Mn, Sr, and Rb) were analyzed to address three questions: (1) whether natal and adult habitats differed chemically within individuals, (2) whether adult habitat chemistry varied by region, and (3) whether natal habitat chemistry varied by region. To evaluate differences between natal (core) and adult (edge) otolith chemistry within individuals, a multivariate analysis of variance (MANOVA) was used to assess the overall multivariate chemical signatures and then paired t-tests were conducted for each element. Multivariate analysis of variance and ANOVA were used to determine whether chemical signatures (multiple markers) or individual markers, respectively, differed among geographic regions, with year incorporated as an interactive factor ($\alpha = 0.05$) (RStudio, package 'nlme'). Tukey's highly significant difference (HSD) tests with a 'Shaffer' corrected alpha ($\alpha = 0.05$) were run on interaction models (Region × Year) to analyze region-specific differences for individual markers with significant ANOVA results. To assess the discriminatory power of otolith chemistry for classifying fish to region of capture, quadratic discriminant function analysis (QDFA) was performed separately on edge and core chemistry using elements that showed significant or near-significant regional effects ($\alpha = 0.10$ for edge; $\alpha = 0.14$ for core) in the univariate ANOVAs. Classification accuracy was evaluated using jackknife (leave-one-out) cross-validation to provide an unbiased estimate of predictive performance. To visualize multivariate differences in otolith elemental chemistry among regions, redundancy analysis (RDA) was performed separately for otolith core and edge.

Response variables included element concentrations for elements identified as influential in discriminant function analysis.

RESULTS

Otoliths from 33 (East = 15, Southeast = 7, West = 11) greater amberjack (FL = 948.0 ± 36.2 mm, mean ± SD) that fit the selection criteria were available for geochemical analysis. Five elements (Li, Mn, Sr, Rb, and Ba) were consistently over detection limits and were used for statistical testing. Chemical signatures between core and edge varied significantly across all five elements (MANOVA; $p < 0.001$, Pillai = 0.96). Higher concentrations of Ba (mean difference = 3.29 ppm), Li (mean difference = 0.140 ppm) and Mn (mean difference = 1.62 ppm) were detected in the otolith core relative to the edge, while lower concentrations of Sr (mean difference = -519.70 ppm) were detected in the core (paired t-test; $p < 0.001$; Figure 2). No significant differences in Rb (mean difference = 0.001) were detected between the core and edge (paired t-test; $p = 0.862$; Figure 2).

Chemical signatures in the otolith core of greater amberjack varied significantly among the 3 regions (MANOVA, $p = 0.03$, Pillai = 0.64) and between years ($p = 0.03$, Pillai = 0.40), with no significant effect of the interaction term (Table 1). Univariate contrasts indicated that otolith core Li varied significantly by the interaction term (ANOVA; $p = 0.03$), with a significant main effect among regions ($p = 0.003$). Otolith core Rb concentration differed among regions ($p = 0.004$; Figure 3) with no significant interaction effect. There was no difference in the concentration of Mn and Sr between regions (ANOVA; $p > 0.05$). Although the MANOVA indicated significant regional differences in otolith core chemistry (Pillai = 0.79, p = 0.004), pairwise comparisons using Tukey-HSD revealed no individual region pairs reached were significant after correction for multiple comparisons (all p > 0.05). Overall jackknifed classification success from QDFA to the 3 regions using core element ratios was 57.58%, with variation in classification success among the regions: East (80.0%), West (45.5%), and Southern Gulf (28.6%).

Chemical signatures in the otolith edge of greater amberjack varied significantly by region (MANOVA; $p = 0.004$, Pillai = 0.79) and year ($p = 0.01$). Univariate contrasts indicated that otolith edge Ba, Mn, Rb varied significantly among the 3 regions (ANOVA; $p < 0.05$) with no significant interaction effect (all $p > 0.05$). Otolith edge Rb was significantly higher in the Southeast (0.1194 ppm, 95% CI: 0.1037–0.135) and West (0.1384 ppm, 95% CI: 0.1059–0.171) than East (0.0862 ppm, 95% CI: 0.0693–0.103) (Tukey-adjusted p = 0.025 and p = 0.029, respectively), while Southeast and West did not differ from one another (p = 0.533). For Ba, West tended to be higher than East (1.98 vs. 1.58 ppm) but this difference was marginally non-significant after Tukey adjustment (p = 0.086). Mn followed a similar pattern with West higher than SFL (0.263 vs. 0.185 ppm; p = 0.096), though no pairwise comparisons reached significance. Otolith edge Li varied significantly by the interaction term ($p = 0.030$), with no significant main effects of region (p = 0.685) or year (p = 0.095). Tukey-adjusted pairwise comparisons among interaction combinations revealed no individual pairs reaching significance after adjustment (all $p > 0.12$). Otolith edge Sr differed significantly by year ($p = 0.024$), with a non-significant interaction term ($p = 0.055$) and no main effect of region ($p = 0.855$). Overall jackknifed classification success from QDFA to the 3 regions using edge element ratios (Ba, Mn, Rb) was 69.70%, with variation in classification success among the regions: East (80.0%), West (72.7%), and Southeast (42.9%). Redundancy analysis confirmed that otolith edge elemental composition differed significantly among regions ($p = 0.001$), while core chemistry showed no significant regional structure ($p = 0.097$, Figure 4).

CONCLUSIONS

Population connectivity of greater amberjack may be dependent on larval or juvenile dispersal and not adult movements within the Gulf. Otolith chemistry showed that greater amberjack natal and adult chemical signatures are different, likely reflecting differing habitats between early life (oceanic) and adult (continental shelf) life stages. There was no clear pattern in natal signatures by region; however, regional signatures in otolith edge were more evident. Together, this is consistent with known life history of the species and may reflect greater mixing during early life, with more spatial structuring between the three regions in adult populations.

TABLES AND FIGURES

Table 2: Multivariate analysis of variance (MANOVA) degrees of freedom (number, denominator) F-values, and p-values (Pillai's trace) of natal vs adult habitat (corse vs edge) and the three region (East, West, south FL) groupings of greater amberjack (*Seriola dumerili*) otoliths (core, edge) in the Gulf of America.

Effect	Pillai	F	df	p
Core vs. Edge				
Part	0.962	141.15	5,28	< 0.001***
Fish ID	2.907	1.39	160,160	0.019*
Otolith Core				
Region	0.641	2.26	10,48	0.029*
Year	0.398	3.04	5,23	0.030*
Region × Year	0.510	1.64	10,48	0.123
Otolith Edge				
Region	0.793	3.15	10,48	0.004**
Year	0.459	3.90	5,23	0.010*
Region × Year	0.540	1.78	10,46	0.091.

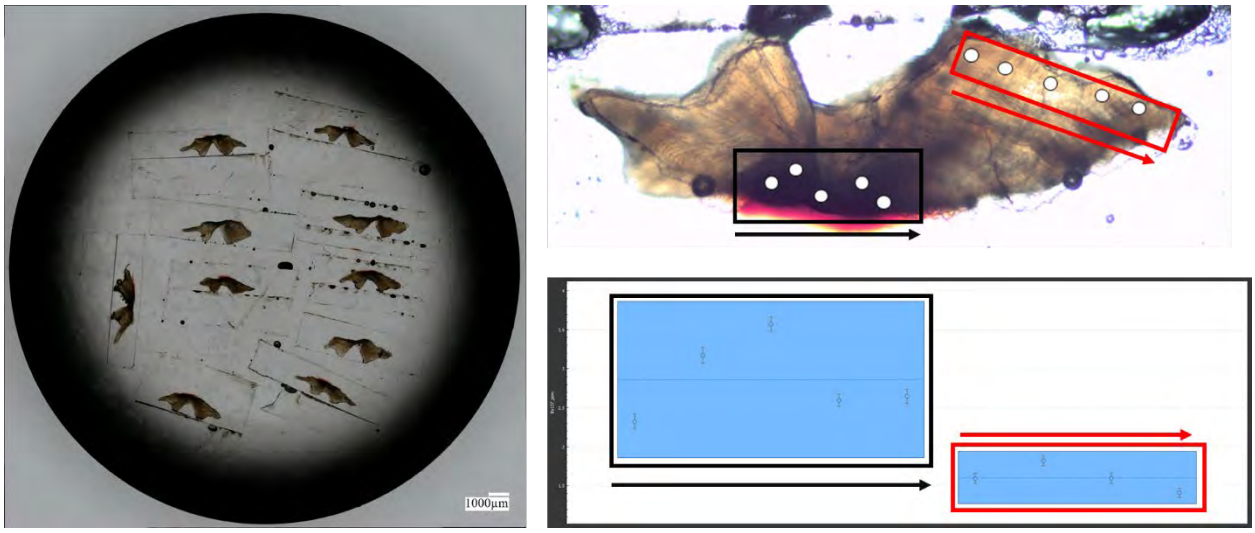


Figure 3: Left: Greater amberjack (*Seriola dumerili*) transverse otolith sections in epoxy in preparation for processing in the LA-ICP-MS. Upper right: Example of core (natal) and edge (adult) laser transect. Lower right: Example of element concentration (^{137}Ba, ppm) in program Iolite. Arrow shows the laser direction.

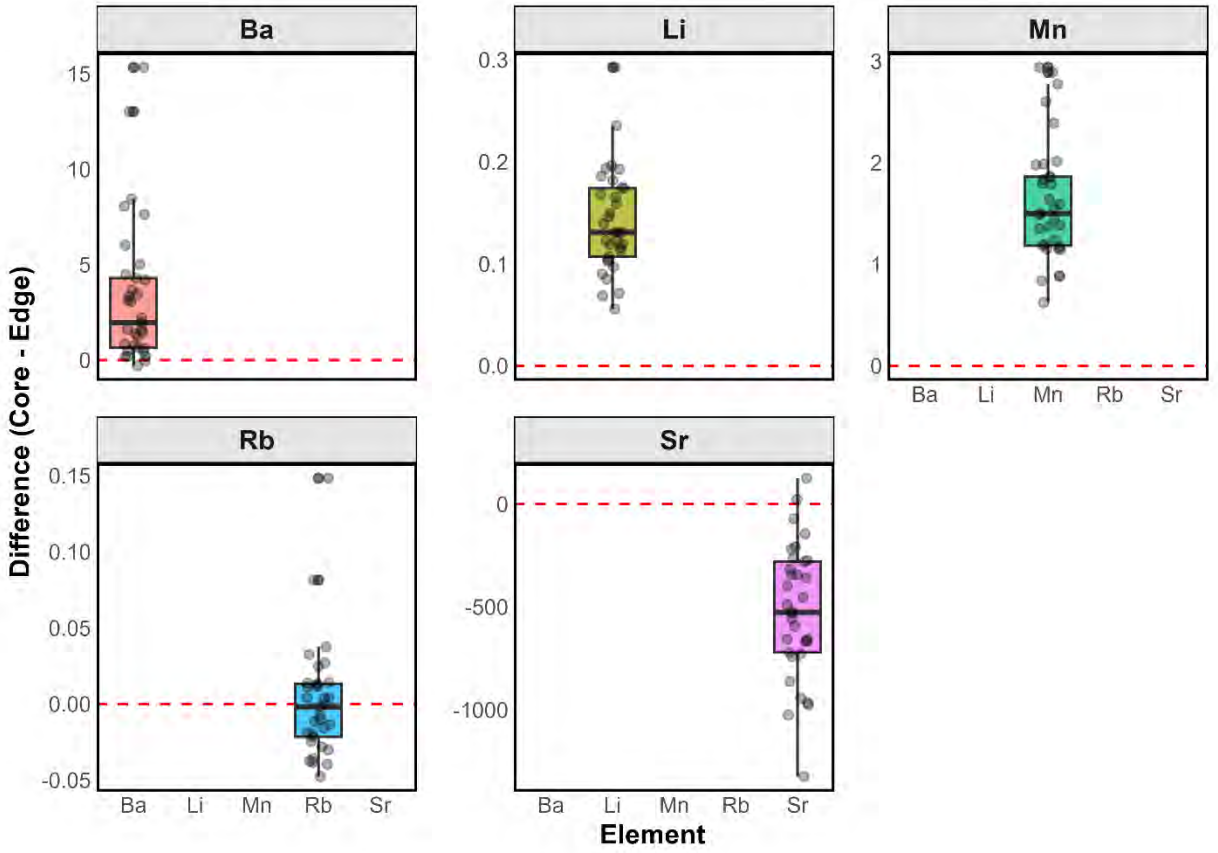


Figure 4: Differences between core and edge element concentrations in greater amberjack (*Seriola dumerili*) otoliths from the Gulf of America. Values above 0 indicate higher concentrations in natal habitat.

