Summary Report

Bluefish and Spiny Dogfish Research Track Stock Assessment Peer Review

Members of the Research Track Peer Review Panel

Yan Jiao, Chair
Robin Cook
Paul Medley
Joe Powers

Prepared for the Northeast Regional Stock Assessment Workshop

National Marine Fisheries Service
National Oceanic and Atmospheric Administration Woods Hole, Massachusetts

Meeting dates: December 5-9, 2022
Report date: December 23, 2022

Table of Contents

1. Introduction

1.1 Background

The most recent full update stock assessment of the spiny dogfish (*Squalus acanthias*) was in 2018 and the method used is the Stochastic Estimator approach (NEFSC 2006, SARC 43). Based on that assessment, spiny dogfish was not overfished and overfishing was not occurring. There were observed survey SSB decreases in recent years, especially in 2021; the smoothed survey SSB was lower than that projected in 2018.

The most recent full update stock assessment for the bluefish (*Pomatomus saltatrix*) was in 2021, and it was based on an ASAP (age-structured assessment program, Legault and Restrepo 1999) model peer-reviewed in the 2015 benchmark assessment (SARC 60, NEFSC 2015). The MRIP (Marine Recreational Information Program) calibration resulted in an increase in the estimated recreational catch and caused scale changes in both biomass and reference points. The bluefish population was overfished, and overfishing was not occurring according to the 2021 assessment (NEFSC 2021).

Both spiny dogfish and bluefish were selected for research track peer review in 2021. Two working groups were created with staff from NEFSC, MAFMC, ASMFC, state agencies, and academia in 2021. The Terms of References (TORs) for the spiny dogfish and bluefish working groups are provided in Appendix 1. The Research Track assessments allow for evaluating and using new datasets, models, or stock structures. The stock assessments are expected to provide the basis for future management track assessments. The 2022 research track spiny dogfish assessment changed the base model from Stochastic Estimator to Stock Synthesis (SS3, Methot et al., 2020) model. The 2022 research track bluefish assessment changed the base model from ASAP to a Woods Hole Assessment Model (WHAM, Miller et al., 2016, 2018).

The Research Track Peer Review meeting met via WebEx from December 5-9, 2022, to review the most recent stock assessments for spiny dogfish and bluefish (see agenda in Appendix 2). The review committee includes Yan Jiao (MAFMC SSC and Virginia Polytechnic Institute and State University, Review Panel Chair) and three scientists affiliated with the Center for Independent Experts: Robin Cook, Paul Medley, and Joe Powers.

The peer review was assisted by Michele Traver (NEFSC's Stock Assessment Process Lead) and Russ Brown (Chief, NEFSC Population Dynamics Branch). Supporting documentation for the spiny dogfish stock assessment was prepared by the Spiny Dogfish Working Group (SDWG) and presentation of the assessment was made by Conor McManus (SDWG co-chair, NEFMC), Cami McCandless (SDWG co-chair, NEFSC), Kathy Sosebee (NEFSC), Dvora Hart (NEFSC), and Jui-Han Chang (NEFSC). SDWG members and Mid-Atlantic Fishery Management Council members and staff contributed substantially to the discussions on various topics. Toni Chute, Chris Legault, Brian Linton and Liz Brooks (all NEFSC) acted as rapporteurs throughout the meeting. Technical documents for the bluefish stock assessment were prepared by the Bluefish

Working Group (BWG) and presentations were made by Michael Celestino (NJDFW, Chair of BWG), Katie Drew (ASMFC), Abby Tyrell (NEFSC), Sarah Gaichas (NEFSC), Sam Truesdell (MADMF), Tony Wood (NEFSC) and Tim Miller (NEFSC). Larry Alade, Chuck Adams, Russ Brown, and Alex Hansell (all NEFSC) served as rapporteurs. A total of 52 individuals attended this Research Track Peer Review meeting, representing NEFSC, MAFMC, ASMFC, GARFO, MADMF, MDNR, NJDFW, NCDMF, NYSDEC, RIDMF, various academic institutions, non-governmental organizations, and fisheries stakeholder organizations (see Appendix 4 for materials provided and Appendix 5 for meeting attendees). Their contributions to the Bluefish and Spiny Dogfish Research Track Stock Assessment Peer Review process are gratefully acknowledged.

1.2 Review of Activities

Approximately one-two weeks before the meeting, the assessment documents and supporting materials were made available to the Peer Review Panel through an NEFSC website (https://appsnefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php). Before the meeting, the review panel members met with Michele Traver and Russell Brown to review and discuss the meeting agenda, reporting requirements, meeting logistics, and the overall process. The meeting opened on the morning of December 5 with welcoming remarks by Michele Traver and Russell Brown and Panel chair, Yan Jiao. Following introductions of the Review Panel, the SDWG and BWG, and other participants, the remainder of first two days were devoted to presentations of the spiny dogfish research track assessment and discussion of the first 8 TORs (Terms of Reference, see Appendix 1), and the third and fourth days were devoted to presentations of the bluefish research track assessment and discussion of the first 8 TORs. The final day of the meeting was dedicated to the review Panelists for report writing. The review panel Chair compiled and edited this Panel Summary Report with assistance (by correspondence) from the CIE Panelists before submission of the report to the NEFSC. Additionally, each CIE panelist will submit their separate reviewer's reports to the Center for Independent Experts.

The presentations during the meeting for each assessment followed the TORs, allowing the review panel to gain a deeper understanding of each assessment. The review panel asked each WG for additional information and clarifications to explore sensitivities and alternative model configurations, and the efforts by working group members to quickly generate those tables, figures and model runs were greatly appreciated. The tone of the meeting was collegial, and considerable time was devoted to facilitating dialog among Panelists, working group members, and MAFMC and ASMFC staff. The review panel was able to conduct a thorough review of both assessments.

The review panel was able to reach a consensus on both assessments. The review panel's evaluation of the working groups' 8 TORs is provided below. The review panel also provided future research recommendations. Since the last peer-reviewed assessments of each species, considerable research advancements have been made in each assessment. The assessments conducted by the SDWG and BWG were both new, and it was apparent that each working group devoted a significant amount of time and effort to data analysis and synthesis, model construction and fitting, diagnostic and evaluation of uncertainty, and report preparation.

2. Review of Spiny Dogfish

2.1 General Comments

The SDWG developed a new Stock Synthesis (SS3) model, which is different from the previous Stochastic Estimator approach. The WG constructed the base model based on the updated landings and discards, size frequencies of landings and discards, the life history processes studied in TOR1, and survey indices studied in TOR3. Sensitivity runs were used to explore assumptions in growth, mortality, SR relationship, time blocks for biological processes, and survey selectivity of NEFSC spring bottom trawl survey and surveys included. Many of the model runs had convergence problems when SS3 was used. The base model and the sensitivity runs did not fit the indices well, because of the strong influence from the length-frequency data. The review panel agreed that all the TORs were met, but some were met with reservations. The review panel recommended continuing to explore the sensitivity of the SS3 model parameterization and configuration before the following management stock assessment review.

2.2 Evaluation of the Terms of Reference for Spiny Dogfish

TOR 1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

The review panel agreed that this TOR had been met.

The SDWG explored the species distribution changes by estimating the center of gravity and effective area using a VAST (Vector Autoregressive Spatio-Temporal, Thorson 2019) model, but no significant changes were observed over time. The vessel trip report data over time was also explored to diagnose the spatial change of the spiny dogfish. The SDWG examined life history-related models such as SR (spawning stock and recruitment), maturity, and growth. SR with environmental factors did not improve model fitting, but a change in maturity was detected compared to the study done in 2010; the 50% maturity-at-length had significantly decreased over time. The SSB and recruitment estimates used for the SR were derived from the swept area survey data rather than estimated in the SS3 model. Exploration of the SR changes using the model estimated values were suggested because of the maturity at age and growth decreasing by the review panel but later realized that the SS3 base model run strongly fixed the SR relationship based on the observations from the swept area and with low variations. The new ageing data were rejected by the SDWG because of the high measurement error. Nevertheless, the spine ageing analysis and the tag-recapture analysis suggested a decreased growth curve after mid-2005 compared with the spiny dogfish sampled in the

1980s. These findings supported the changes in the biological models that were used in the SS3 base model configuration. The review panel agreed with these changes.

TOR 2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

The review panel agreed that this TOR had been met.

The commercial and recreational landings and discards were estimated by SDWG according to the standard method. There had been some pooling and borrowing for developing length and sex compositions of landings and discards but these were found appropriate by the review panel. CPUE from US commercial otter trawl was developed but not considered in the SS3 model because of its short time series and it was not sex-specific. Continued exploration of such resources to inform population trends is recommended. The landings and discards uncertainty were quantified and reported in the assessment report. However, these uncertainties were not accounted for in the SS3 model runs because of convergence issues. The review panel recommended that uncertainty should be considered in the future SS3 model configuration (TOR4).

TOR 3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

The review panel agreed that this TOR had been met.

Nine fishery-independent surveys within the stock boundaries were analyzed, and only the NEFSC bottom trawl survey index was recommended in the SS3 base model run. Indices from the NEAMAP inshore trawl survey, MADMF bottom trawl survey, and ME-NH inshore groundfish trawl survey were considered in the SS3 sensitivity run but their potential influence was hard to assess because of no data weighting and poor fits to the indices.

The VAST model was used to develop indices, and results were compared with the design-based estimates. When VAST was used to combine multiple surveys, the NEFSC bottom trawl dominated, which is not surprising given its wide spatial coverage compared with the other surveys. It was noticed that the VAST index has considerable differences from the survey design-based index. A sensitivity run was undertaken to test its influence in the SS3 model but it was found that the VAST index made little difference to the results, probably because of the strong influence from length-frequency data and the general lack of fitting to the indices for all the model runs.

Because of the importance of estimating cohort signals and given the pup/recruitment data seen in the NEFSC spring bottom trawl survey length frequency, which was not well explained in the SS3 model fit, the review panel recommended that a pup index may be considered in the future as a recruitment index.

TOR 4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

The review panel agreed that this TOR had been met, but had some reservations.

The previous stock assessment (2018) model used a Stochastic Estimator based on the swept area of the NEFSC spring bottom trawl survey. This had been used for management purposes since the 2010s.

The SDWG chose to move to the SS3 framework for this research track assessment primarily because it allowed for sex-specific analyses. The underlying population model was fitted to length data. Attempts were made to utilize age-reading data from the recent period. However, significant problems were detected. There were problems in the age-reading, especially for the younger fish. These analyses indicated a decreased growth rate in 2006-2014 compared to Nammack (1985) which formed the basis of earlier understanding of growth. For these reasons the SDWG rejected the age data and constructed time-blocks of biological parameters. These blocks were based on analyses of TOR1 and the age-length analyses presented under TOR4. These blocks along with the catch data, indices/fisheries and selectivities, and a stock-recruitment model were integrated into the SS3 framework. The review panel agreed with this general construction but noted some areas that might be revisited as the model moves to a management track.

There were some concerns about the SDWG recommended SS3 base model run because of the lack of meaningful model comparisons among sensitivity runs and the model run itself because of data weighting and fixed parameters. All the model runs use a 6-fleet model set which includes 2 landing fleets, 3 discarding fleets and 1 survey fleet: NEFSC spring bottom trawl survey. The landing and discard from the same fishery were separated into 2 fleets. The base run only included NEFSC spring bottom trawl index, and the model did not fit the survey index well (shown as a flat line and did not capture the historical decrease well). The base model only down-weighted the length frequency of the NEFSC survey fleet but not the landings and discards fleets, which was probably the reason for the poor fit to the survey index. The model did not capture the recruitment signal well, which was mainly from the NEFSC survey and the fit to the small-sized group length compositions was consistently poor. Seventeen sensitivity runs tested the influence of the growth model setup, the mortality assumption, the SR relationships, the biology time blocks, NEFSC bottom trawl selectivity time blocks, the model starting year, and the use of the survey indices. The sensitivity runs did not re-weight the data, so none fitted the survey indices well, and it was hard to diagnose the influence of each model change. According to the SDWG, many model runs did not converge, if the data weighting was turned on or manually fixed, which made diagnostics on alternative model runs not possible during the research track review. Additionally, the base model generated stock-recruitment data in the early years that appear anomalous (Fig 4.34 of the assessment report).