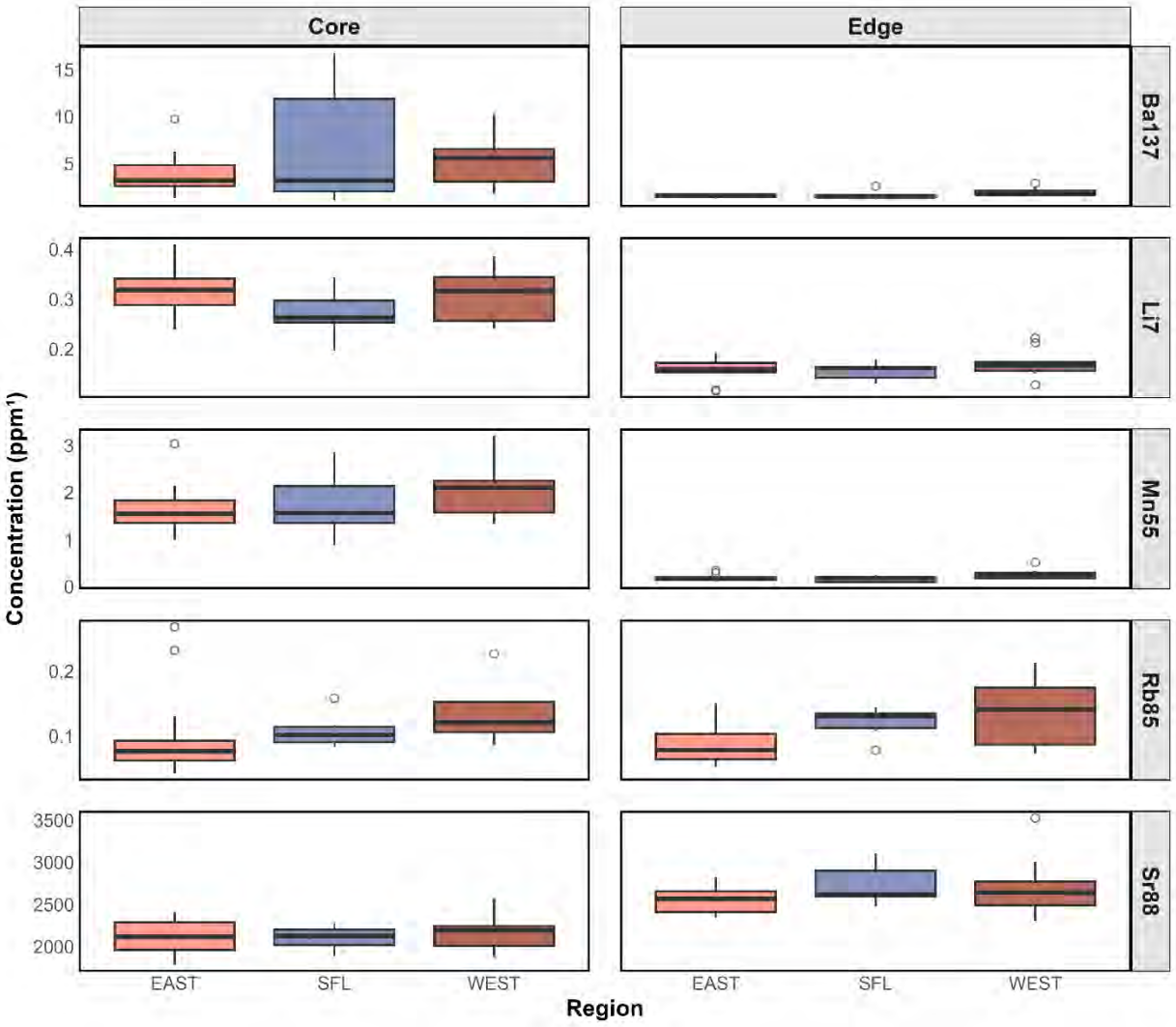


Figure 5: Boxplot of element concentrations (ppm) in greater amberjack (*Seriola dumerili*) otolith edge and core signatures for Ba137, Li7, Mn55, Rb85, and Sr88. The horizontal line in each box indicates the median, the box dimensions represent the 25th to 75th percentile ranges, the whiskers show the 10th to 90th percentile ranges, and dots are outliers.

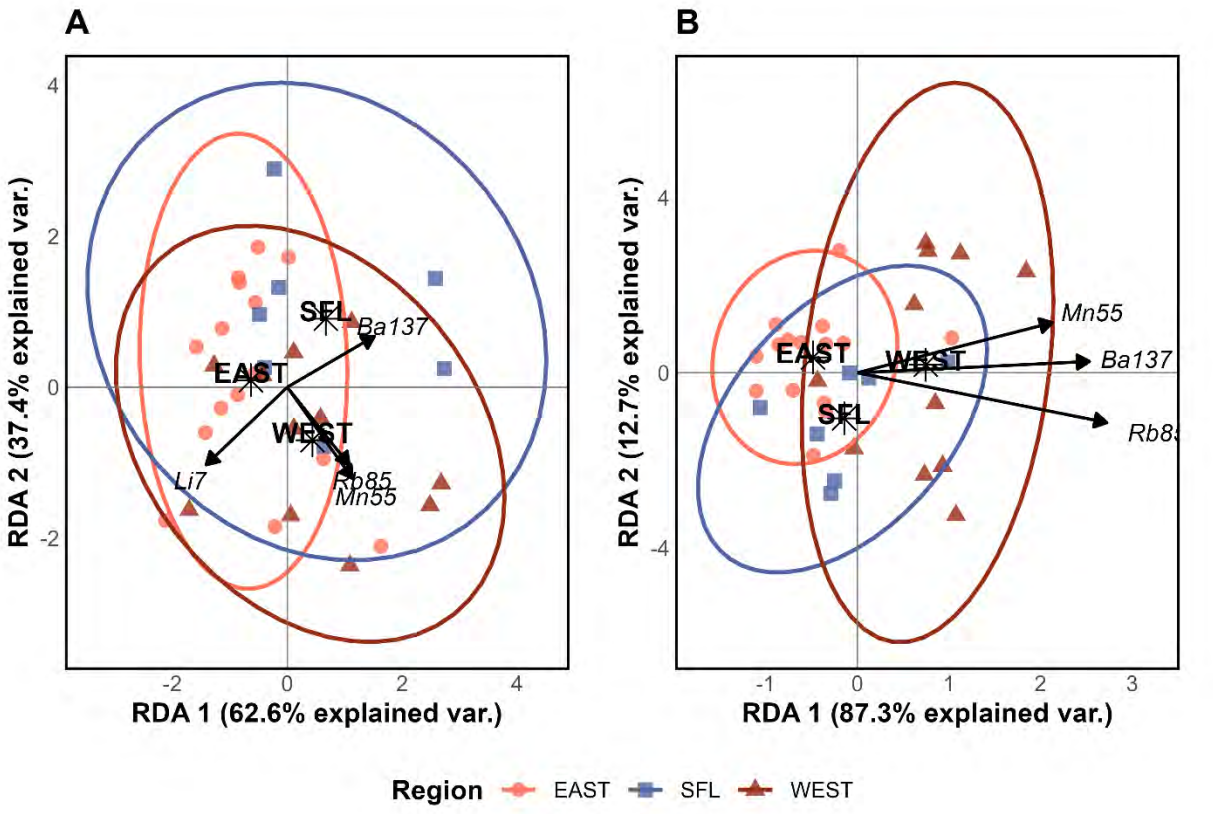


Figure 6: Redundancy plot (RDA) showing regional differentiation in otolith elemental chemistry of greater amberjack (*Seriola dumerili*). Otolith edge chemistry (B) was significantly structured by region (p = 0.001), while core chemistry (A) showed no significant regional constraint (p = 0.097)

REFERENCES

Cadrin, S. X. (2020). Defining spatial structure for fishery stock assessment. *Fisheries Research*, *221*. https://doi.org/10.1016/j.fishres.2019.105397

David, C. L., Marzloff, M. P., Knights, A. M., Cugier, P., Nunes, F. L. D., Cordier, C., Firth, L. B., & Dubois, S. F. (2022). Connectivity modelling informs metapopulation structure and conservation priorities for a reef-building species. *Diversity and Distributions*, *28*(10), 2056-2070. https://doi.org/https://doi.org/10.1111/ddi.13596

Hart, D. R., & Cadrin, S. X. (2004). *Yellowtail Flounder {Limanda ferruginea) off the northeasternUnited States*. Oxford University Press: New York.

Heath, M. R., Kunzlik, P. A., Gallego, A., Holmes, S. J., & Wright, P. J. (2008). A model of meta-population dynamics for North Sea and West of Scotland cod—The dynamic consequences of natal fidelity. *Fisheries Research*, *93*(1), 92-116. https://doi.org/https://doi.org/10.1016/j.fishres.2008.02.014

Hilborn, R., & Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4615-3598-0

James, M. K., Armsworth Pr Fau - Mason, L. B., Mason Lb Fau - Bode, L., & Bode, L. (2002). The structure of reef fish metapopulations: modelling larval dispersal and retention patterns. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1505), 2079–2086. https://doi.org/10.1098/rspb.2002.2128

Kent Smedbol, R., & Wroblewski, J. S. (2002). Metapopulation theory and northern cod population structure: interdependency of subpopulations in recovery of a groundfish population. *Fisheries Research*, *55*(1), 161-174. https://doi.org/https://doi.org/10.1016/S0165-7836(01)00289-2

Kritzer, J. P., & Sale, P. F. (2006). CHAPTER 2 - The Metapopulation Ecology of Coral Reef Fishes. In J. P. Kritzer & P. F. Sale (Eds.), *Marine Metapopulations* (pp. 31-67). Academic Press. https://doi.org/https://doi.org/10.1016/B978-012088781-1/50005-4

Sanchez, P. J., Rooker, J. R., Zapp Sluis, M., Pinsky, J., Dance, M. A., Falterman, B., & Allman, R. J. (2020). Application of otolith chemistry at multiple life history stages to assess population structure of Warsaw grouper in the Gulf of Mexico. *Marine Ecology Progress Series*, *651*, 111-123. https://www.int-res.com/abstracts/meps/v651/meps13457

Seijo, J. C. (2007). Considerations for management of metapopulations in small-scale fisheries of the Mesoamerican barrier reef ecosystem. *Fisheries Research*, *87*(1), 86-91. https://doi.org/https://doi.org/10.1016/j.fishres.2007.06.016

14.4. Appendix D: Greater Amberjack Recreational Survey Instrument

Savannah H. Swinea, Sean P. Powers, Mark A. Albins, J. Marcus Drymon, Steven B. Scyphers

Start of Block: 1 - Intro and Informed Consent

Dear Coastal Angler,

I'm Steven Scyphers from Northeastern University, and I am leading a survey to better understand some of the major issues facing saltwater fishing in the Gulf of Mexico. As recreational and commercial fishers, your insights and response to our survey are essential for our study's success. We would be sincerely grateful for your participation. The survey should only take about 10 minutes to complete.

If you have any questions or would prefer to have a printed copy of the survey mailed to you, please email us at s.scyphers@northeastern.edu or call (781) 581-7370 ext. 375. Many thanks for your consideration and for supporting our research! This study has been reviewed and approved by Northeastern University's Institutional Review Board (Approval #13-07-16).

Additional Information & Frequently Asked Questions about the Survey

1. Who is conducting the research and why? In addition to Northeastern University, other study collaborators include Mississippi State University, Texas A&M - Corpus Christi, University of South Alabama, University of Florida, and several other universities collaborating on a fisheries research study. **2. What are the benefits or risks for my participation in the study?** The decision to participate in this research project is up to you, and you can quit the survey at any time. There are no direct benefits or risks for your participation.

3. How will my survey responses and information be used? Any reports or publications based on this research will use only group data and will not identify you or any individual as being a part of this study. Your part in this study will be handled in a confidential manner.

4. Who funded and approved the study? This study was funded by Mississippi-Alabama Sea Grant Consortium. The research has been reviewed and approved by Northeastern University's Institutional Review Board (Approval #13-07-16).

5. How did you find and access my contact information? If you received this recruitment email directly from our research team, your contact information was acquired through your state's fishing license database. Alternatively, you were contacted through your state's fisheries mailing list or Qualtrics Panels. 6. Who can I contact if I have questions about the survey or research project? You may contact Professor Steven Scyphers,

lead researcher on the survey project. You can reach him at (781) 581-7370 ext. 375 or s.scyphers@northeastern.edu.

6. Who can I contact if I have questions about my rights as a survey participant? You may contact Nan C. Regina, Director of Human Subject Research Protection at Northeastern University. You can reach her at irb@neu.edu.

7. Who can I contact if I have questions about electronic privacy? Please feel free to contact NU's Director of Information Security at privacy@neu.edu. You may download a copy of this information here: Consent.pdf

We care about the quality of our data. In order for us to get the most accurate measures of your opinions, it is important that you thoughtfully provide your best answers to each question in this survey.

Do you commit to thoughtfully provide your best answers to each question in this survey?

- O I will provide my best answers (1)
- O I will not provide my best answers (4)
- O I can't promise either way (5)

End of Block: 1 - Intro and Informed Consent

Start of Block: 2 - Fishing Screeners and State Info

During the past 12 months, have you been *freshwater* fishing?

- O Yes (2)
- O No (4)

During the past 12 months, have you been *saltwater* fishing?

- O Yes (2)
- O No (4)

What state do you currently live in?

- O North Carolina (26)
- O South Carolina (27)
- O Georgia (28)
- O Florida (2)
- O Alabama (1)
- O Mississippi (4)
- O Louisiana (3)
- O Texas (5)
- O Other (16)

Of the 8 states in the South Atlantic and Gulf of Mexico, where have you fished the most during the past 12 months?

- O North Carolina (2)
- O South Carolina (1)
- O Georgia (4)
- O Florida - Gulf of Mexico (3)
- O Florida - Atlantic Ocean (23)
- O Alabama (5)
- O Mississippi (20)
- O Louisiana (21)
- O Texas (22)
- O Other (24)

In addition to ${Q2.4/ChoiceGroup/SelectedChoices}, are there other states where you have been saltwater fishing in the Gulf of Mexico during the past 12 months? (Select all that apply)

- ☐ No, I only fished in one state during the past 12 months. (1)
- ☐ North Carolina (2)
- ☐ South Carolina (3)
- ☐ Georgia (4)
- ☐ Florida - Gulf of Mexico (5)
- ☐ Florida - Atlantic Ocean (6)
- ☐ Alabama (7)
- ☐ Mississippi (8)
- ☐ Louisiana (9)
- ☐ Texas (10)
- ☐ Other (11)

End of Block: 2 - Fishing Screeners and State Info

Start of Block: 3 - General Fishing Characteristics and Behaviors

How often do you fish from each of the following?...

	Never (2)	Less than once per year (3)	Yearly (4)	Monthly (5)	Weekly (6)	Daily (7)
Private Boat with Motor (1)	O	O	O	O	O	O
Kayak or Canoe (4)	O	O	O	O	O	O
Shore, Bridge, or Pier (5)	O	O	O	O	O	O
Charter Boat or Headboat (6)	O	O	O	O	O	O

Which of the following statements best describes you?

O Fishing is an enjoyable, but infrequent activity that is incidental to other travel and outdoor interests. I am not highly skilled in fishing, rarely read fishing articles, and do not own much fishing equipment beyond the basic necessities. (4)

O Fishing is an important, but not exclusive outdoor activity. I occasionally read fishing articles and purchase additional equipment to aid in fishing, my participation in fishing is inconsistent, and I am moderately skilled in fishing. (8)

O Fishing is my primary outdoor activity. I purchase ever-increasing amounts of equipment to aid in fishing, go fishing every chance that I get, consider myself to be highly skilled in fishing, and frequently read fishing articles. (12)

How many years have you been saltwater fishing?

▼ 0 (7) ... 99 (106)

What type(s) of fishing gear do you typically use?

	Never (1)	Sometimes (2)	About half the time (3)	Most of the time (4)	Always (5)
Hook and Line (1)	O	O	O	O	O
Cast Net (2)	O	O	O	O	O
Spear or Powerhead (3)	O	O	O	O	O
Bandit Gear (6)	O	O	O	O	O

In the past 12 months, how many days did you participate in nearshore (less than 3 miles from shore) saltwater recreational fishing off the coast of **${Q2.4/ChoiceGroup/SelectedChoices}**?

▼ 0 (4) ... 365 (369)

In the past 12 months, how many days did you participate in offshore (3 or more miles from shore) saltwater recreational fishing off the coast of **${Q2.4/ChoiceGroup/SelectedChoices}**?

▼ 0 (4) ... 365 (369)

Do you own the boat you typically fish from? If so, what is the length of your boat (in feet)?