Despite these concerns, the review panel recommended the SS3 setup with suggestions for continued model re-configuration and evaluation before the management track review. For example, all the sensitivity runs should include data weighting, so that they are comparable with the base model run. Prior distributions rather than fixed parameter values may be used in future SR model configurations. The review panel felt there are many advantages with using an integrated model, such as the developed SS3 base model, compared with the empirical Stochastic Estimator approach. Other modeling frameworks or directly coding of alternative sub-models could be explored in the future if SS3 continues to have a convergence problem. The review panel also suggested that the ageing-length data collection and analysis should be continued considering its importance in both the assessment model, BRPs and projections (TORs 4, 5 and 6).

TOR 5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

The review panel agreed that this TOR had been met with some concerns.

The most recent stock assessment update for spiny dogfish was in 2018, in which the SSB_{MSY} (159,288 tonnes, based on SSB_{max} that results in max R in the Ricker SR model) and F_{MSY} (0.24/year, estimated as the F to reach stable SSB_{max}) were based on Rago and Sosebee (2010) and have not been updated since then. The data update based on the same approach resulted in a much larger SSB_{MSY} (445,349 tonnes) and a much lower F_{MSY} (0.03/year).

The SDWG found that the fish growth has been decreasing and the 50% maturity at length decreased after the early 2010s, and these were the major sources of uncertainty in estimating biological reference points when yield per recruitment and pup per recruitment were used. The SDWG conducted both landings per recruitment|F and pups per recruitment|F, and recommended using the spawning output (pups) of $SPR_{60\%}$ as the SSB_{target} or SSB_{MSY} proxy, and using the $F_{60\%}$ as the F_{limit} or F_{MSY} proxy. The recommendation of $SPR_{60\%}$ is based on the population responses to the fishing intensity between 2000 and 2019, which indicated that when F was lower than $F_{60\%}$ (during 2002-

2010) the population increased, and when F was higher than $SPR_{50\%}$ level, the population showed a decreased trend in spawning potential (during 2012-2019, shown as million pups). Based on $SPR_{60\%}$ reference points, the 2019 stock size was >½ $B_{SPR60\%}$ and the fishing mortality rate was < $F_{SPR60\%}$.

The review panel was concerned with the SR relationship from the assessment model because of the fixed parameters used and the anomalous data points mentioned under TOR4. It is recommended that the final selection of parameters be revisited in the context of additional years of data in the management track assessment. For these and other reasons the SDWG chose to utilize $SPR_{60\%}$ as a more appropriate surrogate for MSY than that generated directly from the stock-recruitment relationship. The review panel supports that decision for the next management track, but notes that there remains an inconsistency between $SPR_{60\%}$ and the underlying dynamics generated by the stock-recruitment model.

TOR 6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

The review panel agreed that this TOR had been met with some concerns.

The SDWG conducted 3-year (2022-2024) short-term projections under 4 F levels (F=0, $SPR_{70\%}$, $SPR_{60\%}$, and $SPR_{50\%}$) using the SS3 internal projection tool processes and uncertainty in recruitment and numbers-at-age. Fleet selectivity, maturity, natural mortality, SR relationship, and growth are the same as estimated from the 2012-2019 period from the SS3 model run. The 3-year projection showed a sharp decrease in 2020 but increased after that, likely due to the maturation of many females in the large 2009-2012 year classes. There are concerns from the review panel on the projection method related to the definitions of fleets in SS3. These are in effect “pseudo fleets” that separate catch components into landings and discard “fleets” while combining gears in different groups. As a result, forward projections with different F multipliers assume particular fleet selectivity and discard selection that may be unrealistic. Furthermore, given the artificial nature of the model fleets, it is unclear how these relate to management, making the interpretation of potential interventions problematic. The SPWG should explicitly address this issue when carrying out projections.

The NEFSC bottom trawl survey swept area estimated SSB_{2021} indicated a large decline; the projection did not capture this decline. Combining the concerns from the SS3 model runs not fitting the survey abundance index (indices), the review panel recommended that future diagnostics on both the assessment model and projection be evaluated between the research track review and management track review.

TOR 7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new

recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

The review panel agreed that this TOR had been met.

SDWG has made substantial progress on several of the research recommendations stemming from the 43rd SAW Stock Assessment Report (NEFSC 2006), MAFMC 2020-2024 Research Priorities (2019), and MAFMC SSC Research Recommendations in 2019. The review panel was highly impressed by the SDWG's progress in addressing these research recommendations, many of which were incorporated into the research track assessment model. The review panel recommended continued efforts on high-priority research topics from these lists. The SDWG also developed four new research recommendations, but thought the first one below is the most important:

- Consistently collect, process, and age spines of spiny dogfish to understand growth and support future age-based assessments.
- Continue exploration into the spatial distribution of spiny dogfish (e.g., off-shelf abundance).
- Further explore the sensitivity of the SS3 model parameterization and configuration.
- Conduct directed studies that estimate discard mortality rates for spiny dogfish by commercial and recreational harvesting gear type.
- Develop state-space models that can fit to length data.
- Investigate drivers in the decline in maturity over time.
- Continue developing the VAST models presented.
- Investigate datasets enumerating the abundance or diet of known spiny dogfish predators for insight into natural mortality rates.

The review panel agreed with the above new research recommendations and would like to emphasize the importance of consistently collecting and aging spiny dogfish. In addition, the review panel recommended exploring the use of other survey abundance indices and fishery catch rate that may inform either YOY or large spiny dogfish.

TOR 8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

The review panel agreed that this TOR had been met.

The SDWG considered four backup approaches, including a stochastic estimator, Depletion-corrected Average Catch (MacCall 2009), Depletion-based Stock Reduction Analysis (Dick and MacCall 2011) and Ismooth (Legault et al., 2022). The stochastic estimator approach is based on the swept area and is the current stock assessment approach. The SDWG recommended the stochastic estimator approach as the backup assessment approach if the research track review rejects the newly developed SS3 model run.

The review panel felt that the stochastic estimator was appropriate as a backup method. The SS3 model framework should be used subject to further consideration on data weighting and sensitivity analyses before application in management.

TOR 9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

N/A

3. Review of Bluefish

3.1 General Comments

The BWG developed a new WHAM model with the data and parameter configuration bridged from the last benchmark assessment in 2015, in which an ASAP model was used. The WG started with the ASAP 2021 MT (management track) run, RT (research track) continuity run, several model runs on new data or new data analysis methods, new M, new selectivity blocks and other parameter configurations, then moved to the new WHAM model setup. The report is well written, and the assessment is thorough and sound. The review panel unanimously agreed that all the TORs were met and accepted the WHAM model BF28W-m7 for use as the basis for bluefish stock assessment, and the WG's recommended BRPs and the estimation approach for BRPs and future population projections may be used for management advice.

3.2 Evaluation of the Terms of Reference for Bluefish

TOR 1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

The review panel agreed that this TOR had been met.

The BWG had extensively reviewed the existing research and synthesized the existing data on the social-economic, ecosystem, and life history. The BWG developed new analyses, including VAST species distribution changes, ecosystem indicators, VAST forage fish index, and applied their findings to the stock assessment model runs. The findings suggested that Gulf of Mexico catch data should be omitted, used seasonal length frequencies and length-weight relationships at a minimum, and used a seasonal-regional level of data where possible. The BWG also developed age-specific natural mortality, which was used in the recommended BF28W-m7 model run for management purposes. The BWG also addressed several previous research recommendations on life history, species distribution, and recruitment with environmental factors.

Although the forage fish index was not used in the recommended BF28W-m7 model run, its influence was tested in one model run as a covariate of MRIP catchability and was suggested for further research. The review panel also suggested that it may be considered in a catch rate standardization step before being used in the WHAM model to better understand the catchability changes of the MRIP CPUE.

The review panel thought that a tremendous amount of work was done to address this TOR and the work was extremely well done.

TOR 2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

The review panel agreed that this TOR had been met.

Several new research items were included for developing commercial and recreational landing and discard total number, total weight, length distribution, and release mortality. They are all scientifically sound according to the review panel.

The commercial discard was ignored in the past but included in this assessment, and a release mortality of 32% was used based on literature review. The recreational release mortality was updated from 15% to 9.4% based on literature reviews, including the most recent research. The recreational effort was recalibrated based on APAIS and FES; the recreational length frequency was calculated by accounting for the differences among seasons and regions, which was further used in developing seasonal catch-at-age to account for fish size variation among seasons and regions.

Discussions on whether the hook type changes in the past and whether discard mortality is size specific should be considered in future studies.

TOR 3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

The review panel agreed that this TOR had been met.

Several new research items were included for developing relative abundance indices, and age length keys (ALK). A Bayesian hierarchical model was used following Conn (2010) to develop a composite YOY index instead of using 6 separate YOY indices.

All the survey indices except the SEAMAP indices of ages 0 and 1, were developed using both a designed-based approach and a model-based approach (GLM framework used), and the BWG decided to use the result from the design-based approach. The trends of the survey indices are consistent between the two methods, and which method was used did not influence the output much according to the corresponding sensitivity runs. The BWG felt that the design-based approach would be easier to maintain consistency for future updates. The review panel found this a reasonable argument, although details were not discussed during the review.

The MRIP CPUE has been updated using a guild approach to select trips where a trip was considered a bluefish trip if it caught either bluefish or a species that was significantly positively associated with bluefish. This was from a previous research recommendation, and both the BWG and review panel believed that this was an important step forward in improving the recreational CPUE analysis.

A multinomial model was used to estimate probabilities in the age-length key, which avoided having to use an ad hoc “borrowing” method for empty cells when the sample sizes were small. The method was found reasonable for bluefish in this case. The review panel did realize that using multinomial ALK changed the scale of the population size in the stock assessment and suggested further evaluation of this method in the future.

TOR 4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

The review panel agreed that this TOR had been met.

The last benchmark stock assessment (2015) used an ASAP model and has been used for management purposes since then. It was last updated in 2021 in a management track assessment. The BWG moved the assessment from ASAP to a WHAM framework in this research track assessment review. The step-wise migration from ASAP to WHAM with the inclusion of new data or parameter configuration was clear, well thought out and reasonable. The BWG included a continuity ASAP run followed by a bridge model built with new data, smoothed age-length keys, age dependent M, new selectivity blocks, and other parameter configurations. This was then moved to the new WHAM model framework. This step-by-step approach, including results and diagnostics, helped the review process.

The most significant structural changes for the selected model run (BF28W-m7) included the process error on number at age in the model (NAA) and the use of fixed natural mortality varying with weight (Lorenzen M). The new model used the MRIP guild CPUE and multinomial ALK data. The model fit was generally good, with no serious retrospective patterns that needed a correction to the final results. The review panel felt that the BF28W-m7 was appropriate for management purposes to provide scientific advice.

Previous assessments were dependent on the MRIP CPUE. The new assessment is no longer as reliant on this index. The assessment is, however, now more sensitive to the PSIGNS index, the removal of which results in lower SSB and higher F. The PSIGNS index contained most of the information on the older fish abundance. However, given that this survey is limited in geographical coverage, some care is merited in interpreting the results.

The review panel noted that multinomial ALKs used to derive age composition data may have the effect of implying these data are more precise than is actually the case. A potential issue with pre-processing the data in this way may be to over-weight the composition data relative to the abundance data. A sensitivity run using ALKs applying the older "borrowing" method resulted in poorer model convergence but lower F and higher SSB.

The BWG investigated the use of a forage index to account for changes in survey catchability. The review panel saw this as an innovative approach that merits further analysis before being used in an assessment for management purposes.

The review panel suggested that the WHAM framework-based model may further consider processes such as natural mortality and fish spatial distribution changes based on what the BWG found in TOR1.

TOR 5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

The review panel agreed that this TOR had been met.