O Yes (1) _____

O No (3)

On an average fishing day, how many miles do you travel by boat?

▼ Less than 1 (1) ... More than 200 (225)

When you fish by boat, on a clear day, what percentage of the time are you close enough to see the shore?

▼ Less than 1 (11) ... 100 (526)

End of Block: 3 - General Fishing Characteristics and Behaviors

---

Start of Block: 4 - Reef Fishing Characteristics and Behaviors

The next set of questions focuses on **reef fish**, including snapper, grouper, amberjack, triggerfish, and other species commonly targeted though offshore bottom fishing.

In the context of all of your saltwater fishing, how important to you is offshore fishing for reef fish?

- O Extremely important (31)
- O Very important (32)
- O Moderately important (33)
- O Slightly important (34)
- O Not at all important (35)

How many total years have you been fishing for reef fish?

▼ 1 (7) ... 99 (105)

Which of the following fisheries species do you consider to be ***important*** for your offshore fishing for reef fish? (Select all that apply)

- ☐ Hogfish (5702)
- ☐ Almaco Jack (5703)
- ☐ Banded Rudderfish Jack (5704)
- ☐ Greater Amberjack (5705)
- ☐ Lesser Amberjack (5706)
- ☐ Atlantic Goliath Grouper (5707)
- ☐ Black Grouper (5708)
- ☐ Coney Grouper (5740)
- ☐ Gag Grouper (5709)
- ☐ Graysby Grouper (5739)
- ☐ Nassau Grouper (5741)
- ☐ Red Grouper (5710)
- ☐ Rock Hind (5737)
- ☐ Scamp Grouper (5711)
- ☐ Snowy Grouper (5712)
- ☐ Red Hind (5738)
- ☐ Speckled Hind Grouper (5713)
- ☐ Warsaw Grouper (5714)
- ☐ Yellowedge Grouper (5715)
- ☐ Yellowfin Grouper (5716)
- ☐ Yellowmouth Grouper (5717)

- [ ] Blackfin Snapper (5718)
- [ ] Cubera Snapper (5719)
- [ ] Gray Snapper (5720)
- [ ] Lane Snapper (5721)
- [ ] Mutton Snapper (5722)
- [ ] Queen Snapper (5723)
- [ ] Red Snapper (5724)
- [ ] Silk Snapper (5725)
- [ ] Vermilion Snapper (5726)
- [ ] Wenchman Snapper (5727)
- [ ] Yellowtail Snapper (5728)
- [ ] Blueline Tilefish (5729)
- [ ] Goldface Tilefish (5730)
- [ ] Tilefish (golden) (5731)
- [ ] Gray Triggerfish (5732)
- [ ] None of the above (5736)

What fisheries species do you consider to be ***very important*** for your offshore fishing for reef fish? (Select all that apply)

- ☐ Hogfish (1)
- ☐ Almaco Jack (2)
- ☐ Banded Rudderfish Jack (3)
- ☐ Greater Amberjack (4)
- ☐ Lesser Amberjack (5)
- ☐ Atlantic Goliath Grouper (6)
- ☐ Black Grouper (7)
- ☐ Coney Grouper (8)
- ☐ Gag Grouper (9)
- ☐ Graysby Grouper (10)
- ☐ Nassau Grouper (11)
- ☐ Red Grouper (12)
- ☐ Rock Hind (13)
- ☐ Scamp Grouper (14)
- ☐ Snowy Grouper (15)
- ☐ Red Hind (16)
- ☐ Speckled Hind Grouper (17)
- ☐ Warsaw Grouper (18)
- ☐ Yellowedge Grouper (19)
- ☐ Yellowfin Grouper (20)
- ☐ Yellowmouth Grouper (21)

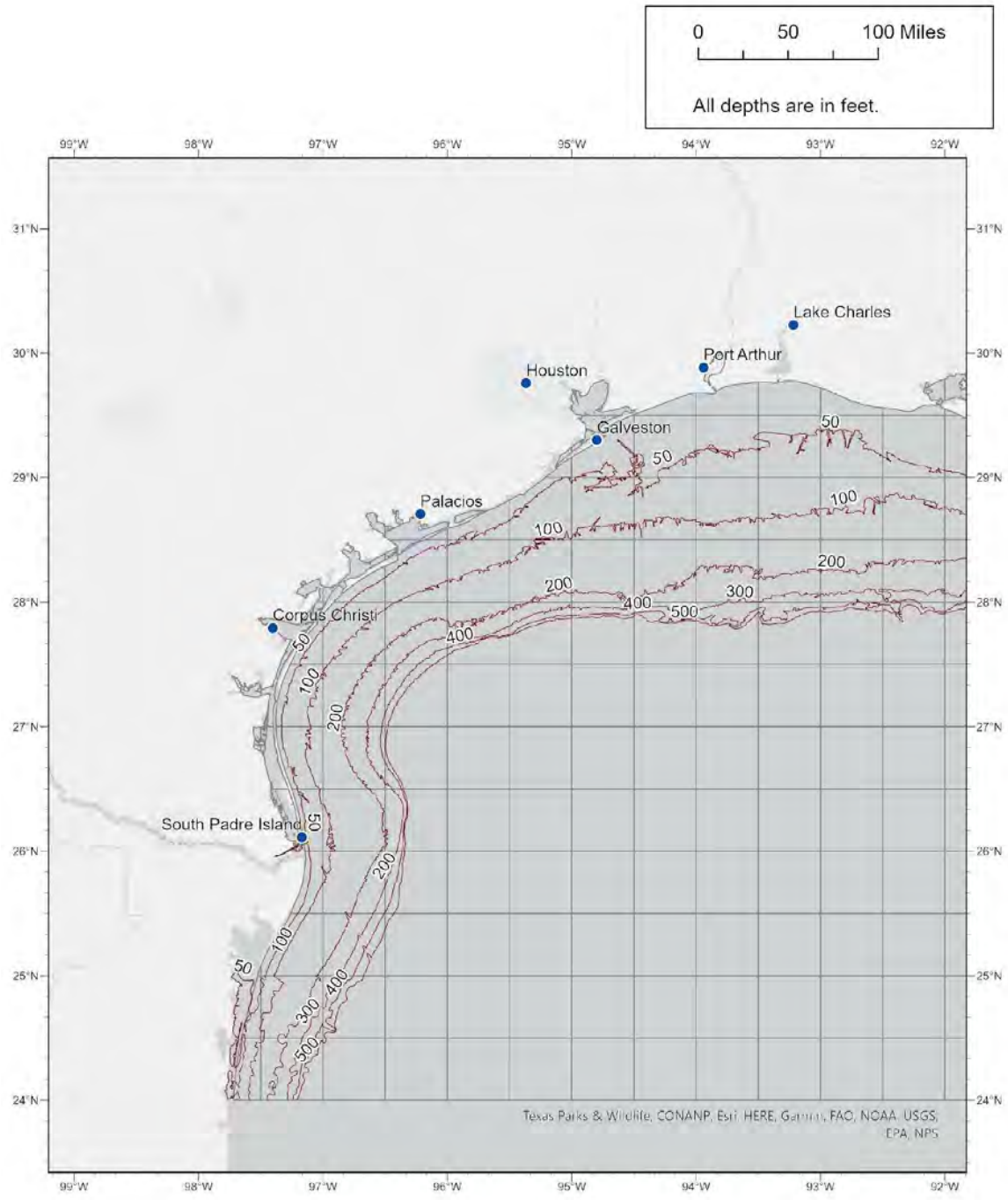
- [ ] Blackfin Snapper (22)
- [ ] Cubera Snapper (23)
- [ ] Gray Snapper (24)
- [ ] Lane Snapper (25)
- [ ] Mutton Snapper (26)
- [ ] Queen Snapper (27)
- [ ] Red Snapper (28)
- [ ] Silk Snapper (29)
- [ ] Vermilion Snapper (30)
- [ ] Wenchman Snapper (31)
- [ ] Yellowtail Snapper (32)
- [ ] Blueline Tilefish (33)
- [ ] Goldface Tilefish (34)
- [ ] Tilefish (golden) (35)
- [ ] Gray Triggerfish (36)
- [ ] None of the above (37)

What fisheries species do you consider to be the ***single most important species*** for your offshore fishing for reef fish? (Choose one)

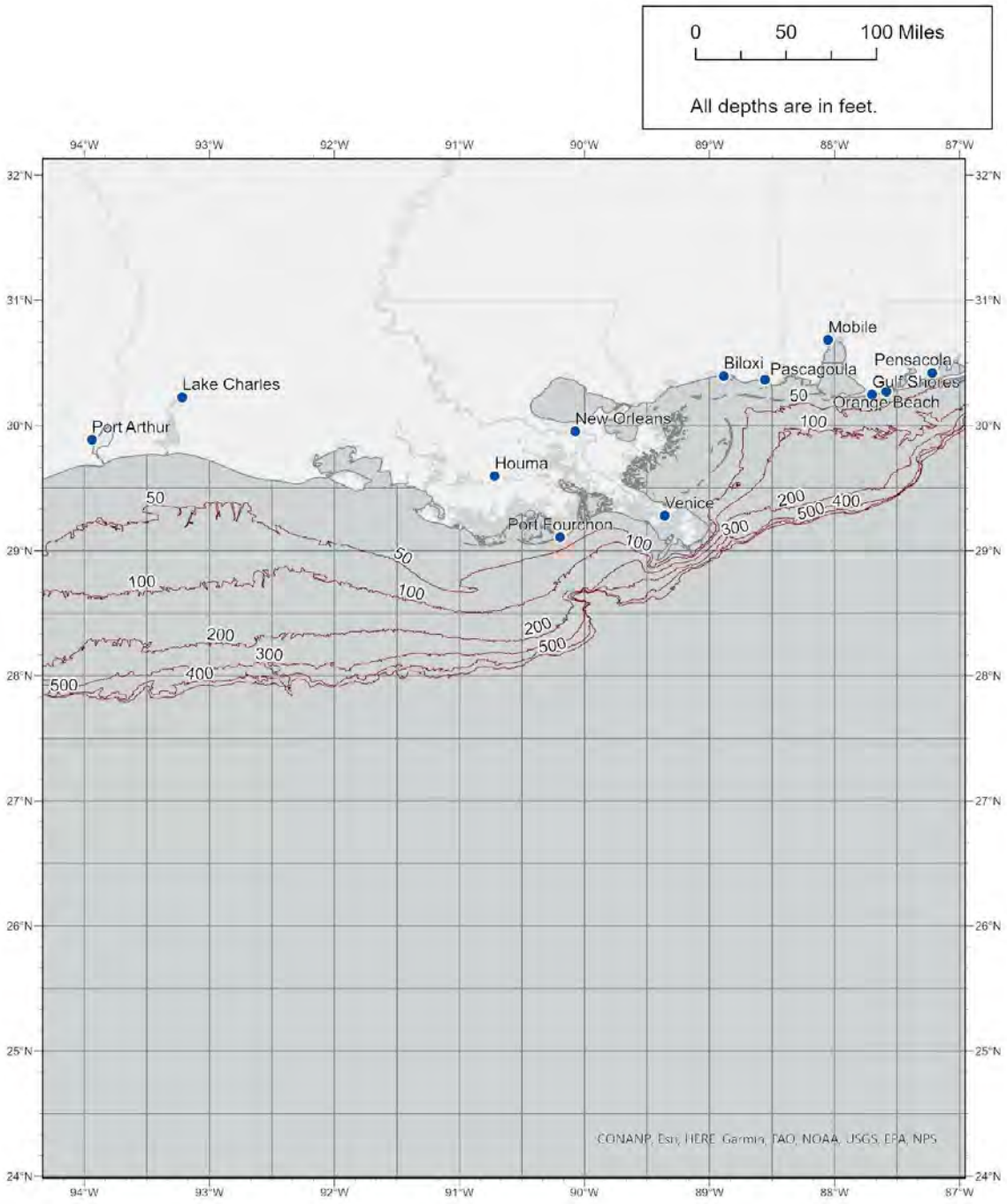
- O Hogfish (1)
- O Almaco Jack (2)
- O Banded Rudderfish Jack (3)
- O Greater Amberjack (4)
- O Lesser Amberjack (5)
- O Atlantic Goliath Grouper (6)
- O Black Grouper (7)
- O Coney Grouper (8)
- O Gag Grouper (9)
- O Graysby Grouper (10)
- O Nassau Grouper (11)
- O Red Grouper (12)
- O Rock Hind (13)
- O Scamp Grouper (14)
- O Snowy Grouper (15)
- O Red Hind (16)
- O Speckled Hind Grouper (17)
- O Warsaw Grouper (18)
- O Yellowedge Grouper (19)
- O Yellowfin Grouper (20)
- O Yellowmouth Grouper (21)
- O Blackfin Snapper (22)
- O Cubera Snapper (23)
- O Gray Snapper (24)
- O Lane Snapper (25)

- O Mutton Snapper (26)
- O Queen Snapper (27)
- O Red Snapper (28)
- O Silk Snapper (29)
- O Vermilion Snapper (30)
- O Wenchman Snapper (31)
- O Yellowtail Snapper (32)
- O Blueline Tilefish (33)
- O Goldface Tilefish (34)
- O Tilefish (golden) (35)
- O Gray Triggerfish (36)
- O None of the above (37)