The BWG continued the use of $SSB_{35\%}$ as the SSB_{MSY} proxy and used the last five-year average Weight at Age (WAA) and selectivity for reference points estimation. The WG agreed that the literature (Rothschild et al. 2012; Thorson et al. 2012) supported the use of $F_{35\%}$ for bluefish and continued the use of $F_{35\%}$ as the F_{MSY} proxy. It was acknowledged that it was the generally accepted approach in this region to use SPR analysis for the reference points estimation. Both $F_{35\%}$ and $SSB_{35\%}$ were calculated internally in WHAM using average recruitment over the time series (1985-2021), and 5-year averages for fishery selectivity and weights-at-age for SSB per recruit calculations

$F_{35\%}$ = 0.248. $SSB_{35\%}$ was calculated using SPR at 35% (0.718), and the mean of the full time series of recruitment (127,924 tonnes) $SSB_{35\%}$ = 91,897 tonnes. Natural mortality and maturity were assumed constant over time in the model. Uncertainties of the BRP estimations were included in the assessment report shown as CIs and were calculated internally in WHAM. The Kobe plot showing the uncertainty envelope of current stock status relative to reference points is particularly useful.

The review panel discussed whether $SPR_{35\%}$ was the best proxy of MSY. Based on the plot of YPR|F and SPR|F, the $SPR_{35\%}$ is less than F_{max} but may be close to $F_{0.1}$ (not estimated in the report). Combined with the literature and the bluefish SPR and YPR analyses, the review panel agrees that $SPR_{35\%}$ is a reasonable proxy of SSB_{MSY}. Future exploration of SR relationship and MSY reference points by combining YPR and stock recruitment relationship may be explored (Shepherd 1982).

TOR 6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

The review panel agreed that this TOR had been met.

The BWG conducted short-term projections under 3 F levels in WHAM, which incorporated auto-regressive processes and uncertainty in recruitment and numbers-at-age. The projections used the entire time series of recruitment (1985-2021), 5-year averages for natural mortality (assumed age varying but constant cross years), maturity (constant), fishery selectivity, and weights-at-age. The life history study from TOR1 found that the maturity changes over time are limited, and the changes in weight-at-length are trivial.

The projection algorithm in dealing with multi-fleet fishery matched the operational model setup. The review panel found it reasonable for projections under alternative fishing mortality level based on the council's risk policy and appropriate for management advice identification.

TOR 7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 1 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

The review panel agreed that this TOR had been met.

BWG made considerable progress on several research recommendations stemming from the 2015 assessment (SAW/SARC 60) and MAFMC SSC recommendations in 2015 and 2021. The review panel was highly impressed by the BWG's progress in addressing these research recommendations, and many of them were incorporated into the research track assessment model. The review panel recommended continued efforts on high priority research topics from these lists. The BWG also developed four new research recommendations:

- Expand collection of recreational release length frequency data. The recreational release length frequency spatially stratified; borrow if n < 30.
- Continue coastwide collection of length and age samples from fishery dependent and fishery independent sources
- Continued development and refinement of forage fish index; incorporate into the base model for management
- Initiate fishery-independent or fishery dependent sampling programs to provide information on larger, older bluefish

The review panel agreed with the above new research recommendations and suggested more be added to the list:

- Continue exploring the appropriate application of the WHAM model, including alternative ALK estimation.
- Explore the reasons for bimodal length frequency observed in bluefish harvest.
- Continue the forage fish index study and explore the potential application in catch rate standardization to remove the forage fish influence on catchability
- Explore WHAM process error in simulating key parameter changes caused by climate or environmental changes, such as M and fish spatial distribution changes over time.

TOR 8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

The review panel agreed that this TOR had been met.

The logic for the selection of the Ismooth (Legault et al. 2022) was clear. The BWG indicated this approach worked as well as any other Index Based method. The BWG also addressed reasons for not selecting other candidate approaches, including swept area, Depletion-corrected Average Catch (MacCall 2009), and Depletion-based Stock Reduction Analysis (Dick and MacCall 2011).

The review panel felt that Ismooth method is appropriate as a backup method even though it is not needed in this case. The review panel recommended the WHAM model run BF28W-m7 for management purposes.

TOR 9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

N/A

4. Supporting Materials for Research Track Peer Review

4.1 Spiny Dogfish

<u>Assessment Report</u>

"Spiny_Dogfish_SAW_SARC_2022_FINAL.pdf" = Main assessment document

<u>Background</u>

"Read Me.pdf" – Document and materials guide, as well as a repository of any report revisions.

"plots_v3.6.2_1.5_fnum_a12.zip" - This zip file contains the base case model figures and files produced from SS3. Within this zip file, there is a file labeled '_SS_output.html', which is an html that allows for viewing SS3 produced plots and results in an organized fashion (i.e. by various data type or model result).

<u>Working Papers</u>

Anstead K. 2022a. Natural mortality estimates for spiny dogfish.

Anstead K. 2022b. Two data poor methods applied to spiny dogfish.

Chang J-H, Hart D and McManus MC. 2022. Stock synthesis for Atlantic spiny dogfish.

Hansell A and McManus C. 2022. Spatio-temporal dynamics of spiny dogfish (*Squalus acanthias*) in US waters of the northwest Atlantic.

Hart DR, and Chang J-H. 2022. Per recruit modeling and reference points for spiny dogfish.

Hart DR, and Sosebee K. 2022. Length/Weight/Fecundity relationships for Atlantic spiny dogfish.

Jones AW. 2022. Exploring vessel trip report and observer based fishery information for spiny dogfish.

Jones AW, Didden JT, McManus MC, and Mercer AJ. 2022. Exploring commercial CPUE indices for the spiny dogfish in the northeast U.S.

McCandless C. 2022. Preliminary spiny dogfish movements and growth estimates from NEFSC mark recapture data.

McManus MC, Sosebee K, and Rago P. 2022. Biological Reference Points for Spiny Dogfish: Revisiting Rago and Sosebee (2010).

Neiland JL and McElroy WD. 2022. NEFSC Gulf of Maine Bottom Longline Survey Data and Analyses for Spiny Dogfish

Passerotti MS, and McCandless CT. 2022. Updated age and growth estimates for spiny dogfish Squalus acanthias.

Sosebee KA. 2022a. Maturity of spiny dogfish in US waters from 1998-2021.

Sosebee KA. 2022b. Spiny dogfish catch summary and derivation of catch at length and sex.

4.2 Bluefish

<u>Assessment Report</u>

Bluefish_SAW_SARC_2022_FINAL.pdf = Main assessment document

<u>Background</u>

readme.docx – document guide and repository of any report revisions.

Background Documents
NEFSC. 2015. 60th Northeast Regional Stock Assessment Workshop (60th SAW) assessment report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-08; Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026 http://doi.org/10.7289/V5W37T9T
Stock, B.C., and Miller, T.J. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research, 240: 105967. Doi: 10.1016/j.fishres.2021.105967
Legault, C.M. and Restrepo, V. 1999. A flexible forward Age-Structured Assessment Program. ICCAT Coll. Vol. Sci. Pap. 49.
Legault, C. M. 2012. User manual for ASAP 3. 22 p.
Ng, E.L., Deroba, J.J., Essington, T.E., Grüss, A., Smith B.E., and Thorson, J.T. 2021. Predator stomach contents can provide accurate indices of prey biomass. ICES Journal of Marine Science 78(3):1146–1159.
Thorson, J.T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210:143–161.

Working Papers

Tyrell et al. 2022. Bluefish Ecosystem and Socioeconomic Profile.
Valenti 2022a. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from American Littoral Society Fish Tagging Data
Tyrell 2022. Bluefish VAST Index Exploration.
Gaichas et al. 2022. Vector Autoregressive Spatio-Temporal (VAST) modeling of piscivore stomach contents, 1985-2021.
Truesdell et al. 2022. Life History Analyses for Bluefish.
Tyrell and Truesdell 2022. Natural mortality of bluefish.
Celestino et al. 2022a. Index of abundance exploration and development by the Bluefish Working Group's Fishery Independent Data Group.
Wood 2022a. TOR 2: Commercial and Recreational Data Collection and Analysis.
Drew 2022a. Recreational Data Changes for Bluefish, 2012-2021.
Drew 2022b. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from MRIP Data.
Valenti 2022b. Catch-and-Release Recreational Angling Mortality of Bluefish (*Pomatomus saltatrix*): Updated Analysis for 2022
Drew 2022c. Development of the Composite YOY Index for Bluefish.
Drew 2022d. A Fishery-dependent CPUE index for bluefish derived from MRIP data.
Celestino et al. 2022b. Development of Bluefish Age-Length Keys.
Wood 2022b. Bluefish Model Bridge-Building in ASAP.
Wood 2022c. ASAP diagnostic plots.
Wood 2022d. WHAM diagnostic plots.
Truesdell 2022. Alternative assessment plan.

Other References

Conn, PB. 2010. Hierarchical analysis of multiple noisy abundance indices. Canadian Journal of Fisheries and Aquatic Sciences 67:108-120.

Dick, E.J. and MacCall, A.D. 2011. Depletion-based stock reduction analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110(2): 331-341.

Legault C.M., and Restrepo V.R. 1998. A flexible forward age-structured assessment program. ICCAT. Col. Vol. Sci. Pap. 49:246-253.

Legault, C.M., Wiedenmann, J., Deroba, J.J., Fay, G., Miller, T.J., Brooks, E.N., Bell, R.J., Langan, J.A., Cournane, J.M., Jones, A.W., Muffley, B. 2022. Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. Canadian Journal of Fisheries and Aquatic Sciences e-First https://doi.org/10.1139/cjfas-2022-0045

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

Methot, R.D., Jr., Wetzel, C.R. Taylor, I.G. and Doering, K. 2020. Stock Synthesis User Manual Version 3.30.15. U.S. Department of Commerce, NOAA Processed Report NMFS-NWFSC-PR-2020-05. https://doi.org/10.25923/5wpn-qt71

Miller, T.J., Hare, J.A., and Alade, L.A. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to Southern New England yellowtail flounder. Canadian Journal of Fisheries and Aquatic Sciences 73(8): 1261-1270. Do: 10.1139/cifas-2015-0339

Miller, T.J., O'Brien, L., and Fratantoni, P.S. 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of Georges Bank Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences 75(12): 2159-2171. Doi:10.1139/cifas-2017-0124

NEFSC. 2021.Atlantic Bluefish Operational Assessment for 2021. Updated Through 2019. https://apps-st.fisheries.noaa.gov/stocksmart?stockname=Bluefish%20-%20Atlantic%20Coast&stockid=10388

Northeast Fisheries Science Center (NEFSC). 2006. 43rd Northeast Regional Stock Assessment Workshop (43rd SAW): 43rd SAW assessment report. Northeast Fisheries Science Center Reference Document 06-25; 400p.

Northeast Fisheries Science Center (NEFSC). 2015. 60th Northeast Regional Stock Assessment Workshop (60th SAW) Assessment Report. US Dept Commerce, Northeast Fish Sci Cent Ref Doc. 15-08; 870 p.

Northeast Fisheries Science Center (NEFSC). 2019. Operational Assessment of the Black Sea Bass, Scup, Bluefish, and Monkfish Stocks, Updated Through 2018. US Dept Commerce, Northeast Fish Sci Cent Ref Doc. 20-01; 164 p.

Rago, P.J., Sosebee, K.A. 2010. Biological Reference Points for Spiny Dogfish. Northeast Fish Sci Cent Ref Doc. 10-06; 52 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at: http://www.nefsc.noaa.gov/nefsc/publications/

Rothschild, B.J., Jiao, Y., and Hyun, S.Y. 2012. Simulation study of biological reference points for Summer Flounder. Transactions of the American Fisheries Society 141(2):426-436

Thorson, J.T., Cope, J.M., Branch, T.A., and Jensen, O.P. 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.

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Attachment 1: Performance Work Statement (PWS)

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Bluefish and Spiny dogfish Research Track Peer Review

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation’s marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[1]. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

Scope

The Research Track Peer Review meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The research track peer review is the cornerstone of the Northeast Region Coordinating Council stock assessment process, which includes assessment development, and report preparation (which is done by Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees), assessment peer review (by the peer review panel), public presentations, and document publication. The results of this peer review will be incorporated into future management track assessments, which serve as the basis for developing fishery management recommendations.

The purpose of this meeting will be to provide an external peer review of the Spiny dogfish and Bluefish stock. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: **Appendix 1**: TORs for the research track, which are the responsibility of the analysts; **Appendix 2:** a draft meeting agenda; **Appendix 3:** Individual Independent Review Report Requirements; and **Appendix 4:** Peer Reviewer Summary Report Requirements.

[1] http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf

Requirements
NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the PWS, OMB Guidelines, and the TORs below. All TORs must be addressed in each reviewer's report. The reviewers shall have working knowledge and recent experience in the use and application of index-based, age-based, and state-space stock assessment models, including familiarity with retrospective patterns and how catch advice is provided from stock assessment models. In addition, knowledge and experience with simulation analyses is required.

Tasks for Reviewers

- Review the background materials and reports prior to the review meeting
  - Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.
- Attend and participate in the panel review meeting
  - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the Peer Review Panel (co)Chair with contributions to the Peer Reviewer Summary Report
- Deliver individual Independent Reviewer Reports to the Government according to the specified milestone dates
- This report should explain whether each research track Term of Reference was or was not completed successfully during the peer review meeting, using the criteria specified below in the "Tasks for Peer Review Panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

- During the peer review meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the Peer Review Panel chair shall identify or facilitate agreement among the reviewers for each research track TOR.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for Peer Review Panel chair and reviewers combined:
Review the Report of Spiny Dogfish and Bluefish Research Track Working Group.

The Peer Review Panel Chair, with the assistance from the reviewers, will write the Peer Reviewer Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each research track Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the peer review meeting. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such opinions.

The chair's objective during this Peer Reviewer Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express their opinion on each research track Term of Reference, either as part of the group opinion, or as a separate minority opinion. The Peer Reviewer Summary Report will not be submitted, reviewed, or approved by the Contractor.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, country of current residence, dual citizenship (yes, no), passport number, country of passport, travel dates.) to the NEFSC Assessment Process Lead for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Place of Performance
The place of performance shall be held remotely, via WebEx video conferencing.

Period of Performance
The period of performance shall be from the time of award through December 23, 2022. The Chair's duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Approximately November 21, 2022	NOAA/NMFS provides the pre-review documents to the reviewer panel (Reviewers and Chair)
December 5-9, 2022	Panel peer review meeting
December 23, 2022 (approximately 2 weeks later)	Chair submits a draft summary peer review report to NEFSC and NEFMC
January 6, 2023 (within 2 weeks of receiving draft reports)	Chair submits final reports to the Government (to NEFSC and NEFMC)

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
No travel is necessary, as this meeting is being held remotely.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
Michele Traver, NEFSC Assessment Process Lead
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov

Appendix 1. Generic Research Track Terms of Reference

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

*Any additional TORs will require review and approval.

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Research Track TORs:

General Clarification of Terms that may be used in the Research Track Terms of Reference

Guidance to Peer Review Panels about "Number of Models to include in the Peer Reviewer Report":

In general, for any TOR in which one or more models are explored by the Working Group, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the Working Group and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Research Track Working Group:

Anyone participating in peer review meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Appendix 2. Peer Review Meeting Agenda

Spiny Dogfish/Bluefish Research Track Assessment Peer Review Meeting

December 5-9, 2022

WebEx link: https://noaanmfs-meets.webex.com/noaanmfs-meets/j.php?MTID=m537714866febfc8ede459d55b0482239

Meeting number (access code): 2764 137 9769

Meeting password:

AhFMe8W3DS5 Phone: +1-415-

527-5035 US Toll

AGENDA* (v. 11/17/2022)

**All times are approximate, and may be changed at the discretion of the Peer Review Panel chair. The meeting is open to the public; however, during the Report Writing sessions we ask that the public refrain from engaging in discussion with the Peer Review Panel.*

Monday, December 5, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics Introductions/Agenda/ Conduct of Meeting	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief Yan Jiao, Panel Chair	
9:15 a.m. - 9:30 a.m.	Introduction/Executive Summary	Conor McManus (WG co-chair) (Spiny Dogfish)	
9:30 a.m. - 10 a.m.	Term of Reference (TOR) #1	Conor McManus	Ecosystem Data
10 a.m. - 10:30 a.m.	TOR #3	Cami McCandless (WG co-chair)	Survey Data
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2	Kathy Sosebee	Catch Data
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m	Lunch		

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Time	Topic	Presenter(s)	Notes
1:30 p.m. - 2:30 p.m.	TOR #4	Dvora Hart Jui-Han Chang	Models
2:30 p.m. - 3 p.m.	End of Day Wrap-up/ Discussion/Summary	Review Panel	
3 p.m. - 3:15 p.m.	Public Comment	Public	
3:15 p.m.	Adjourn		

Tuesday, December 6, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:05 a.m.	Welcome/Logistics Introductions/Agenda	Michele Traver, Assessment Process Lead Yan Jiao, Panel Chair	
9:05 a.m. - 10:30 a.m.	TOR #4 cont.	Dvora Hart Jui-Han Chang	Models
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 12 p.m.	TORs #5 and #6	Dvora Hart Jui-Han Chang	Reference Points Projections
12 p.m. - 12:30 p.m.	Discussion/Summary	Review Panel	
12:30 p.m. - 12:45 p.m.	Public Comment	Public	
12:45 p.m. - 1:45 p.m	Lunch		
1:45 p.m. - 2:30 p.m.	TOR #8	Dvora Hart	Alternative Assessment Approach
2:30 p.m. - 2:45 p.m.	TOR #7	Conor McManus	Research Recommendations
2:45 p.m. - 3:15 p.m.	End of Day Wrap-up/ Discussion/Summary	Review Panel	
3:15 p.m. - 3:30 p.m.	Public Comment	Public	
3:30 p.m.	Adjourn		

Wednesday, December 7, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:05 a.m.	Welcome/Logistics Introductions/Agenda	Michele Traver, Assessment Process Lead	

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Time	Topic	Presenter(s)	Notes
		Yan Jiao, Panel Chair	
9:05 a.m. - 9:30 a.m.	Introduction/Executive Summary	Mike Celestino (WG chair) (Bluefish)	
9:30 a.m. - 10:30 a.m.	TOR #2	Mike Celestino Katie Drew	Catch Data
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #3	Mike Celestino Katie Drew	Survey Data
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m	Lunch		
1:30 p.m. - 3 p.m.	TOR #1	Abby Tyrell Sarah Gaichas	Ecosystem Data
3 p.m. - 3:30 p.m.	End of Day Wrap-up/ Discussion/Summary	Review Panel	
3:30 p.m. - 3:45 p.m.	Public Comment	Public	
3:45 p.m.	Adjourn		

Thursday, December 8, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:05 a.m.	Welcome/Logistics Introductions/Agenda	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief Yan Jiao, Panel Chair	
9:05 a.m. - 10:30 a.m.	TOR #4	Tony Wood Tim Miller	Models
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:15 a.m.	TOR #4 cont.	Tony Wood Tim Miller	Models
11:15 a.m. - 11:45 a.m.	Discussion/Summary	Review Panel	
11:45 a.m. - 12 p.m.	Public Comment	Public	

Time	Topic	Presenter(s)	Notes
12 p.m. - 1 p.m	Lunch		
1 p.m. - 2:30 p.m.	TORs #5, #6, #8, and #7	Tony Wood Sam Truesdell Mike Celestino	Reference Points, Projections, Alternative Assessment Approach Research Recommendations
2:30 p.m. - 3 p.m.	Meeting Wrap-up/ Discussion/Summary	Review Panel	
3 p.m. - 3:15 p.m.	Public Comment	Public	
3:15 p.m.	Adjourn		

Friday, December 9, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 5 p.m.	Report Writing	Review Panel	

New England Fishery Management Council
50 WATER STREET | NEWBURYPORT, MASSACHUSETTS 01950 | PHONE 978 465 0492 | FAX 978 465 3116
Eric Reid, *Chair* | Thomas A. Nies, *Executive Director*

Appendix 3. Individual Independent Peer Reviewer Report Requirements

1. The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the Peer Reviewer Summary Report.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Peer Reviewer Summary Report that they believe might require further clarification.

   d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Performance Work Statement
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Appendix 4. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the Research Track Peer Review Panel chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the peer review meeting. Following the introduction, for each assessment /research topic reviewed, the report should address whether or not each Term of Reference of the Research Track Working Group was completed successfully. For each Term of Reference, the Peer Reviewer Summary Report should state why that Term of Reference was or was not completed successfully.

   To make this determination, the peer review panel chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and peer review panel chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

   The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the peer review meeting, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.

   The report shall also include as a separate appendix the assessment Terms of Reference used for the peer review meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 5 - Meeting attendees at the Spiny Dogfish/Bluefish Research Track Peer Review
December 5-9, 2022

GARFO - Greater Atlantic Regional Fisheries Office
MADMF - Massachusetts Division of Marine Fisheries
MAFMC - Mid Atlantic Fisheries Management Council
MDNR - Maryland Department of Natural Resources
NEFMC - New England Fisheries Management Council
NEFSC - Northeast Fisheries Science Center
NCDMF - North Carolina Division of Marine Fisheries
NJFW - New Jersey Fish and Wildlife
NYSDEC - New York State Department of Environmental Conservation
RIDEM - Rhode Island Department of Environmental Management
~~~~~~
Yan Jiao - Chair
Joe Powers - CIE Panel
Robin Cook - CIE Panel
Paul Medley - CIE Panel

Russ Brown - NEFSC, Population Dynamics Branch Chief
Michele Traver - NEFSC, Assessment Process Lead

Abby Tyrell - NEFSC
Alan Bianchi - NCDMF
Alex Dunn - NEFSC
Alex Hansell - NEFSC
Alexei Sharov - MDNR
Andy Jones - NEFSC
Anna Mercer - NEFSC
Brandon Muffley - MAFMC staff
Brian Linton - NEFSC
Cami McCandless - NEFSC
Charles Adams - NEFSC
Charles Perretti - NEFSC
Chris Legault - NEFSC
Conor McManus - RIDEM
Cynthia Ferrio - GARFO
Dave McElroy - NEFSC
Dvora Hart - NEFSC
Eric Robillard - NEFSC
Greg DiDomenico - Lunn's Fisheries
Hannah Hart - MAFMC staff
James Fletcher - United National Fishermen's Association
Jason Didden - MAFMC staff
~~~~~~

John Maniscalco - NYSDEC
Jose Montanez - MAFMC staff
Jui-Han Chang - NEFSC
Julie Nieland - NEFSC
Karson Cisneros - MAFMC staff
Kathy Sosebee - NEFSC
Katie Drew - ASMFC staff
Kiersten Curti - NEFSC
Kristen Anstead - ASMFC
Larry Alade - NEFSC
Liz Brooks - NEFSC
Mark Terceiro - NEFSC
Mike Celestino - NJFW
Michelle Passerotti - NEFSC
Paul Nitschke - NEFSC
Rich McBride - NEFSC
Ricky Tabandera - NEFSC
Sam Truesdell - MADMF
Samantha Werner - NEFSC
Sarah Gaichas - NEFSC
Scott Large - NEFSC
Tim Miller - NEFSC
Toni Chute - NEFSC
Tony Wood - NEFSC

Research Track Assessment of Northwest Atlantic

Spiny Dogfish

Spiny Dogfish Research Track Working Group

November 23, 2022

PARTICIPANTS

Working Group

NAME	AFFILIATION
Kristen Anstead	ASMFC
Jui-Han Chang	NEFSC
Jason Didden	MAFMC
Alex Hansell	NEFSC
Dvora Hart	NEFSC
Cami McCandless, Co-Chair	NEFSC
Conor McManus, Co-Chair	RIDEM
Halie O'Farrell	FLFWC
Kathy Sosebee	NEFSC
Ben VanDine	NEFSC

Additional Contributors and Meeting Participants

NAME	AFFILIATION
Alan Bianchi	NCDEQ
Jessica Blaylock	NEFSC
Russ Brown	NEFSC
Alex Dunn	NEFSC
James Fletcher	UNFA
Cynthia Ferrio	GARFO
Catherine Foley	NEFSC
Vladlena Gertseva	NWFSC
Andy Jones	NEFSC
Kristof Ketch	NEFSC
Anna Mercer	NEFSC
Dave McElroy	NEFSC
Julie Neiland	NEFSC
Michelle Passerotti	NEFSC
Paul Rago	NEFSC (retired)
Ian Taylor	NWFSC
Michele Traver	NEFSC
John Whiteside	SFA

TABLE OF CONTENTS

EXECUTIVE SUMMARY

A research track assessment for spiny dogfish was planned for peer review in 2022, with several terms of reference (TORs) established to be addressed. This is the Spiny Dogfish Working Group's report to fulfill the TORs.