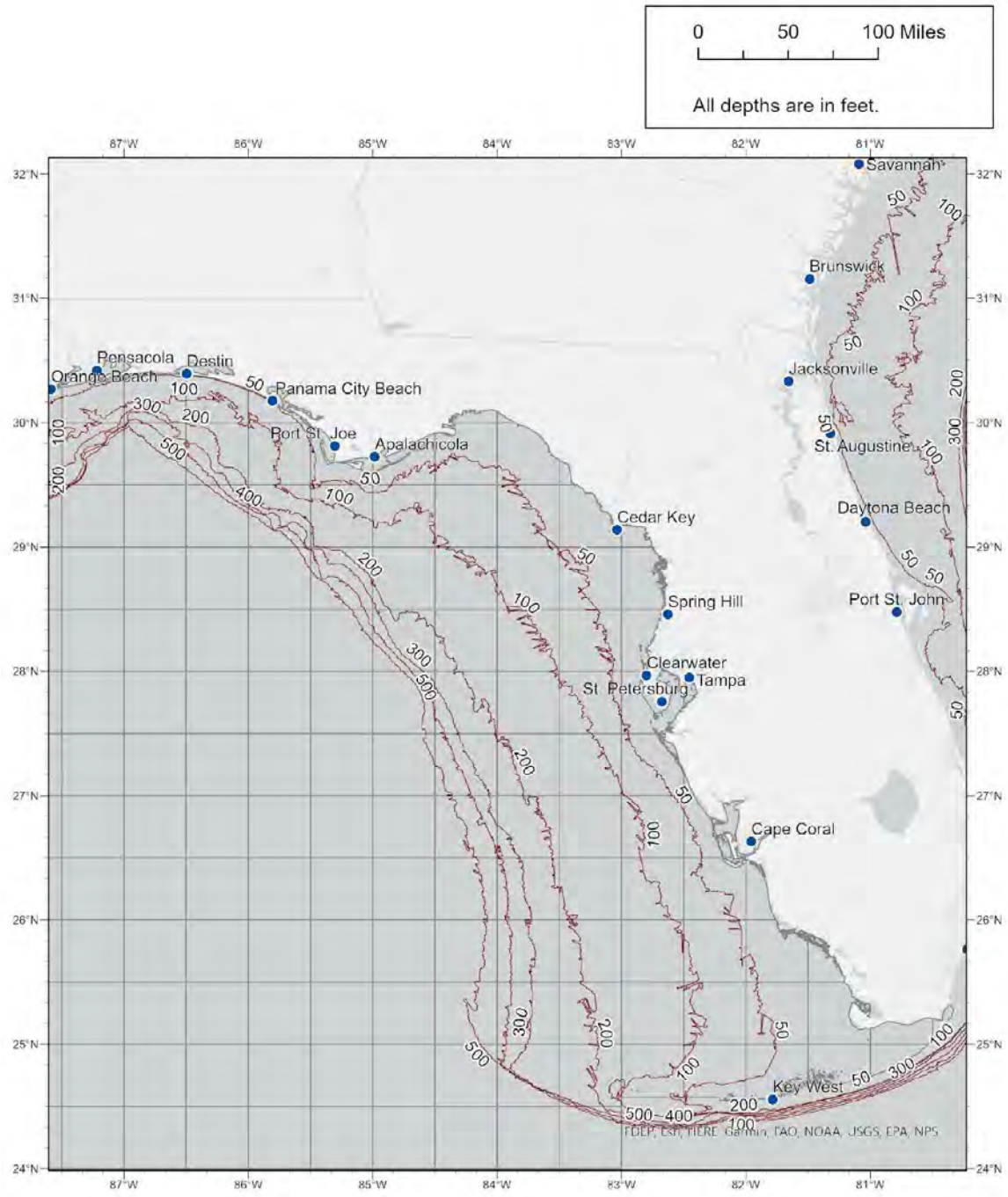
If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



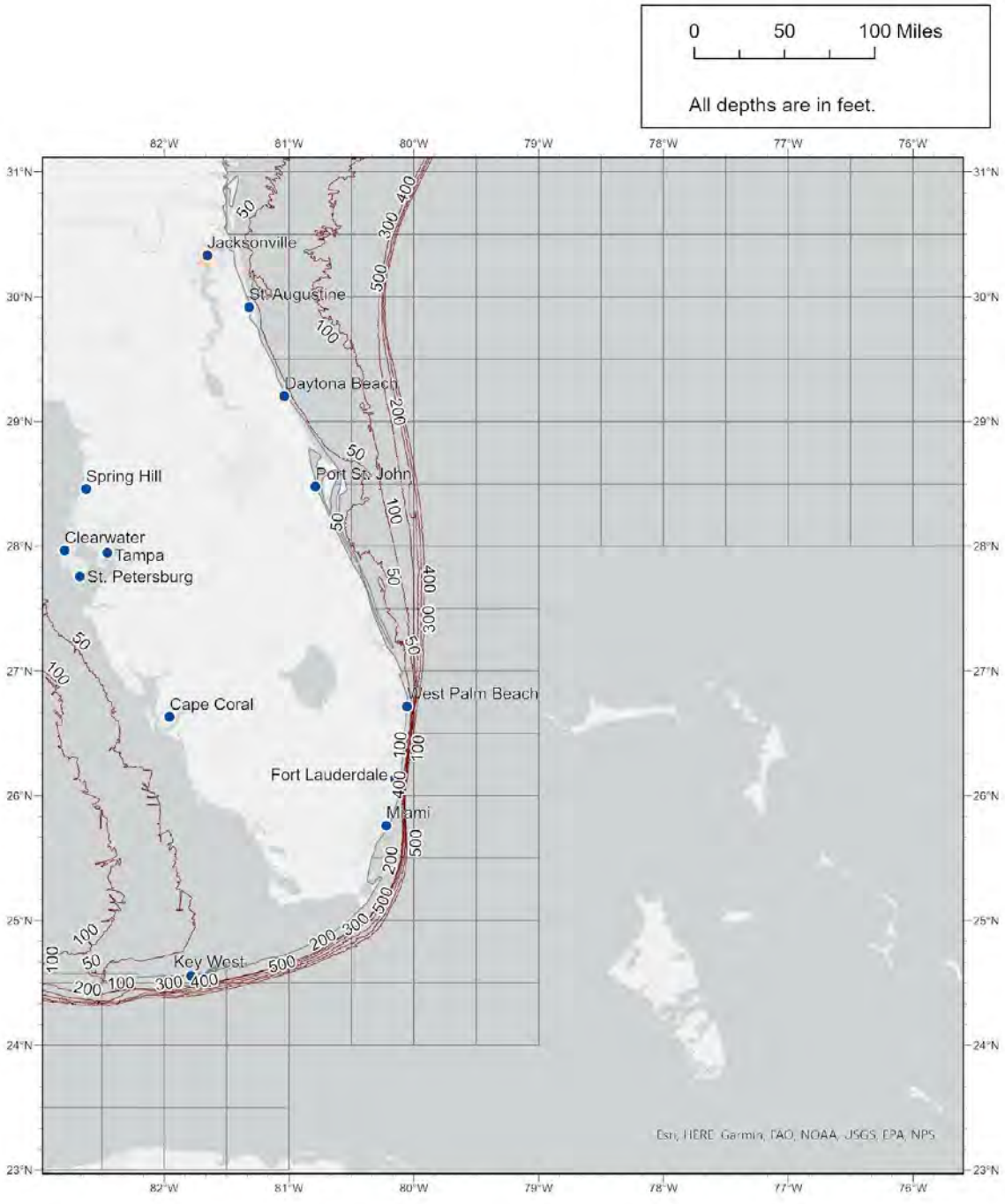
If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



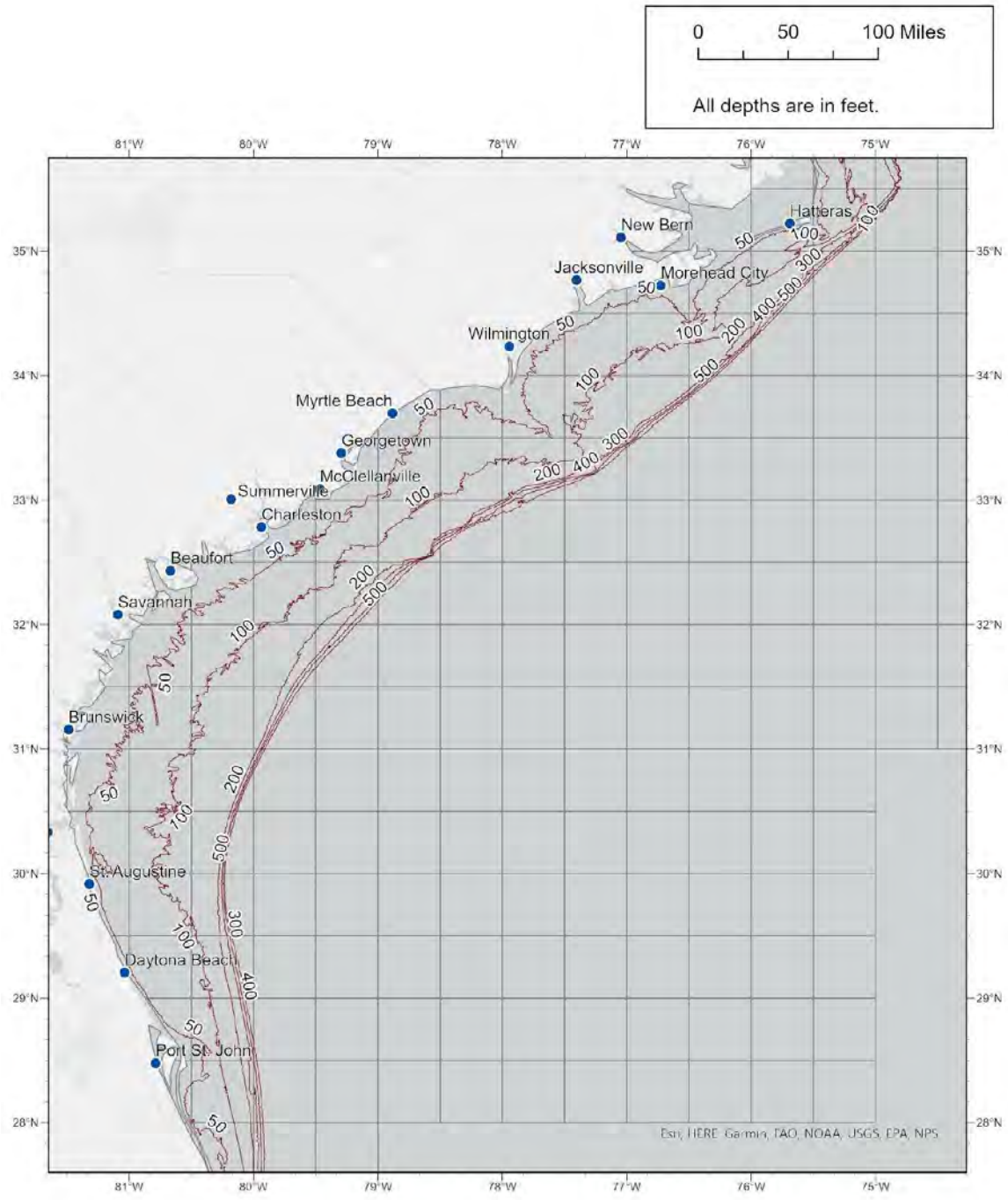
If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



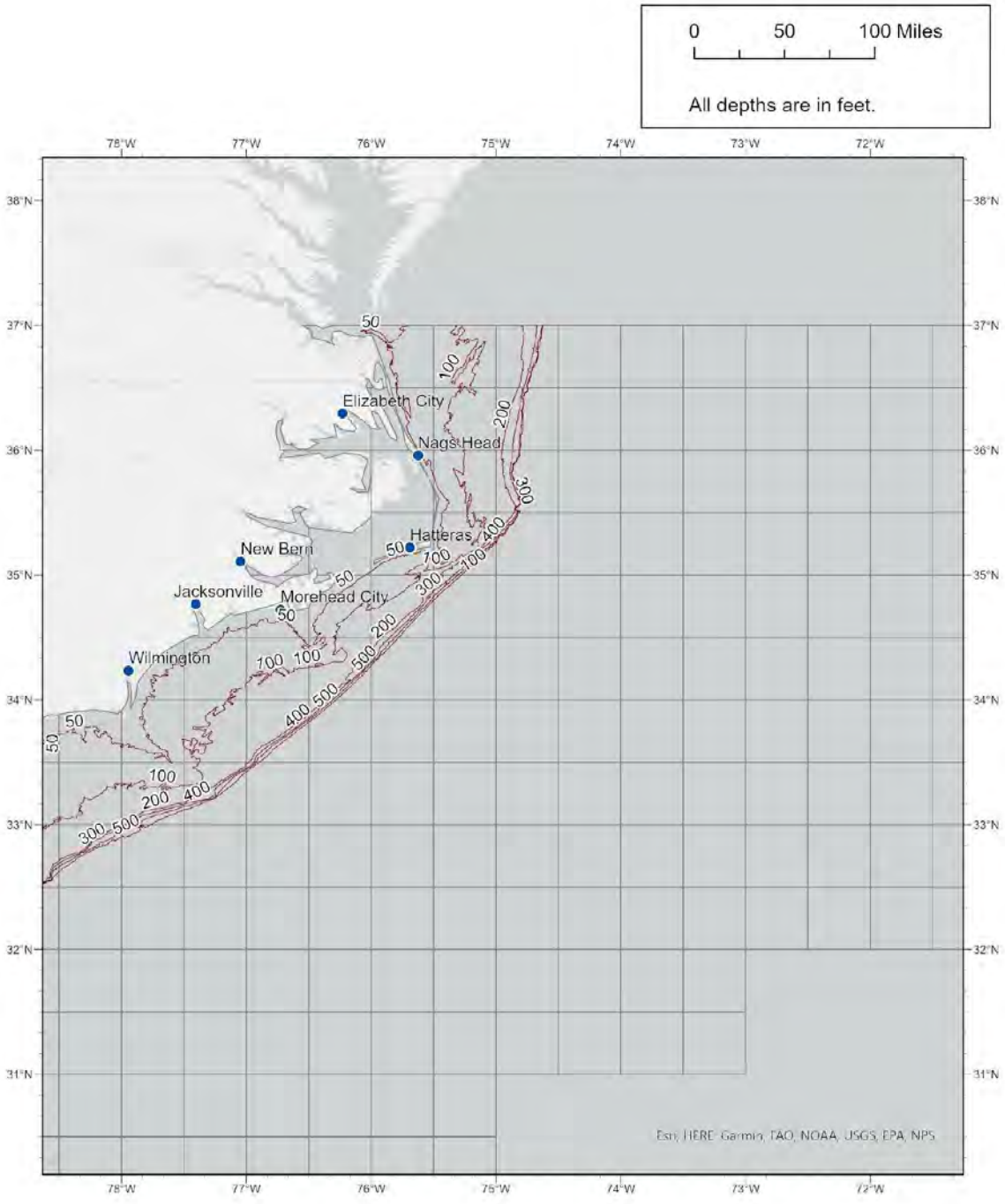
If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.


0 50 100 Miles
All depths are in feet.
82°W
81°W
80°W
79°W
78°W
77°W
76°W
31°N
30°N
29°N
28°N
27°N
26°N
25°N
24°N
23°N
Jacksonville
St. Augustine
Daytona Beach
Port St. John
Spring Hill
Clearwater
Tampa
St. Petersburg
Cape Coral
West Palm Beach
Fort Lauderdale
Miami
Key West
50
100
200
300
400
500
Esri, HERE, Garmin, FAO, NOAA, USGS, EPA, NPS

If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



If you were going fishing specifically for ${Q4.6/ChoiceGroup/SelectedChoices}, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



How would you describe your overall level of satisfaction

with ${Q4.6/ChoiceGroup/SelectedChoices} ***population levels?***

- O Very Satisfied (2)
- O Somewhat Satisfied (3)
- O Neither Satisfied nor Dissatisfied (4)
- O Somewhat Dissatisfied (7)
- O Very Dissatisfied (8)

How would you describe your overall level of satisfaction with ***current fishing regulations*** for ${Q4.6/ChoiceGroup/SelectedChoices}?