Terms of Reference (TOR) 1: "*Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.*"

Ecosystem and climate influences on the Northwest Atlantic spiny dogfish stock (simply "spiny dogfish" hereafter) were assessed by the Working Group in the context of their distribution and life history processes. The literature on spiny dogfish distribution was reviewed to provide context on its historical range, migration patterns, and perceived stock structure. Spatial distribution of the species was described specifically for within the Northeast U.S. Continental Shelf, and the geographic, climate, and environmental variables that have been known to influence spiny dogfish. To assess how climate has influenced the stock's abundance and distribution, a Vector Autoregressive Spatiotemporal (VAST) model was developed from the Northeast Fisheries Science Center (NEFSC) spring bottom trawl survey to calculate the center of gravity and effective area occupied for male and female dogfish. Largely, these metrics suggested that the annual distribution of dogfish has not changed significantly over time. Temperature and depth were explored as covariates in the VAST model, as they were the most common variables associated with spiny dogfish abundance and distribution from the literature. Results indicated that depth was the only significant factor in predicting occurrence and abundance.

The Working Group also discussed the environment and potential effects on life history characteristics: recruitment, growth, maturity, and diet. The Working Group explored the correlation between environmental conditions (e.g., spring bottom temperature, the North Atlantic Oscillation) on recruitment and recruits per spawner indices from the NEFSC spring bottom trawl survey, with little correspondence. Temperature was also evaluated in the context of a stock-recruit relationship, which indicated no statistical improvement over a non-environmentally explicit relationship. While environmental and climate influences on growth may be occurring, the lack of time series growth information prevented the Working Group from conducting related formal analyses. Updated maturity time-series data indicated a decline in maturity over time, but several causes are possible, including either harvest or environmental forcings. As such, better understanding the drivers in the declining maturity over time is considered a research recommendation.

TOR 2: Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

Commercial and recreational landings and discards are estimated over time, with methods for deriving them presented. Commercial landings increased rapidly from the late 1960s to 1974, with substantial spiny dogfish harvest by foreign trawling fleets beginning in 1966. After 1978, landings by foreign fleets were curtailed, and landings by U.S. and Canadian vessels increased. The U.S. commercial fishery intensified in 1990, and landings were reduced in the 2000s due to restrictions imposed by federal and interstate fisheries management plans. When the stock was declared rebuilt in 2009, the allowed biological catch, trip limits and landings increased. Otter trawl and gill nets have been the primary U.S. commercial gears used to harvest spiny dogfish. Estimation of discards was uncertain prior to establishment of the at-sea observer program in 1989, which informed the starting year of the assessment model. There is some uncertainty in landings and discards for each fleet’s size and sex composition information based on the available data and thus associated assumptions made to produce catch information for the assessment model. Catch per unit effort indices were developed for the U.S. commercial otter trawl fleet to assess prospective correspondence to fisheries independent surveys.

TOR 3: Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

The Working Group evaluated several fisheries-independent surveys within the stock boundaries to inform modeling efforts of TOR 4: NEFSC Bottom Trawl Surveys, NEFSC

Bottom Long Line Survey, Northeast Area Monitoring and Assessment Program (NEAMAP) Inshore Trawl Survey, Massachusetts Division of Marine Fisheries (MADMF) Bottom Trawl Survey, Atlantic States Marine Fisheries Commission (ASMFC) Shrimp Survey, Rhode Island Coastal Trawl Surveys, the Maine-New Hampshire (ME-NH) Inshore Groundfish Trawl Survey, and Canadian Bottom Trawl Surveys. Where available, indices were evaluated for both male and female spiny dogfish by season. Concerns as to whether surveys that only sampled a portion of the stock unit adequately track temporal population changes led the Working Group to only use the NEFSC spring bottom trawl survey for modeling purposes. Of the available data, this survey best samples the entirety of the stock. Fall indices are not optimal for assessing annual changes because substantial portions of the stock are outside the survey domain during that season.

VAST models were developed to integrate multiple surveys' information and produce a single index and associated length composition for each sex in a given season. VAST models for this exercise included the NEFSC Bottom Trawl Survey, NEAMAP Inshore Trawl Survey, MADMF Bottom Trawl Survey, and ME-NH Inshore Groundfish Trawl Survey. A comparison of NEFSC spring bottom trawl relative abundance indices and the VAST model spring indices indicated similar patterns over time. Abundance indices produced by VAST were developed for spiny dogfish by season and sex for use in the assessment model as a sensitivity run. However, VAST model fitting proved challenging for the length composition data and the Working Group was unable to get a converged model at the resolution of the length bins used by the assessment model. Model sensitivity analyses included testing the NEFSC fall bottom trawl survey indices, NEFSC spring and fall bottom long line survey indices, as well as the VAST spring index with interpolated length compositions.

TOR 4: Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

Stock Synthesis 3 (SS3) was chosen as the primary assessment tool, due to its ability to model sexes separately, and to accommodate length-based approaches. The SS3 base case model ran from 1989-2019 because the sea sampling data used to estimate discards was not available prior to 1989. Input data to the model included the NEFSC spring trawl survey, landings, discards, and length compositions for all of these data sources. Growth was modeled as von Bertalanffy, using the parameters estimated by Nammack et al. (1985), except that L_∞ for 2012-2019 was estimated within the model; the estimated female L_∞ for that period (89.24 cm) is considerably smaller than that used for 1989-2011 (100.50 cm). Natural mortality was taken to decline with age (Lorenzen 1996), and was assumed to average 0.102 over the 50 year potential lifespan of Atlantic spiny dogfish. The survival spawner-recruitment relationship was used, which was specifically designed for low fecundity species such as spiny dogfish (Taylor et al. 2013). Alternative stock-recruit models (Beverton-Holt and Ricker) were tested in SS3, but output from these runs appeared to be much less credible than that from the survival spawner-recruitment relationship.

The base case SS3 run showed declines in spawning output from 1989 to 1997; these quantities increased until 2012, then declined again. The estimated base case spawning output trends reasonably matched survey trends during 2000-2019 and exhibited almost no retrospective pattern (Mohn's $\rho = 0.06$). However, the base case estimated smaller declines in spawning output during 1989-1997 than those observed in the NEFSC spring trawl survey. Estimated female fishing mortality (numbers based, age 12+) peaked in 1992 at about 0.17, declined to less than 0.025 between 2002-2010, and averaged about 0.033 during the most recent period (2014-19).

The SS3 base case run was compared to the output from the Stochastic Estimator, the model used in previous spiny dogfish assessments. The Stochastic Estimator is based on swept area calculations under the assumption that the survey trawl efficiency is one, and uses bootstrapping to quantify the uncertainties. The SS3 model generally estimated somewhat higher biomass and spawning output and lower fishing mortality than the Stochastic Estimator because it estimated a slightly lower survey efficiency ($q = 0.83$). The Stochastic Estimator estimated much higher F and a larger decline in female biomass and spawning output in the early portion of the time series.

TOR 5: Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

Per recruit calculations indicate that both yield-per-recruit (YPR) and pups-per-recruit (PPR) calculations are highly sensitive to growth assumptions. Maximum YPR occurred around $F = 0.15$, but using the estimates of L_∞ from the most recent period (89.24 cm for females), fishing above $F = 0.03$ produced less than two pups per recruit, and thus was unsustainable. The Working Group evaluated three SS3 estimated spawners-per-recruit (SPR) reference points: SPR50%, SPR60% and SPR70%. The fishing mortality associated with SPR50% (0.037) would produce less than two PPR. Furthermore, mean fishing mortality was below this value during 2013-2019, but nonetheless, female biomass and spawning output substantially declined during this period. By contrast, these quantities increased when fishing mortality was below $F = 0.025$, the fishing mortality associated with SPR60%, and decreased when $F > 0.025$ during the most recent period. For these reasons, the Working Group recommended adopting the SPR60% reference points: a spawning output target of 370.8 million pups and $F = 0.025$. This spawning output target corresponds to a considerably higher spawning biomass than previous reference points (SSB_{MAX} = 159,288 or 189,553 mt). However, reestimation of the previous reference points using updated data and parameters produced estimates similar to SPR60% (SSB_{MAX} = 445,349 mt and $F = 0.03$, McManus et al. 2022).

TOR 6: Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

The Working Group used the projection tool internal to SS3 for this assessment. The continuity of both the assessment model and projections being conducted with the same software allowed for effective and efficient application of the projection tool. Short-term projections were conducted (2020-2022) under four different fishing mortality rates: one under zero harvest and at $F = 0.017$, 0.025, and 0.037, corresponding to the SPR reference points SPR70%, SPR60%, and SPR50% respectively. Projections indicated a decline in spawning output from 2019 to 2020, and then increases in spawning output under all four alternatives, likely due to maturation of many females in the large 2009-2012 year classes.

TOR 7: "Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations."

The Working Group reviewed the research recommendations presented in the last benchmark stock assessment for spiny dogfish (43rd SAW Stock Assessment Report, NEFSC 2006), and those most recent from the Mid-Atlantic Fisheries Management Council and its Scientific and Statistical Committee. Individual responses were provided to each recommendation on how the work conducted during this assessment addressed them. New research recommendations were also put forth by the Working Group; the highest priority recommendation is in regard for consistent ageing analyses. Movement from data-limited approaches to more sophisticated models often depends on available age or growth information. Ageing programs should be established to allow for the continuous inclusion of such data and better inform growth in the assessment model, which can have significant impacts on model performance. Age samples should be collected across the spectrum of significant variables: by sex, across the size spectrum, by season, and over various areas of the stock bounds.

TOR 8: Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment. A backup assessment approach is required to be in place as a hedge against a scenario where the primary catch-at-age model is not suitable for providing management advice.

The Working Group evaluated several backup approaches, including the Stochastic Estimator, Depletion-Based Stock Reduction Analysis, Depletion-Corrected Average Catch, and the index-based method Ismooth. Each method uses various data streams (e.g., fisheries-independent indices, landings or catch information, life history parameters) to provide inferences on population size and/or stock status. Of the methods reviewed, the Working Group recommended the Stochastic Estimator be used as the backup approach to providing scientific advice to managers if the preferred SS3 assessment model approach does not pass peer review or if SS3 is rejected in a future management track assessment.

WORKING GROUP PROCESS

A research track assessment for spiny dogfish was planned for peer review in July 2022, to be followed by a management track assessment in fall 2022. However, the peer review was rescheduled to allow for data streams for the assessment to become available (of which included new ageing data and analyses for spiny dogfish). The peer review was rescheduled for December 2022, with an anticipated management track assessment in 2023. The Working Group was formed in June 2021 and met over a series of virtual meetings. Working Group meeting agendas were developed prior, based on feedback from the Working Group and non-Working Group members. The Working Group met during the following meetings:

1. July 30, 2021 – Kickoff meeting
2. September 22, 2021 – TORs 2, 3, 4
3. October 12, 2021 – TORs 2, 3, 4
4. November 15, 2021 – TORs 1, 2, 3, 4, 5
5. December 21, 2021 – TORs 2, 3, 4
6. January 19, 2022 – TORs 1, 2, 3, 4
7. February 15, 2022 – Stakeholder session
8. March 9, 2022 – TORs 1, 3, 4, 8
9. April 5, 2022 – TORs 4, 8
10. April 19, 2022 – TORs 2, 3, 4
11. May 2, 2022 – TORs 1, 3, 4, 8
12. May 11, 2022 – TORs 4, 8
13. June 15, 2022 – TOR 4
14. June 29, 2022 – TORs 4, 5, 8
15. August 23, 2022 – TORs 1, 2, 3

16. September 8, 2022 – TORs 1, 3, 4
17. September 22, 2022 – TORs 3, 4
18. October 4, 2022 – TOR 4
19. October 11, 2022 – TORs 4, 5, 6, 7
20. October 24, 2022 - TORs 4, 5, 6
21. November 1, 2022 - TORs 4, 5, 6
22. November 4, 2022 - TORs 1-8
23. November 15, 2022 - TOR 5, Assessment Document
24. November 16, 2022 - Assessment Document
25. November 17, 2022 - Assessment Document

Working Group members met through additional sub-TOR meetings to discuss finer details of various research to support individual TORs, of which discussions and recommendations were brought before the Working Group for consensus. Working Group materials (presentations, agendas, meeting minutes, literature, data, model runs, working papers, and assessment document drafts) were shared using a Google Drive folder. Working Group Co-Chairs and TOR Leads produced the report by compiling information from working papers, meeting minutes and presentations, and the draft report was reviewed and edited by Working Group members.