- O Very Satisfied (1)
- O Somewhat Satisfied (2)
- O Neither Satisfied nor Dissatisfied (3)
- O Somewhat Dissatisfied (4)
- O Very Dissatisfied (5)

Why do you fish for ${Q4.6/ChoiceGroup/SelectedChoices}?

- ☐ Puts up a good fight (5)
- ☐ Catch-and-release (6)
- ☐ Good tasting (7)
- ☐ Tournament fish/state records (8)
- ☐ Taking home a big "keeper" (9)
- ☐ Fills a gap between other fishing seasons (10)
- ☐ Other: (11) _____

How would you describe your level of satisfaction with the following:

	Not at all Satisfied (1)	Somewhat Satisfied (2)	Very Satisfied (3)	Not Applicable or Not Interested (4)
Availability of Reef Fish to Catch (1)	O	O	O	O
Size of Reef Fish Caught (3)	O	O	O	O
Current Reef Fish Regulations (4)	O	O	O	O
Reef Fish Habitat (6)	O	O	O	O

End of Block: 4 - Reef Fishing Characteristics and Behaviors

Start of Block: 5 - General Fisheries Management Attitudes, Beliefs, and Communication

How much do you agree with the following statement?

I feel that I am able to influence decisions made by fisheries managers.

- O Strongly agree (1)
- O Agree (10)
- O Neither agree nor disagree (11)
- O Disagree (6)
- O Strongly disagree (7)

How would you describe your overall level of knowledge on the ***scientific processes*** *involved in assessing reef fish populations*?

- O Extremely knowledgeable (31)
- O Very knowledgeable (32)
- O Moderately knowledgeable (33)
- O Slightly knowledgeable (34)
- O Not knowledgeable at all (35)

How would you describe your overall level of knowledge on the ***management processes*** *involved with setting regulations for reef fish fisheries*?

- O Extremely Knowledgeable (1)
- O Very Knowledgeable (6)
- O Moderately Knowledgeable (7)
- O Slightly Knowledgeable (8)
- O Not Knowledgeable (9)

For information on reef fish fisheries, how often do you get information from each of the following sources?

	Daily to Weekly (1)	Monthly (2)	Rarely to Never (3)
Online Blogs (1)	O	O	O
Online Chat Forums (25)	O	O	O
Print Magazines (19)	O	O	O
Social Media (Twitter, Facebook) (3)	O	O	O
Television (4)	O	O	O
Radio (5)	O	O	O
Youtube (24)	O	O	O
Peer-reviewed Scientific Journals (6)	O	O	O
Email Newsletters (20)	O	O	O
In-person Meetings (21)	O	O	O
Word of Mouth (22)	O	O	O
Other (please briefly describe) (23)	O	O	O

On a scale of 1 to 5, with 1 indicating "Not Familiar at All" and 5 indicating "Extremely Familiar," how familiar are you with...

	1 (Not Familiar at All) (1)	2 (2)	3 (3)	4 (4)	5 (Extremely Familiar) (5)
My State's Fisheries Management Agency (1)	O	O	O	O	O
State Fisheries Management Agencies in Other States (2)	O	O	O	O	O
NOAA Fisheries (3)	O	O	O	O	O
University and Academic Scientists (4)	O	O	O	O	O
Sea Grant (5)	O	O	O	O	O
Regional Fishery Management Councils (6)	O	O	O	O	O
Other Regional Councils (7)	O	O	O	O	O
Interstate Marine Fisheries Commission (8)	O	O	O	O	O
Regional Council's Science and Statistical Committee (SSC) (9)	O	O	O	O	O
Environmental Organizations (10)	O	O	O	O	O

Fishing Organizations (11)	O	O	O	O	O

On a scale of 1 to 5, with 1 indicating "No Trust" and 5 indicating "Complete Trust," how much do you trust...

	1 (No Trust) (1)	2 (2)	3 (3)	4 (4)	5 (Complete Trust) (5)
Recreational Fishermen (8)	O	O	O	O	O
Commercial Fishermen (7)	O	O	O	O	O
My State's Fisheries Management Agency (x1)	O	O	O	O	O
State Fisheries Management Agencies in Other States (x2)	O	O	O	O	O
NOAA Fisheries (x3)	O	O	O	O	O
University and Academic Scientists (x4)	O	O	O	O	O
Sea Grant (x5)	O	O	O	O	O
Regional Fishery Management Councils (x6)	O	O	O	O	O
Other Regional Councils (x7)	O	O	O	O	O
Interstate Marine Fisheries Commission (x8)	O	O	O	O	O
Regional Council's Science and Statistical Committee (SSC) (x9)	O	O	O	O	O

Environmental Organizations (x10)	O	O	O	O	O
Fishing Organizations (x11)	O	O	O	O	O

End of Block: 5 - General Fisheries Management Attitudes, Beliefs, and Communication

---

Start of Block: GAJ Screen

Have you ever caught a Greater Amberjack?

- O Yes (1)
- O No (2)
- O I'm not sure (3)

How often do you target Greater Amberjack?

- O I only target Greater Amberjack. (1)
- O Frequently (2)
- O Sometimes (3)
- O Rarely (4)
- O I never target Greater Amberjack. (5)

End of Block: GAJ Screen

Start of Block: 6 - GAJ LEK

Why do you fish for **Greater Amberjack**?

- ☐ Puts up a good fight (5)
- ☐ Catch-and-release (6)
- ☐ Good tasting (7)
- ☐ Tournament fish/state records (8)
- ☐ Taking home a big "keeper" (9)
- ☐ Fills a gap between other fishing seasons (10)
- ☐ Other: (11) _____

How confident are you that you can identify a Greater Amberjack?

- ○ Not at all confident (1)
- ○ Slightly confident (2)
- ○ Moderately confident (3)
- ○ Very confident (4)
- ○ Extremely confident (5)

According to *NOAA Fisheries*, which of the following best describes the current status of Greater Amberjack stocks in the region where you fish most?

- ○ Overfished and Experiencing Overfishing (1)
- ○ Overfished and Rebuilding (5)
- ○ Experiencing Overfishing but Not Overfished (7)
- ○ Not Experiencing Overfishing and Not Overfished (6)
- ○ I'm not sure (8)

How would you describe the current trend in Greater Amberjack *abundance* off the coast of **${Q2.4/ChoiceGroup/SelectedChoices}**?

- O Increasing in abundance (1)
- O Not changing in abundance (2)
- O Decreasing in abundance (3)
- O I'm not sure (4)

When did the *numbers* of Greater Amberjack start changing?

- O Before 1950 (1)
- O 1950-1960 (2)
- O 1960-1970 (3)
- O 1970-1980 (4)
- O 1980-1990 (5)
- O 1990-2000 (6)
- O 2000-2010 (7)
- O 2010-2020 (8)
- O 2020-present (9)
- O I don't know (10)

What do you think is causing a change in Greater Amberjack *abundance*?

_____

_____

_____

_____

_____

How would you describe the current trend in average Greater Amberjack *size* off the coast of **${Q2.4/ChoiceGroup/SelectedChoices}**?

- O Increasing in size (2)
- O Not changing in size (3)
- O Decreasing in size (4)
- O I'm not sure (5)

When did the average *size* of Greater Amberjack start changing?

- O Before 1950 (1)
- O 1950-1960 (2)
- O 1960-1970 (3)
- O 1970-1980 (4)
- O 1980-1990 (5)
- O 1990-2000 (6)
- O 2000-2010 (7)
- O 2010-2020 (8)
- O 2020-present (9)
- O I don't know (10)

What factors do you think are causing a change in average Greater Amberjack *size*?

_____

_____

_____

_____

_____

How would you describe your overall level of satisfaction with *Greater Amberjack* ***population levels?***

- O Very Satisfied (2)
- O Somewhat Satisfied (3)
- O Neither Satisfied nor Dissatisfied (4)
- O Somewhat Dissatisfied (7)
- O Very Dissatisfied (8)

How would you describe your overall level of satisfaction with ***current fishing regulations*** *for Greater Amberjack*?

- O Very Satisfied (1)
- O Somewhat Satisfied (2)
- O Neither Satisfied nor Dissatisfied (3)
- O Somewhat Dissatisfied (4)
- O Very Dissatisfied (5)

Which of the following habitats do you think is best for Greater Amberjack fishing?

- O Natural Structured Bottom (Live Patches, Reefs) (1)
- O Natural Sand / Mud Bottom (14)
- O Artificial Reefs (Pyramids, Chicken Coops) (12)
- O Shipwrecks (3)
- O Offshore Oil / Gas Rigs (5)
- O Sunken Rigs (7)
- O Bridge Rubble (9)
- O Other: (11) _____
- O None of the Above (13)

If you were fishing at a site with **${Q6.16/ChoiceGroup/SelectedChoices}**, how many Greater Amberjack would you expect to catch during one hour of fishing on an average fishing day?

_____

If you were fishing for Greater Amberjack at a site with **${Q6.16/ChoiceGroup/SelectedChoices}**, what approximate depth (in feet) would you fish from on an average fishing day?

▼ 50 (4) ... 500 (13)

If you were fishing at a site with **Artificial Reefs (Pyramids, Chicken Coops)**, how many Greater Amberjack would you expect to catch during one hour of fishing on an average fishing day?

_____

If you were fishing for Greater Amberjack at a site with **Artificial Reefs (Pyramids, Chicken Coops)**, what approximate depth (in feet) would you fish from on an average fishing day?

▼ 50 (4) ... 500 (13)

If you were fishing at a site with **Natural Sand / Mud Bottom**, how many Greater Amberjack would you expect to catch during one hour of fishing on an average fishing day?

_____

If you were fishing for Greater Amberjack at a site with **Natural Sand / Mud Bottom**, what approximate depth (in feet) would you fish from on an average fishing day?

▼ 50 (4) ... 500 (13)

Did you go offshore fishing for reef fish last month (May 2022)?

O Yes (1)

O No (2)

How many offshore fishing trips did you take last month?

▼ 1 (32) ... 31 (31)

Did you catch or try to catch Greater Amberjack when fishing last month? If so, please describe your trip and catch.

_____

_____

_____

_____

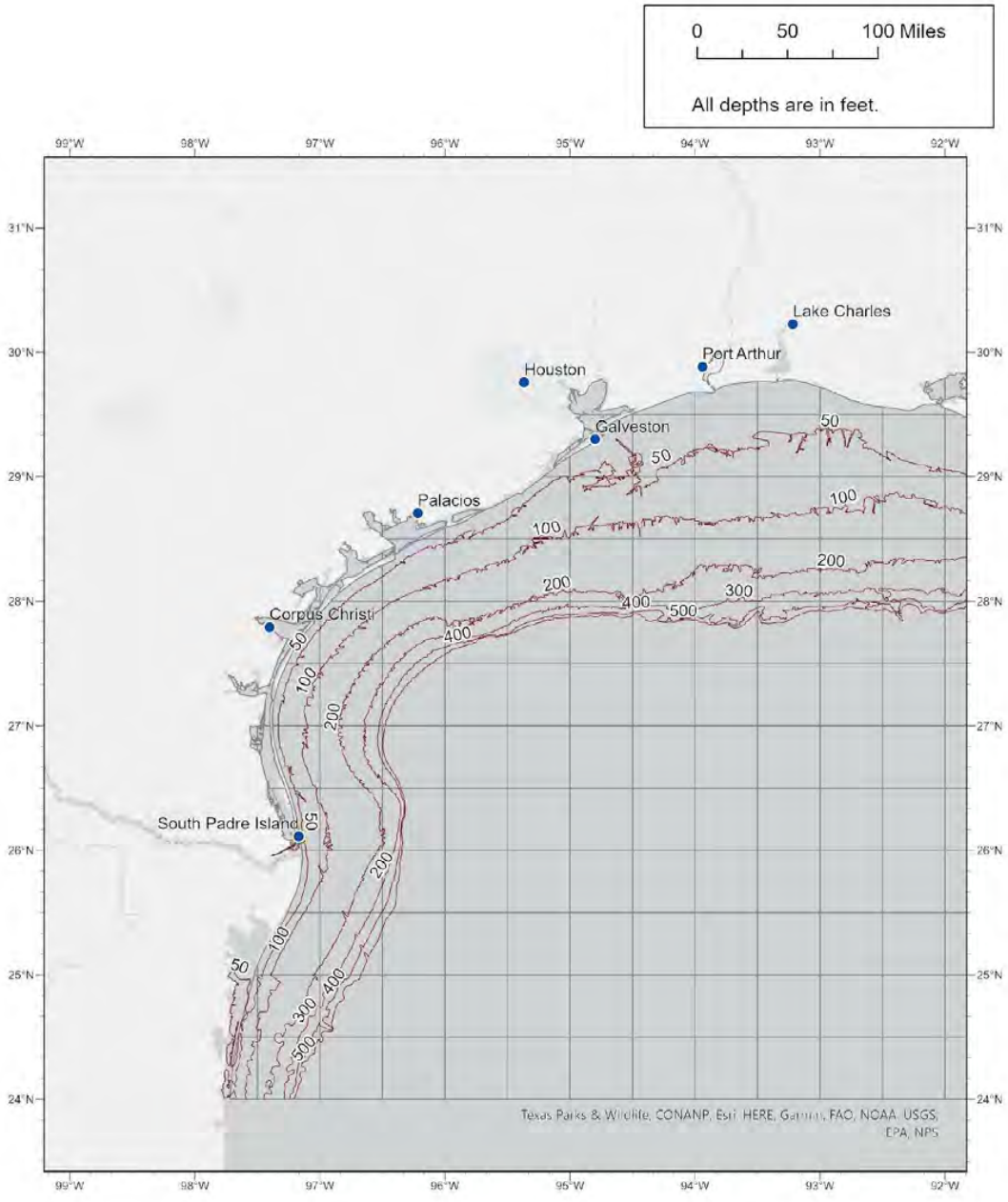
_____

End of Block: 6 - GAJ LEK

Start of Block: 7 - GAJ Participatory Mapping (Tagging)