INTRODUCTION

Atlantic spiny dogfish (*Squalus acanthias*) is a schooling shark that is widely distributed across both sides of the North Atlantic. It is closely related to Pacific spiny dogfish, which previously was considered a subspecies of *Squalus acanthias,* but recently has been reclassified as its own species, *Squalus suckleyi* (Ebert et al. 2010). This assessment is for the Northwest Atlantic spiny dogfish stock (hereafter spiny dogfish refers to the Northwest Atlantic stock unless otherwise indicated).

Spiny dogfish are considered one of the most migratory shark species in the northwest Atlantic (Compagno 1984). It has a wide-ranging diet consisting of fish, such as herring, mackerel and sand lance, as well as invertebrates including ctenophores, squid, crustaceans and bivalves. Spiny dogfish are live bearers with a very long gestation period (18-24 months), and are slow growing with late maturation. Females grow larger than males and as a result, the fishery primarily targets females. In the northwest Atlantic, spiny dogfish occur from Florida to Canada, with highest concentrations from Cape Hatteras to Nova Scotia. In the winter and spring, they are found primarily in Mid-Atlantic waters, and tend to migrate north in the summer and fall, with concentrations in southern New England, Georges Bank, and the Gulf of Maine (though a recent study has created some uncertainty regarding the established migration paradigm, Carlson 2014).

Fishery and Management History

The management unit for spiny dogfish is the northwest Atlantic coast of the United States. Canadian landings are also accounted for by management. The management objectives of the Spiny Dogfish Fishery Management Plan (FMP) can be summarized as avoiding overfishing, avoiding management or regulatory conflicts, facilitating enforcement, and contributing to the protection of biodiversity and ecosystem structure and function.

The fishery was essentially unmanaged before 2000. Prior to about 1979, landings of spiny dogfish by U.S. and Canadian vessels were very low, with most catch likely being discarded. However, there were substantial landings by foreign trawlers, with landings peaking in the early 1970s at about 20,000 mt per year. A domestic fishery began to develop between 1979-1989, with annual landings averaging around 4,000 mt. Landings increased in the 1990s as other groundfish stocks declined, averaging over 20,000 mt per year from 1993-1998.

Observations of declining numbers and sizes of mature females as well as reduced recruitment (Rago et al. 1998) led to a determination in 1998 that this stock was overfished (NEFSC 1998). This led the Mid-Atlantic and New England Fishery Management Councils to develop a joint management plan that initially curtailed most directed fishing in order to rebuild the spiny dogfish stock. Low trip limits and catch reductions in the 2000s led to increases in spawning stock biomass and recruitment. The fishery was declared rebuilt in 2010, which allowed for the resumption of a directed fishery. Current management includes a 7,500 lb (3,402 kg) trip limit and an overall quota of 29.56 million lbs (13,408 mt), although a substantial decrease in the quota is likely for 2023. Table 1 describes the history of quotas and trip limits.

The Atlantic States Marine Fisheries Commission (ASMFC) approved an Interstate FMP to complement the federal plan in 2003, and ASMFC management sets regional and/or state allocations and trip limits. These allocations can restrict fishing at times even if the full quota has not been attained, though late 2019 ASMFC changes have facilitated state transfers that reduce (but do not eliminate) the state allocation constraint on total landings. Boats without federal spiny dogfish permits are not bound by the federal trip limit in state waters, but cannot retain spiny dogfish in federal waters. The federal spiny dogfish permit is not "limited access," so it can be added and/or dropped by fishery participants as they deem the ability to fish either federal waters, or state waters with higher trip limits, to be more advantageous.

Table 1. History of spiny dogfish quotas and trip limits. Note: The Councils have not always agreed on catch limits or trip limits - those listed here are as implemented by NMFS. States can also set their own trip limits for state waters.

Fishing Year	NMFS Commercial quota (mt)	Federal Trip Limit (pounds)	Notes
2000	1,814	600/300	Initially two seasonal quotas and trip limits. 5/1-10/31 and 11/1-4/30
2001	1,814	600/300	
2002	1,814	600/300	
2003	1,814	600/300	
2004	1,814	600/300	
2005	1,814	600/300	
2006	1,814	600	Trip limits for both periods or just annual hereafter
2007	1,814	600	
2008	1,814	600	
2009	5,443	3,000	Closed 9/26-10/31, 2009, and 1/26-4/30, 2010. ASMFC removes seasonal quotas
2010	6,803	3,000	Closed 8/27-10/31, 2010, and April 2011
2011	9,072	3,000	Closed 8/26-10/ 31, 2011, and 1/13-4/30, 2012
2012	16,191	3,000	
2013	18,526	4,000	New trip limit effective May 3, 2013
2014	22,243	5,000	New trip limit effective Sept 8; federal seasonal allocation ends Aug 2014
2015	22,957	5,000	
2016	18,307	6,000	New trip limit effective Aug 15, 2016
2017	17,735	6,000	
2018	17,325	6,000	
2019	9,309	6,000	
2020	10,521	6,000	
2021	13,408	6,000	
2022	13,408	7,500	New trip limit effective May 1, 2022

Assessment History

The following presents the chronology of spiny dogfish benchmark assessments with brief summaries regarding the findings:

<u>Anthony and Murawski (1985)</u>: During Stock Assessment Workshop (SAW) 1, several notes were made regarding key spiny dogfish uncertainties including discard mortality, data confidentiality limitations on calculating catch per unit effort (one company), growth rates, survey variability, predator/prey interactions, and harvest implications of the stock's low mean fertility, natural mortality rate, and long life span.

<u>NEFSC (1990):</u> During SAW 11, spiny dogfish were assessed as part of the small elasmobranchs group, which included skates and dogfish species. Landings, life history, trawl survey, and reference points based yield per recruit analyses were presented. General conclusions were that the population has substantially increased since the 1960s. To better understand the dynamics of spiny dogfish and its response to exploitation, future research recommended included: better evaluation of the Northeast Fisheries Science Center (NEFSC) survey indices as an indicator of stock abundance and biomass and means to estimate absolute population size; evaluating changes in population demographics over time, including size, age, and sex composition and population fecundity; evaluation of stock recruitment relationships from survey data; better understanding of the trophic dynamics of spiny dogfish in the ecosystem; and investigation of discard data to clarify the removals from the stock.

<u>NEFSC (1994):</u> During SAW 18, the assessment scientists addressed terms of references regarding patterns of landings and fishery dependent data, fishery independent abundance data, and biological reference points. Data suggested that the spiny dogfish stock in the Northwest Atlantic had begun to decline as a consequence of the recent increase in exploitation. Pups per recruit and biomass dynamic models were used to derive reference points and understand the population size. Swept-area estimates of the fishable biomass increased threefold from 1968 to 1988, but then declined by over 10%. It was recommended

that a management program with appropriate management targets for stock biomass and fishing mortality rates be quickly established.

NEFSC (1998): During SAW 26, the assessment from SAW 18 was updated with data through 1997. Several analyses were presented as part of the assessment: trends in length composition of landings and surveys, trends in recruitment, application of a Beverton-Holt mortality estimator, comparison of observed length-specific sex ratios and predictions of a mechanistic life history model, and revisions to the previous yield-per-recruit estimates. New biological reference points based on pups per recruit necessary for equilibrium were proposed. Although the stock was deemed to be at a moderate biomass level, a severe reduction in the mature component of the fishery was apparent, which can affect recruitment, and the stock was over-exploited.

NEFSC (2003): During SAW 37, the Beverton-Holt mortality estimator was again applied to derive mortality rates. Fishery selectivity was further explored, and stochastic estimates of fishing mortality and biomass for the stock were conducted for the first time. Fishery-dependent, fishery-independent, and life history information were evaluated. Poor recruitment was identified, with an apparent recruitment failure from 1997-2003.

NEFSC (2006): During SAW 43, the Stochastic Estimator of fishing mortality (*F*) and biomass (*B*), the primary model used in the assessment, was updated to include uncertainty in recreational catch and discarded catch by gear type. Fishing mortality rate reference points were improved by incorporating length specific patterns of fishing mortality into a measure of reproductive potential. Despite lower landings since 2001, fishing mortality rates on the fully recruited female stock component were above the rebuilding target. New biological reference points for spawning stock biomass based on the Ricker stock-recruitment model were developed. At this time, however, recent recruitment patterns did not conform to the Ricker model, soliciting a more detailed consideration of reproductive biology in the future. As such, the existing *F* and *B* reference points were retained. The stock was found to be not overfished and overfishing not occurring. Projections indicated that if recruitment returns to levels consistent with expected size-specific reproduction, the biomass would rebound by 2015.

O'Brien and Worcester (2010), TRAC (2010): Two models were attempted using different stock units and growth estimates. Scientists from Fisheries and Oceans Canada (DFO) presented a two area, two half year (Nov-Apr, May-Oct) time step model that allowed for migration in both directions over the Hague line (Haist et al. 2010). The model deficiencies identified included no recruitment being estimated for the Canadian side of the stock, and poor fits to length and survey data. NEFSC scientists attempted a one stock, annual Stock Synthesis 3 model (Sosebee et al. 2010). The model included both sexes and used <= 35 cm individuals as a recruitment index, and females >= 80 cm as a spawning stock biomass (SSB) index. However, the model did not provide reasonable population estimates for the stock, and also had issues with fit to survey abundance indices and length data. From this assessment, both proposed DFO and NEFSC models were rejected.

TOR1: ECOSYSTEM AND CLIMATE INFLUENCES

"*Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.*"

Distribution and Habitat Use

Spiny dogfish (*Squalus* spp.) are distributed worldwide in boreal and temperate continental shelves as far from shore as the 900 m contour (Stehlik 2007; Dell'Appa et al. 2015). Spiny dogfish can be found throughout the North Atlantic, but the northwest Atlantic Ocean population is not believed to mix with populations from across the Atlantic or other oceans of the world (Stehlik 2007). Although it is not considered the norm, there is evidence of transatlantic migration by individuals, with historic records spanning from off Newfoundland, Canada to southwest of Iceland and north of Scotland (Holden 1967, Templeman 1976) and a recent record from Georges Bank to just south of Ireland (McCandless 2022). Genetic findings have suggested that North Pacific spiny dogfish are distinct from the South Pacific and Atlantic regions (Verissimo et al. 2010). The general distribution of northwest Atlantic spiny dogfish is considered to be from Florida to Greenland, with most concentrated from Cape Hatteras, North Carolina, to Nova Scotia (Rago et al. 1998). Spiny dogfish have been found in water temperatures between 1 and 20°C, but are most often between 6 and 15°C (Dell'Appa et al. 2015, references therein).

The species seasonally migrates north and south on the northwest Atlantic shelf, as well as inshore and offshore with changes in water temperature (Garrison 2000; Dell'Apa et al. 2015). In U.S. waters, mature females typically will overwinter in waters off North Carolina, move north to southern New England and the Gulf of Maine in the spring, with migrations occurring south toward the Carolinas again in the fall (O'Brien and Worcester 2010). A satellite tag study has created some uncertainty regarding this established migration paradigm, suggesting movements may occur more regionally (Carlson et al. 2014). Conventional tagging studies have suggested the northwest Atlantic spiny dogfish population may be comprised of multiple stocks, principally separate U.S. and Canadian stocks, given limited intermixing (between 10 and 38.4 % intermixing rate) along the New England coast (Campana et al. 2007; O'Brien and Worcester 2010; Rulifson et al. 2012). These tagging studies have been used to understand prospective population structure for northwest Atlantic spiny dogfish (Figure 1.1, O'Brien and Worcester 2010). However, these studies primarily consisted of mature females and genetic studies on the species are lacking to determine the true distinctions between U.S. and Canadian Atlantic spiny dogfish.

Spiny dogfish distributions vary with sex (Haugen et al. 2017, McCandless 2022). Mature males overwinter on the outer continental shelf and slope off southern New England and down to the Delaware and Maryland border, whereas females tend to stay further inshore whether traveling down the coast to North Carolina or remaining in the Gulf of Maine (McCandless 2022). Male dogfish can be found with little female presence in the spring along the continental shelf (Chesapeake Bay to Long Island) between depth ranges of 70–80 m and 300–330 m (~ 90–150 km or 240–270 km from shore, Haugen et al. 2017). Male-skewed catches have also been observed in the fall in the western Gulf of Maine and on Georges Bank at depths of 80–250 m. Depth has been found to be the greatest environmental variable in predicting male-skewed dogfish locations (Haugen et al. 2017).