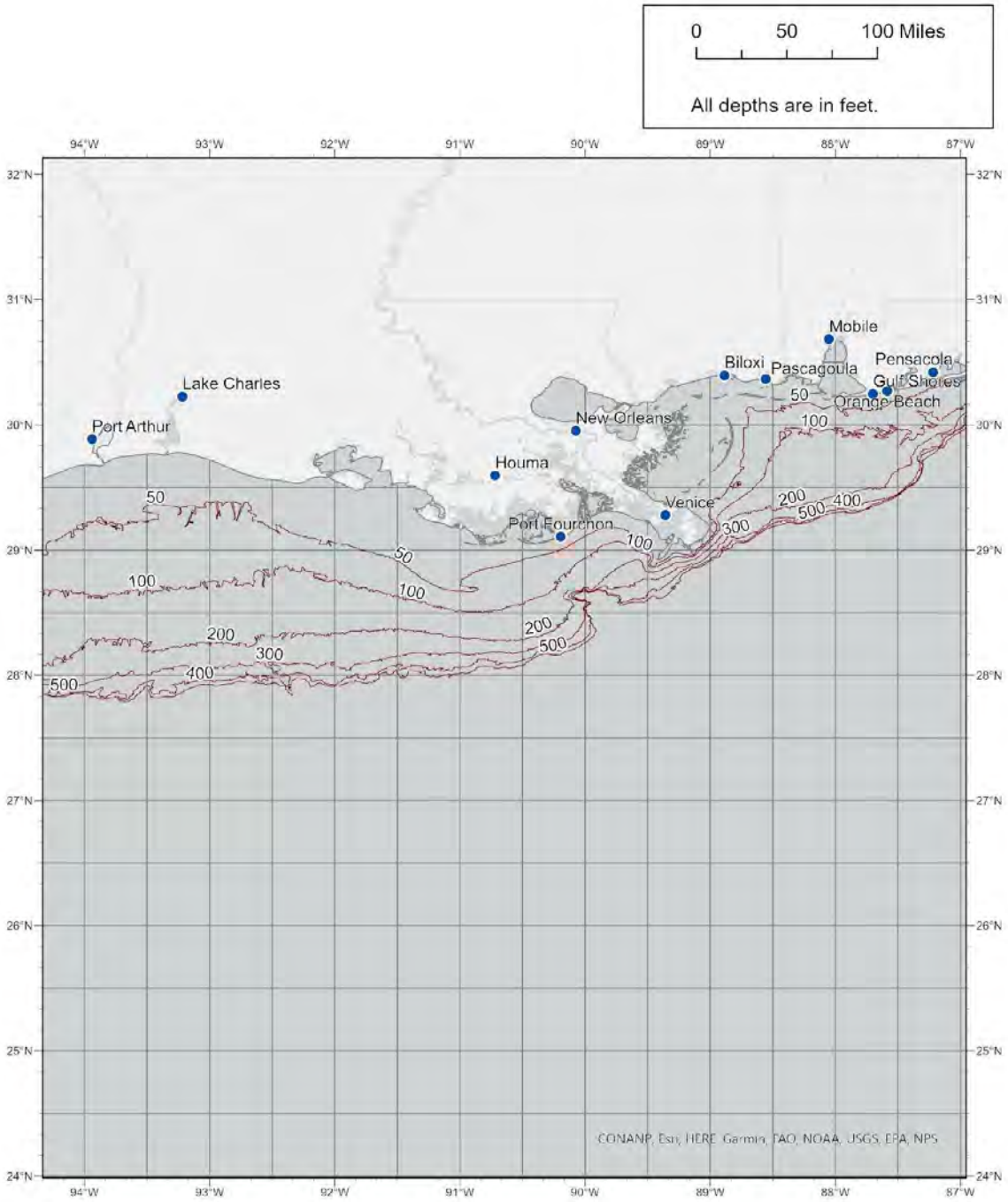
For the next series of questions, please consider all of your saltwater fishing **for only last year** (2021).

If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map as you would fish from.

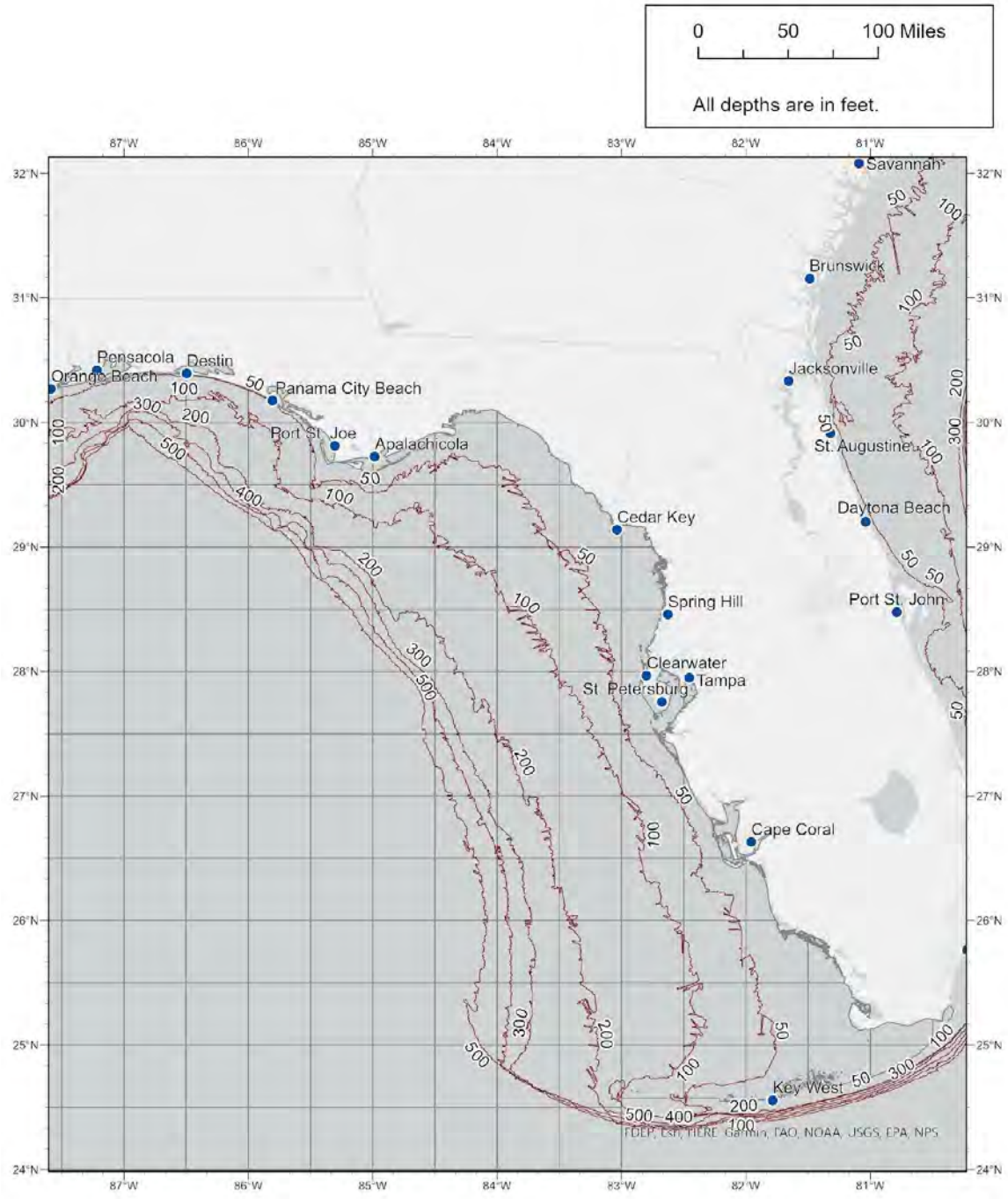


If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map

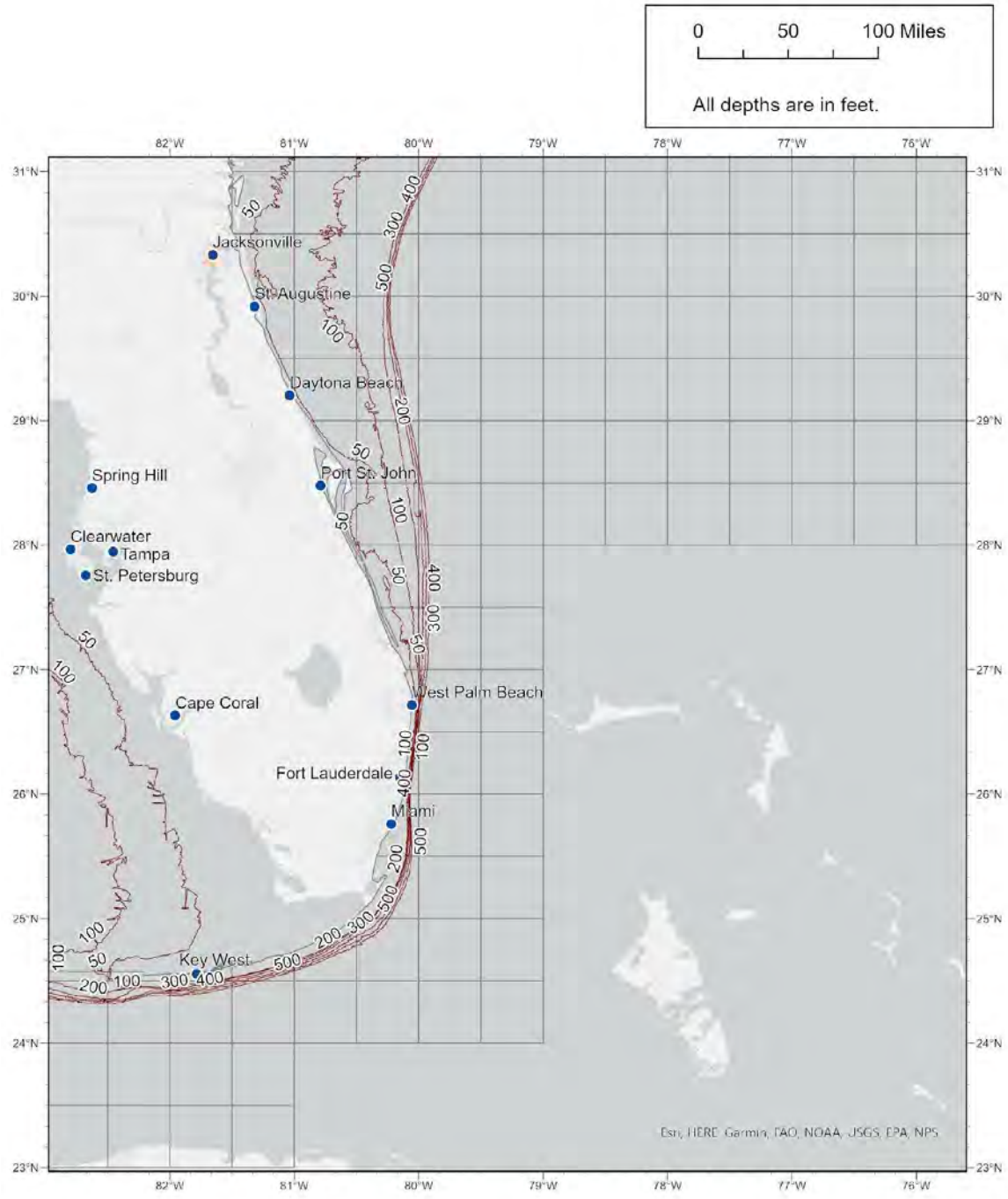
as you would fish from.



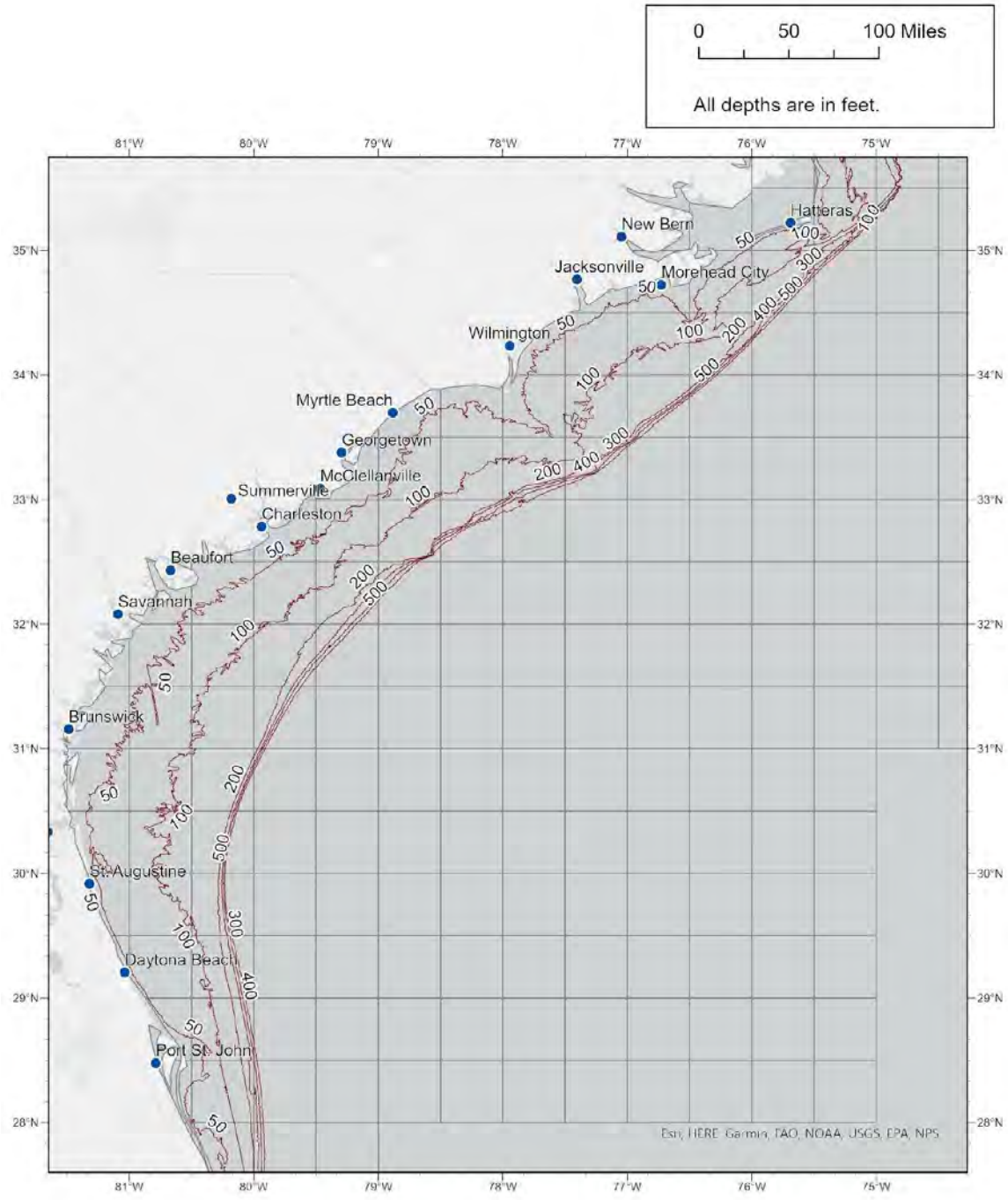
If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map as you would fish from.