Previous work has suggested little or no distributional shifts over time (Nye et al. 2009). However, recent data from the NEFSC Spring Bottom Trawl Survey indicate that the center of abundance for spiny dogfish has moved 1.42 degrees (157.98 km) south, its range has expanded by 0.21 degrees (23.79 km), and has moved 18 meters shallower from 1974-1977 to 2017-2019 (NOAA Fisheries 2022a). The Fall Bottom Trawl Survey data indicate that spiny dogfish has moved 0.61 degrees (68.36 km) north, its range has contracted by 0.37 degrees (41.05 km), and has moved 32.8 meters deeper from 1974-1977 to 2017-2019 (NOAA Fisheries 2022a).

Species distribution modeling using NEFSC Bottom Trawl Survey data has also been used to understand spiny dogfish distributions and their changes over time (NOAA Fisheries 2022b). Spring survey indices are greater than those in the fall, at least in part due migrations into Canada in the summer and fall. Over time, the probability of occurrence in the spring has increased in areas such as the Gulf of Maine, and coastal and shelf break waters in the Mid-Atlantic Bight, whereas it has decreased on Georges Bank. In the fall, the probability of occurrence has increased in large regions of Southern New England, the Mid-Atlantic Bight, Georges Bank, and the Gulf of Maine, except it has decreased in the eastern Gulf of Maine and coastal waters of southern Massachusetts and Rhode Island. Friedland et al. (2020) reported that over the Northeast U.S. Continental Shelf, the areal distribution of occupancy habitat has decreased in the spring over time, and increased in the fall. From a suite of environmental (e.g., physical, primary and secondary productivity, benthic) variables, bottom temperature appeared to be the most important covariate in determining the presence of spiny dogfish. In a similar analysis conducted in nearshore waters from North Carolina and southern New England using the NEAMAP Bottom Trawl survey data, Dell'Apa et al. (2017) identified several variables were significant in predicting catch, including bathymetry, sea surface temperature, salinity, chlorophyll-a (chl-a) concentration, season, and time of survey. Females were predicted to occur more inshore than males, inhabiting warmer, less saline, and higher chl-a waters. Females were also in greater abundance in the spring and morning, with males more abundant in the fall and afternoon times.

Spiny dogfish has been characterized as having an overall low climate vulnerability rank, with high exposure to climate changes and low biological sensitivity (Hare et al. 2016). Correspondence between spiny dogfish distribution and environmental conditions have been identified. Using the NEFSC Bottom Trawl Survey data, Sagarese et al. (2014b) found patterns specific to ontogenetic stages. Neonate, immature, and mature dogfish selected warmer, more saline waters. In the fall, the authors found that larger dogfish occupied relatively warmer, shallower, and less saline waters and that neonates selected higher salinities. Using generalized additive models, seasonal occurrences for various stages of spiny dogfish have been linked to depth, bottom temperature, and prey species (e.g., Atlantic herring, Atlantic mackerel, *Doryteuthis* spp.; Sagarese et al. 2014a). Using these models to forecast distributions under a warming scenario suggest that higher regional probabilities of occurrence for most dogfish stages could result.

As part of the Stakeholder Session held during the Research Track Assessment, several participants described their perspectives on ecosystem drivers for spiny dogfish (Appendix A). With warming waters, a prospective indirect effect of increasing seal populations on spiny dogfish natural mortality was mentioned. The impact of groundfish on the spiny dogfish population was recommended to be investigated, which has since been explored through evaluating a suite of drivers on the retrospective patterns of groundfish stocks (Kerr et al. 2022). A spatial and temporal shift in spiny dogfish abundance and distribution was noted to have occurred over time, which has impacted the distance that harvesters need to travel now to catch the species (Appendix A). Aligning with previous studies (Sagarese et al. 2014a; Sagarese et al. 2016), stakeholders noted similar prey items of significance for spiny dogfish that may influence their distribution such as squid and herring (Appendix A). Although, studies have shown that spiny dogfish are opportunistic predators that prey on more abundant species and will shift their diet when these prey are not readily available, as seen with herring (Overholtz and Link 2006) and ctenophores (Link and Ford 2005). This may also be the case if prey distributions shift, unless those factors affecting the prey distribution, such as temperature, also influence spiny dogfish distribution.

As part of this assessment, catch per unit effort (CPUE) was also analyzed to determine whether ecological and economic conditions influence the catch rates (Jones et al. 2022). An inverse relationship between catch rate and depth was identified, with little variation in this effect between years when models were fit with an explicit year by depth interaction. This consistent relationship suggests that catch rates are consistently higher in shallow areas. A unimodal relationship between catch rate and the hour of the day emerged as well, perhaps either due to increased availability to the gear in this time period. Models also indicated a significant, cyclical relationship between CPUE and month, where there was an increase in CPUE early in the year followed by a decrease.

The Working Group explored the relationship between spiny dogfish abundance and distribution with the environment through species distribution modeling (Hansell and McManus 2022). A vector auto-regressive spatiotemporal model (VAST, Thorson 2015) was used to model the distribution of spiny dogfish over time using the NEFSC bottom trawl survey. VAST is a delta or hurdle model, where the probability of occurrence and positive catch rates are modeled separately as generalized linear mixed models, with resulting predictions integrated. Two seasonal models (spring and fall) were fit to sex specific catch rates from the NEFSC bottom trawl survey to estimate changes in spiny dogfish distribution. Models were only fit to strata that were consistently sampled and explored the influence of local environmental variables. While several environmental variables have been used to describe spiny dogfish abundance and distribution, for these analyses, only bottom temperature and depth were tested. Both seasonal models successfully converged, with depth proving to be a significant variable in the models, and thus it was included as a covariate influencing spiny dogfish abundance (Hansell and McManus 2022). Spatial estimates of probability of occurrence and abundance when present highlighted some degree of interannual variability (Hansell and McManus 2022). From these predictions, the center of gravity and effective area occupied were estimated. For both male and female spiny dogfish in the fall, center of gravity estimates were variable with no clear distribution shifts north/south or east/west (Figures 1.2). In the spring however, it appeared that the center of gravity for both sexes shifted east since the early 2000s (Figure 1.2). Effective areas occupied for both sexes and seasons were variable with no clear indication of a significant change over time (Figure 1.3).

Life History Processes and Rates

Several species' recruitment patterns in the Northeast U.S. Shelf have been found to change over time in concert with environmental changes (e.g., Perretti et al. 2017). The Working Group evaluated whether similar changes have occurred over time, and whether such changes may be driven by the environment. The goal of this exercise was to determine whether environmental influences on recruitment and recruitment per spawner have occurred, and if so, should such considerations be carried forward into the assessment model. As previously examined by Rago and Sosebee (2010), spiny dogfish recruitment and spawning females were analyzed from the NEFSC spring bottom trawl survey (McManus et al. 2022) These analyses were intended to use similar methods and definitions as Rago and Sosebee (2010) for comparability; as such, recruitment was defined here as fish ≤ 35cm, and spawning females were defined as ≥80 cm. Change point analyses on recruits and recruits per spawner did not identify any meaningful regimes over time (McManus et al. 2022). Recruitment correspondence to annual spring mean bottom temperature and the North Atlantic Oscillation (NAO) were also explored to determine if a significant relationship existed between the variables, despite there not having been a dramatic change over time. Both of these environmental indices indicated very little correlation to recruitment or recruits per spawner (McManus et al. 2022). Lastly, the impact of spring bottom temperature on the stock-recruit relationship was also tested using a Ricker model (Figure 1.4). While a model was successfully fit which incorporated temperature and highlighted its influence of recruits per spawning, this model was not a statistical improvement over a model without incorporating the environment, suggesting temperature's impact on the relationship was not significant. Based on these findings, the Working Group did not pursue investigation of environmental drivers on recruitment or the stock-recruit relationship within the assessment model.

Given marine species' growth rates and maturity schedules can be influenced by environmental conditions, the Working Group discussed evaluating such interactions for spiny dogfish. As part of the Research Track Assessment, spiny dogfish spines were aged to determine whether growth has changed in recent years compared to previous growth rates (Passerotti and McCandless 2022). While more recent growth rates were available, age and growth information was only available for select years over the last several decades and substantial uncertainties in the contemporary growth estimates persisted (Passerotti and McCandless 2022). Therefore, the Working Group determined that there was not enough time series information to test whether environmental conditions have changed growth over time. Growth was also investigated using mark-recapture data from fish tagged between 2011 and 2012, with the majority of recaptures within the first couple of years (McCandless 2022). A lower estimate for both L_∞ and k were seen for females when compared to estimates used in past assessments based on ageing data from 1980 - 1981 (Nammack et al. 1985). Male growth parameters did not decrease, indicating the changes seen in females would be more likely due to fishing pressure. Although the estimates from this study were not appropriate for incorporation into the assessment model due to the low sample size, lack of small-sized fish, and measurement error, these estimates do provide supporting evidence of a decrease in large females.

The maturity and fecundity analyses presented in Sosebee (2005) were also updated for this assessment (Sosebee 2022a), particularly for informing the assessment model. The Working Group did not explore whether time series trends in maturity patterns were driven by environmental conditions primarily because the temporal patterns seemed to align with changes in the relative abundance, with the hypothesis that declines in maturity concurrent with abundance were the result of fishing mortality more so than the environment. However, these relationships warrant further investigation.

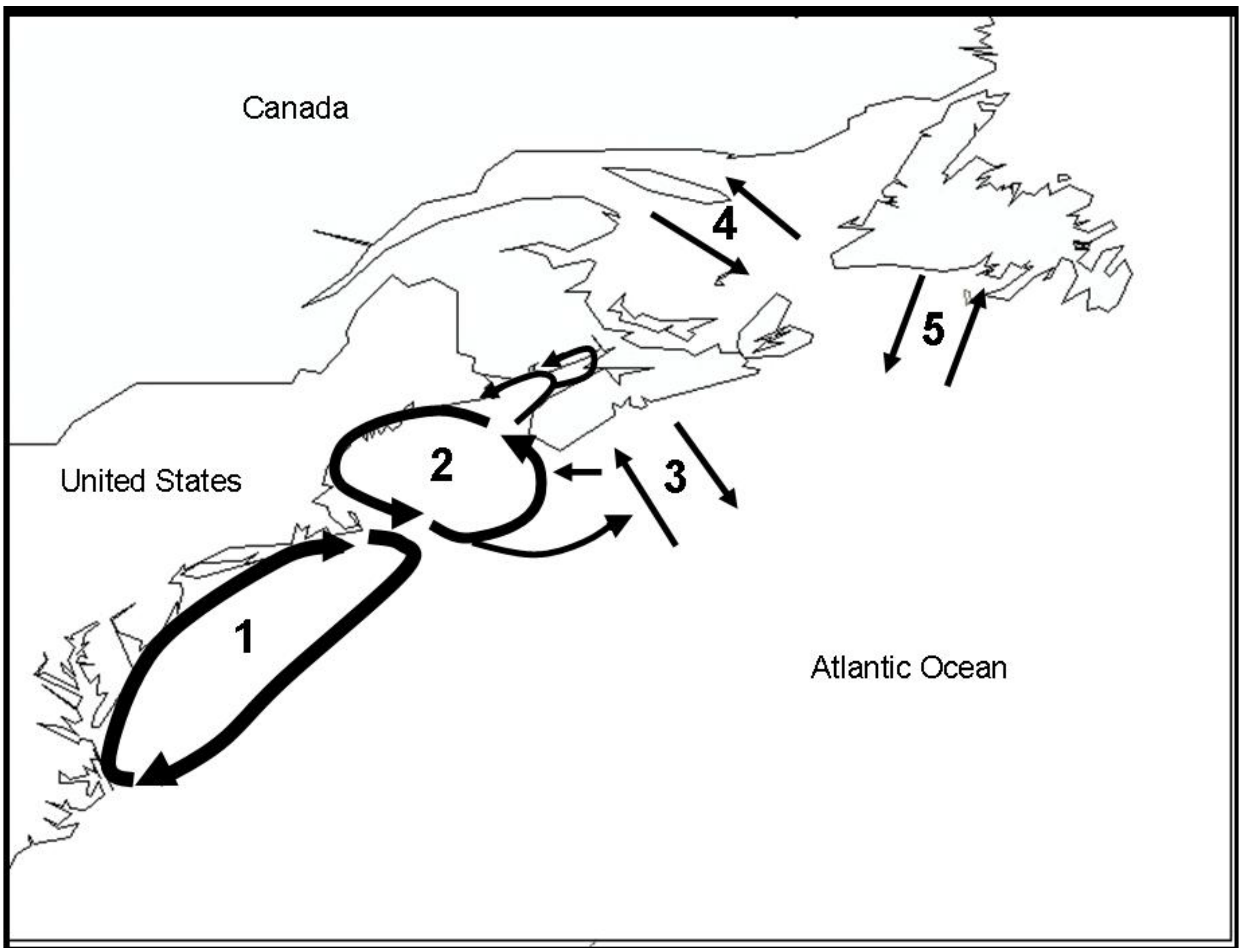


Figure 1.1. Spiny dogfish movements based on tagging data assessed during the most recent Transboundary Resources Assessment Committee benchmark assessment for Northwest Atlantic spiny dogfish (TRAC 2010).