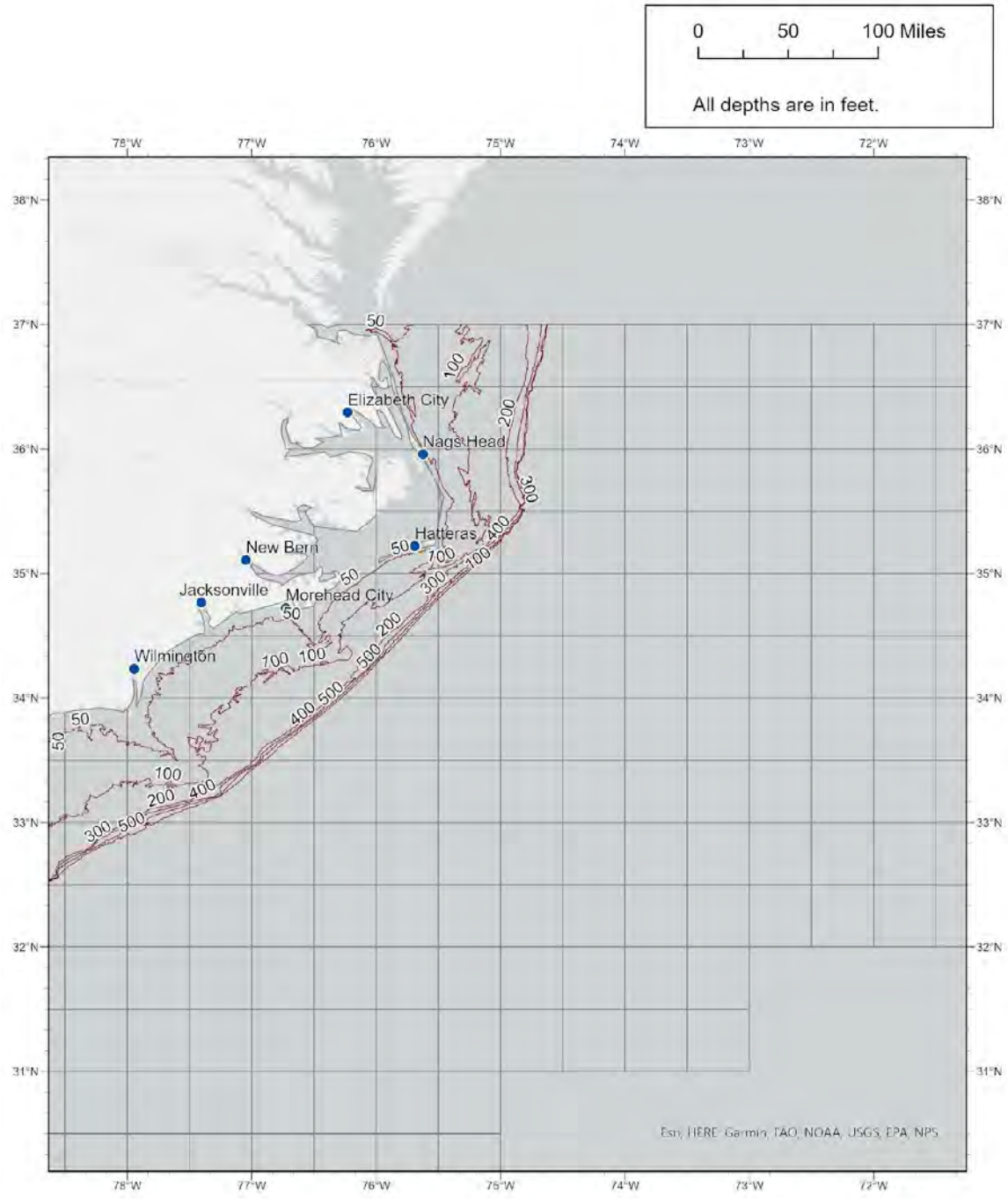
If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map as you would fish from.



If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map as you would fish from.



If you were invited to work with scientists on a tagging project for Greater Amberjack, where would you fish to have the best chance of catching a Greater Amberjack? Please choose as many locations on the map as you would fish from.

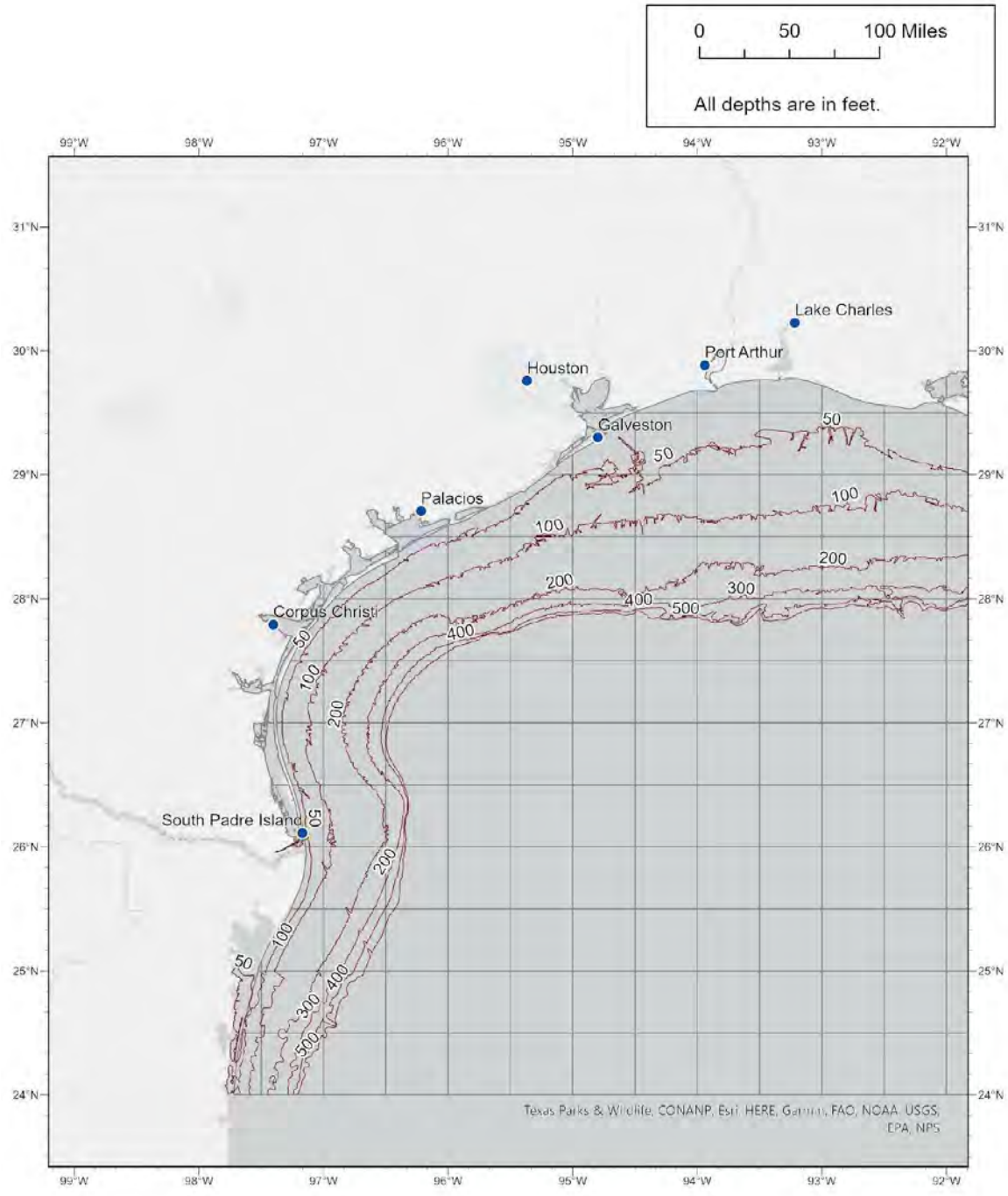


End of Block: 7 - GAJ Participatory Mapping (Tagging)

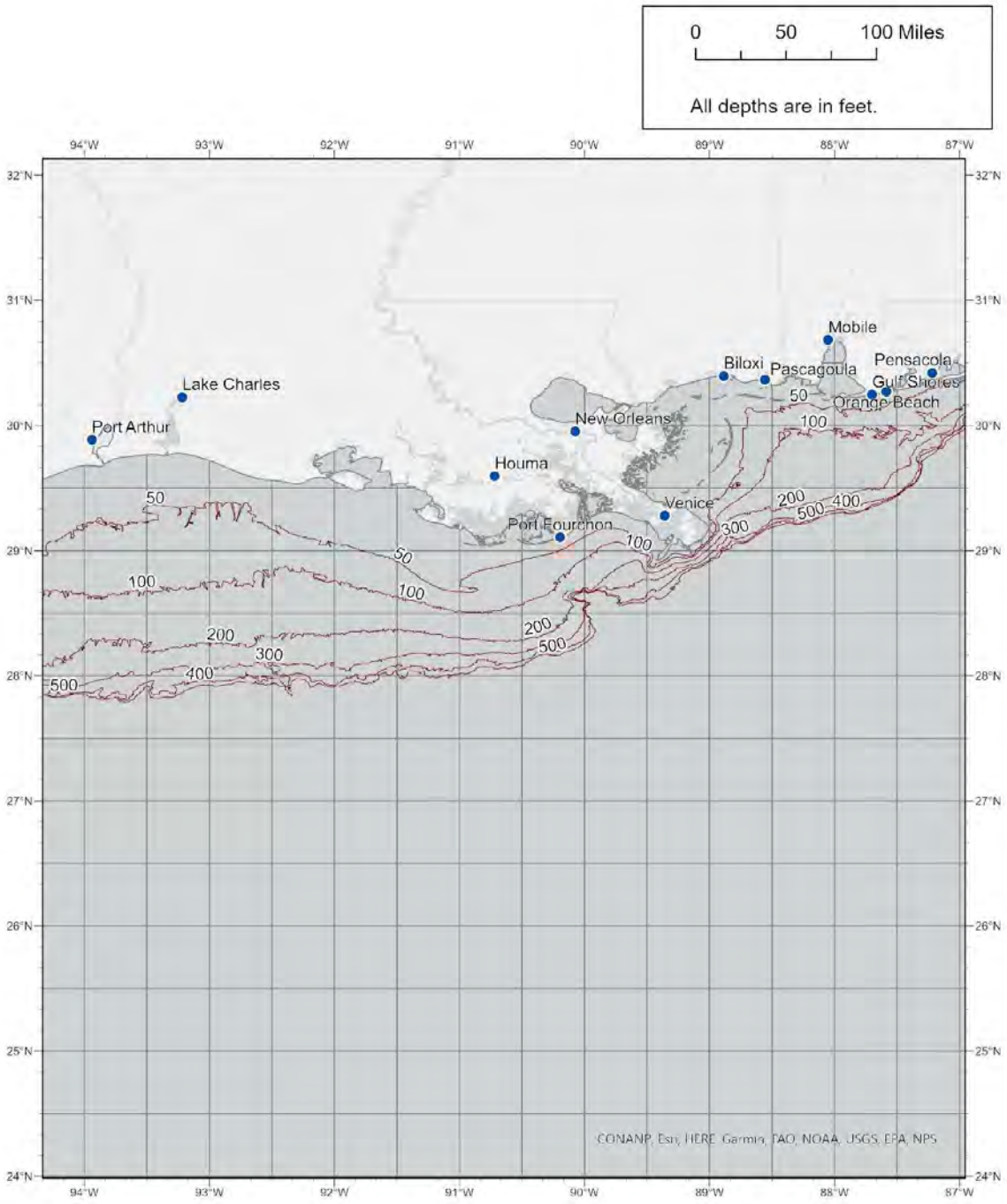
Start of Block: 8 - GAJ Participatory Mapping (Control)

For the next series of questions, please consider all of your saltwater fishing **for only last year** (2021).

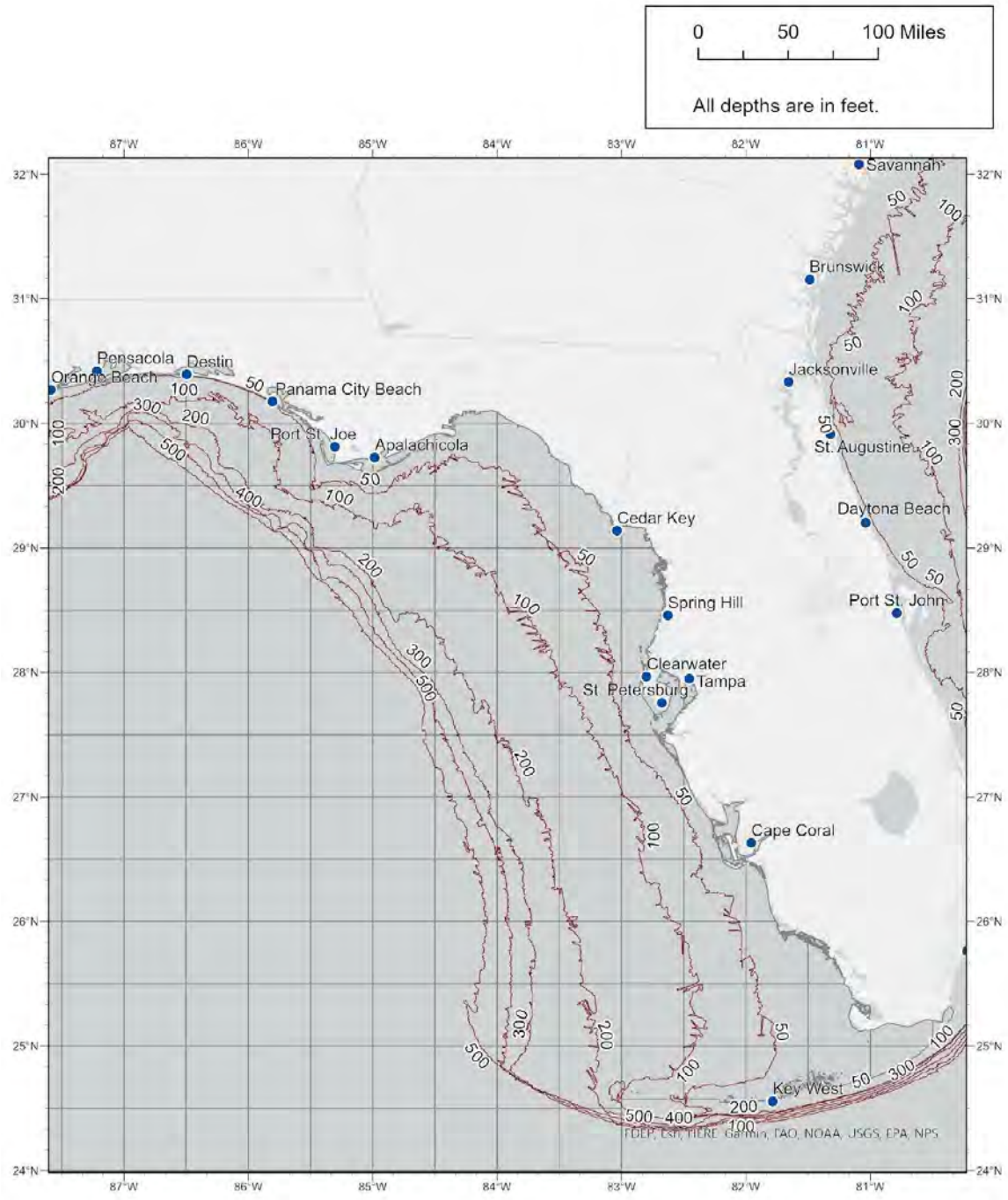
If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



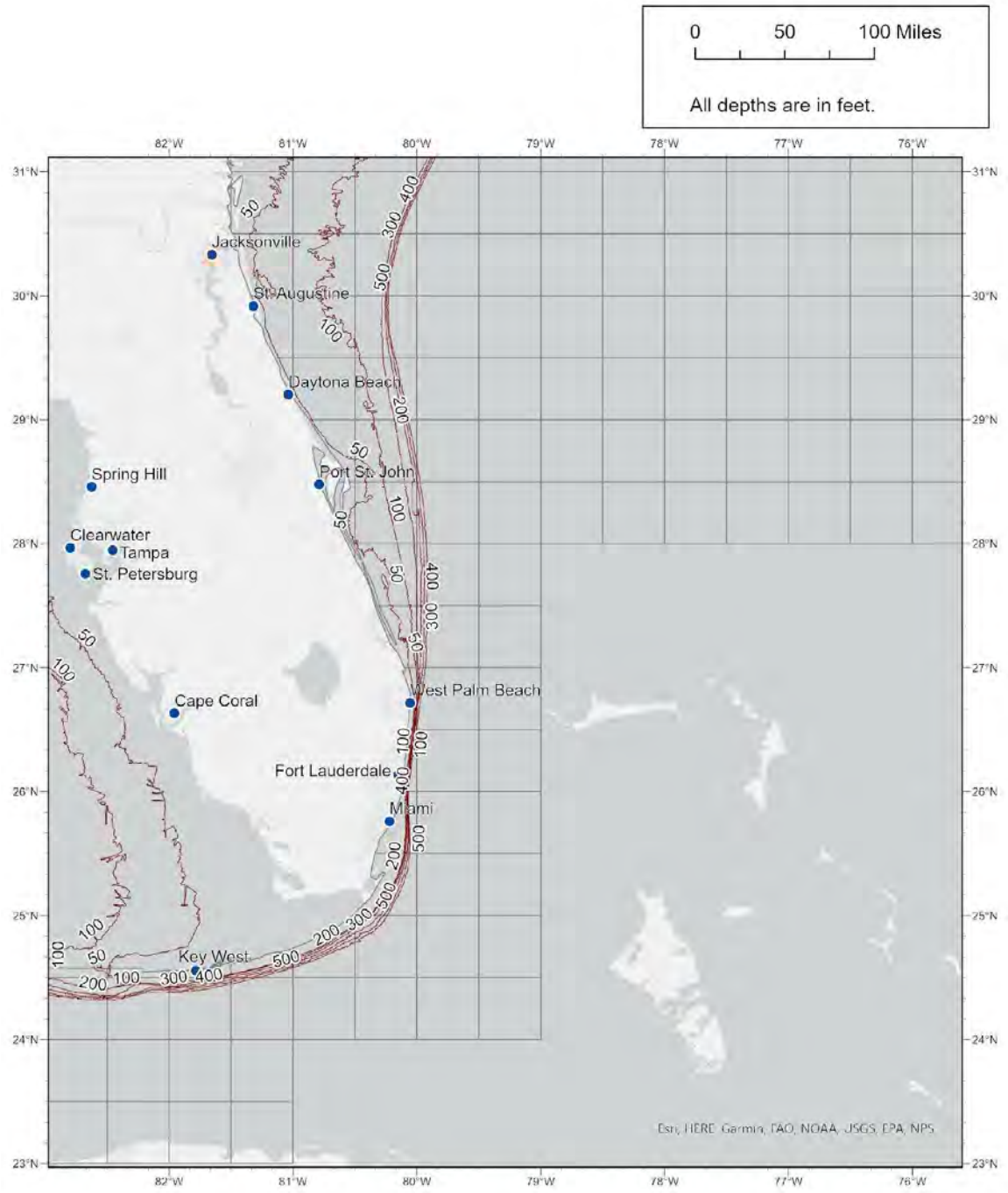
If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



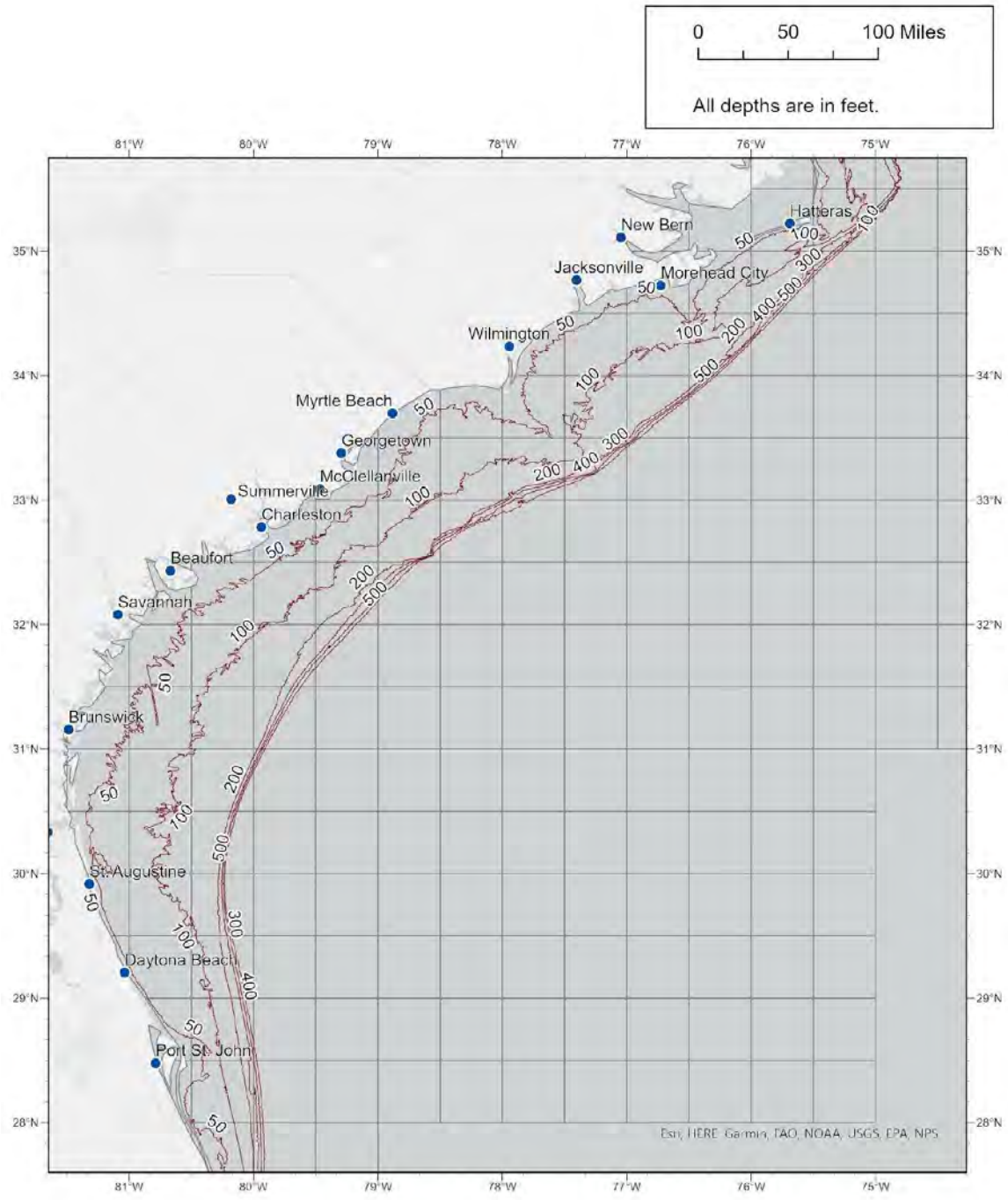
If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



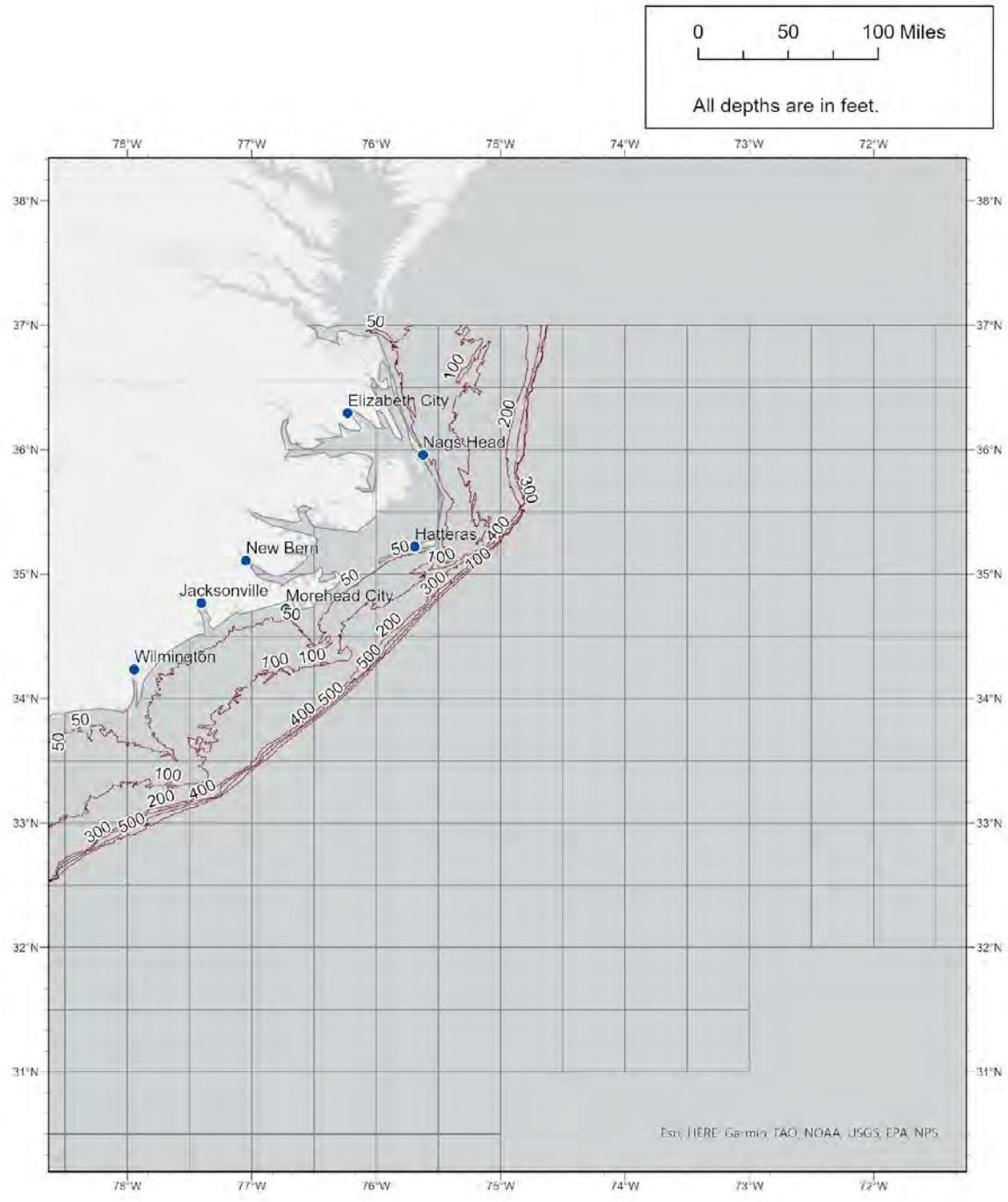
If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



If you were going fishing specifically for Greater Amberjack, where would you fish to have the best chance of catching one? Please choose as many locations on the map as you would fish from.



End of Block: 8 - GAJ Participatory Mapping (Control)

Start of Block: 9 - Demographics

What is your ZIP code?

_____

What year were you born?

▼ 2006 (114) ... Prefer not to answer (116)

What is your highest level of education? (Please select one)

- O Less than high school (1)
- O High school diploma or GED (2)
- O Some college or 2 year degree (3)
- O Bachelor's degree (4)
- O Master's degree (5)
- O Law or MD (6)
- O Doctorate (PhD) (7)
- O Prefer not to answer (8)

What was your total household income in 2021?

- O $25k or less (1)
- O $25,001 to $35k (2)
- O $35,001 to $50k (3)
- O $50,001 to $75k (4)
- O $75,001 to $100k (5)
- O $100,001 to $150k (6)
- O $150,000 to $250k (7)
- O More than $250k (8)
- O Prefer not to answer (9)

What is your gender?

- O Man (1)
- O Woman (2)
- O Other (3)
- O Prefer not to answer (4)

What is your race or ethnicity?

- ☐ White (1)
- ☐ Black or African American (2)
- ☐ American Indian or Alaska Native (3)
- ☐ Asian (4)
- ☐ Native Hawaiian or other Pacific Islander (5)
- ☐ Hispanic or Latino (6)
- ☐ ⊗Prefer not to answer (7)

Some people who work in industries such as fisheries and seafood, tourism, and real estate, feel that their jobs depend on maintaining good environmental conditions in the local waterways, bays, and beaches. How much does your job depend on environmental conditions in the local waterways, bays, and beaches?

- O A great deal (1)
- O A fair amount (2)
- O Only a little (3)
- O Not at all (4)

To which of these industries is your occupation most closely related?

- ☐ Fisheries and Seafood (1)
- ☐ Coastal Tourism (2)
- ☐ Real Estate (3)
- ☐ ⊗Not applicable (4)

In general, do you think of yourself as...

- O Very liberal (1)
- O Somewhat liberal (2)
- O Moderate / Middle of the road (3)
- O Somewhat conservative (4)
- O Very conservative (5)
- O Other (6) _____
- O Prefer not to answer (7)

End of Block: 9 - Demographics