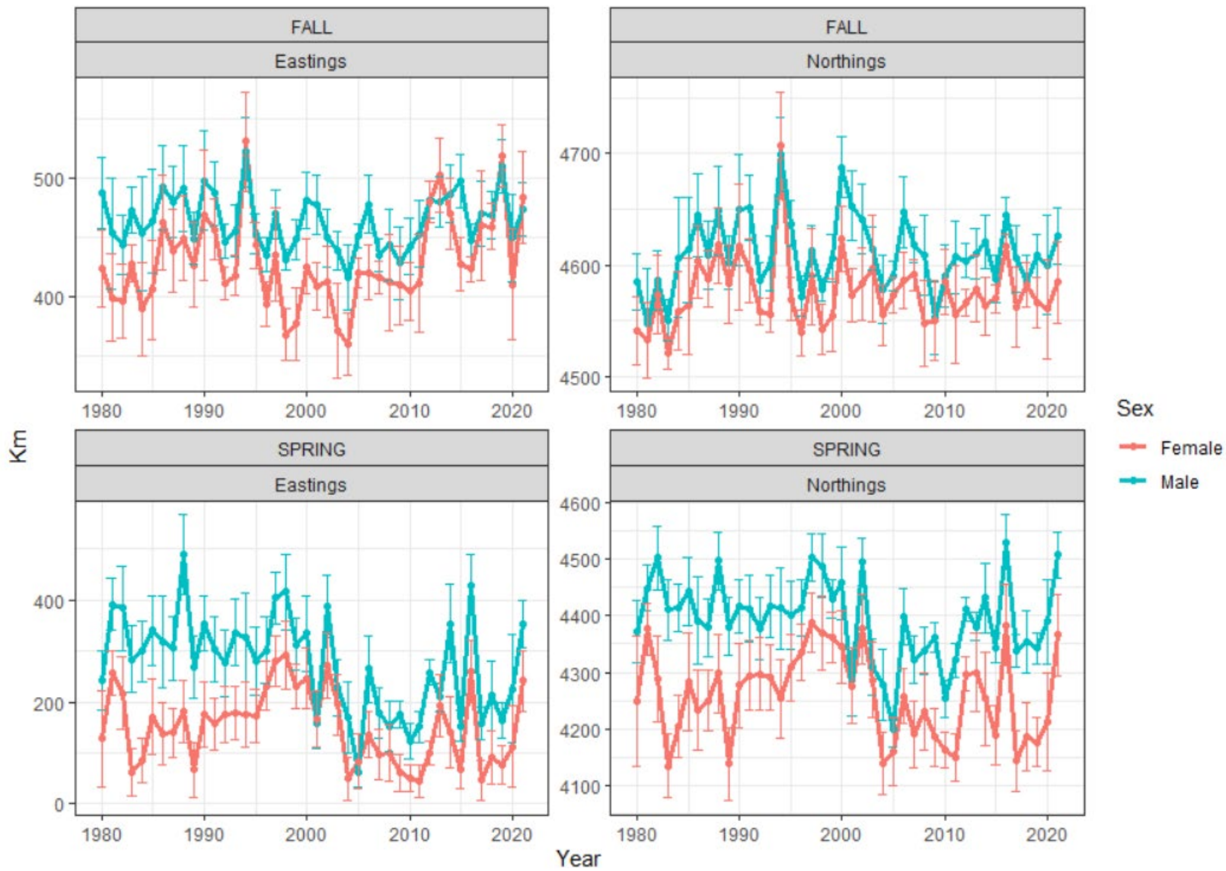


Figure 1.2. Center of gravity estimates from the VAST model for spiny dogfish. Results are presented by season and sex. Higher eastings values indicate the center of gravity is further east whereas higher northings values indicate the center of gravity is further north. Similarly, lower eastings values indicate the center of gravity is further west whereas lower northings values indicate the center of gravity is further south.

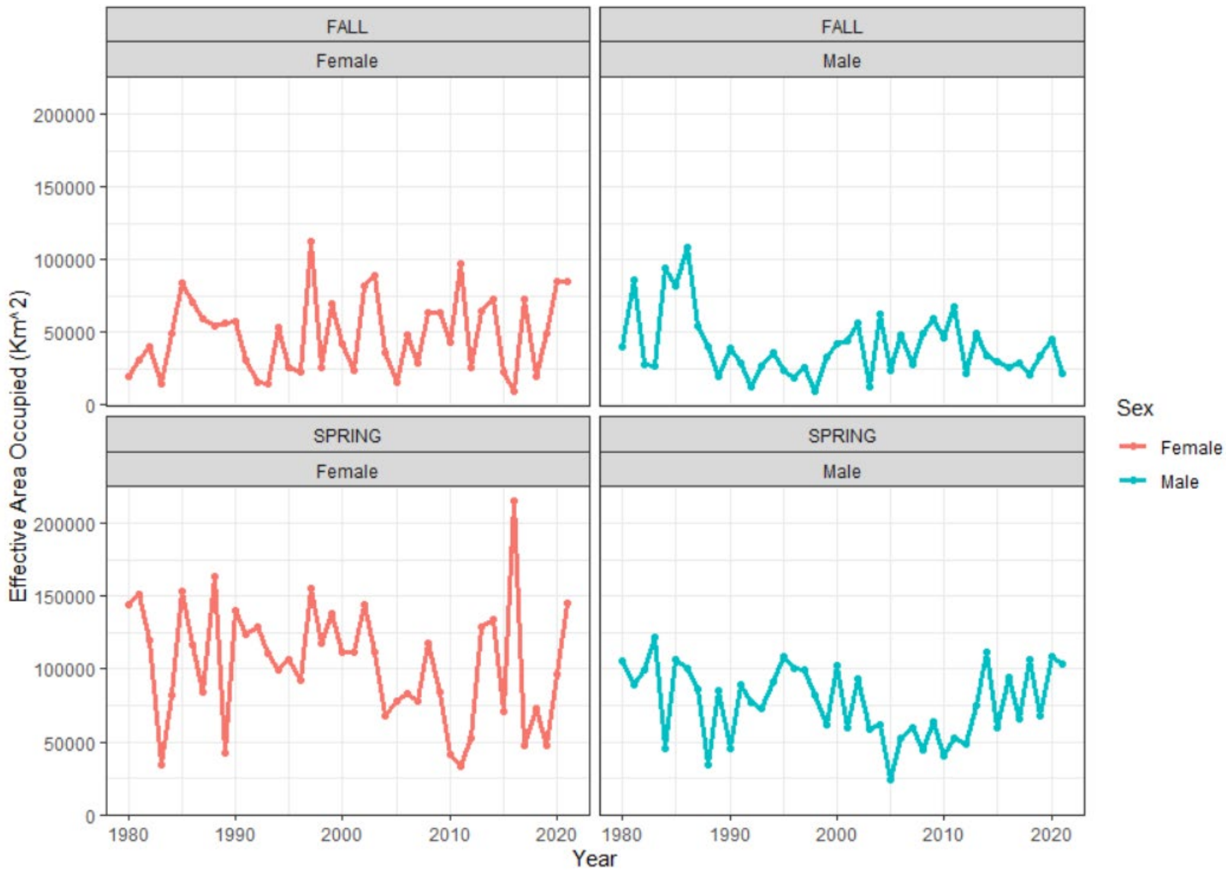


Figure 1.3: Effective area occupied estimates from the VAST model for spiny dogfish. Results are presented by season and sex.

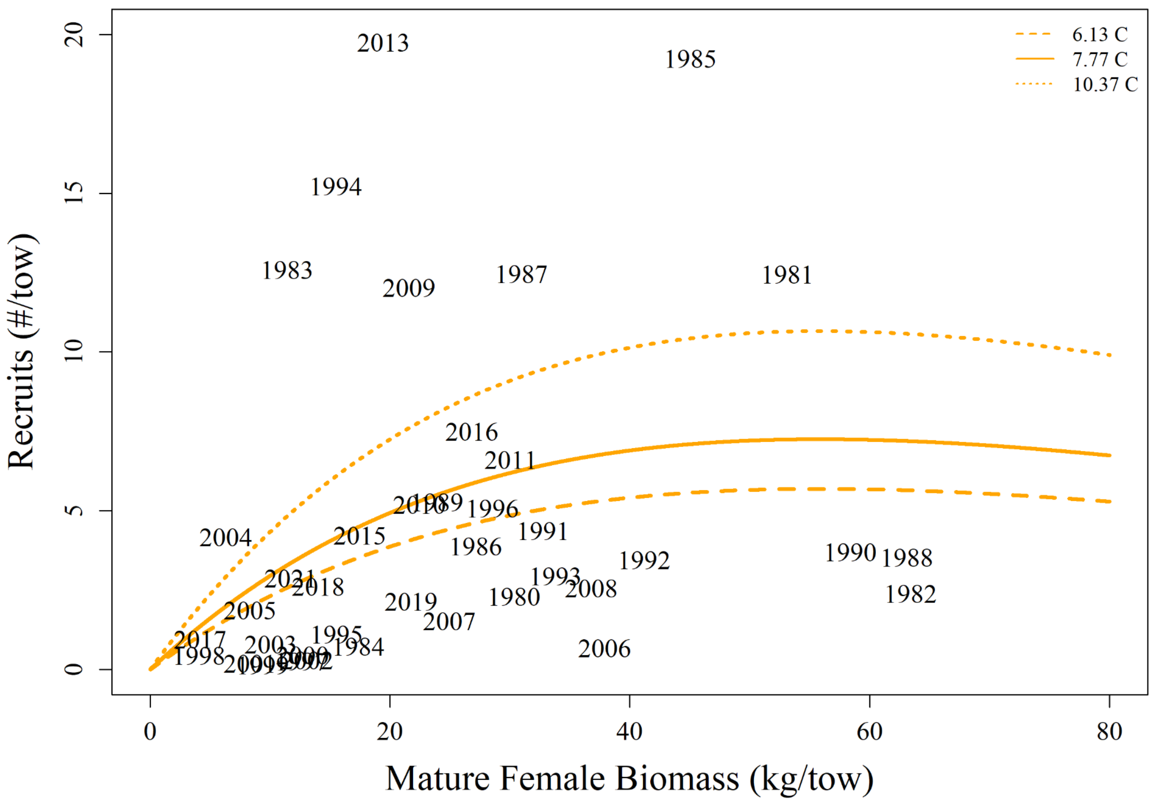


Figure 1.4. Ricker model fit to the mature female index and recruit index from the NEFSC spring bottom trawl survey, as defined in Rago and Sosebee (2010, with mean annual spring bottom temperature as a covariate in the model). Years represent annual data points. The solid line represents model predictions using the 50th percentile of the spring bottom temperature time series, whereas the dotted and dashed lines represent the minimum and maximum time series values, respectively.

TOR2: FISHERY DATA

"*Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.*"

Commercial Landings

Commercial landings data were obtained from the NEFSC commercial fisheries databases (General Canvass, Weighout and logbook), the Northwest Atlantic Fisheries Organization (NAFO) database (https://www.nafo.int/Data/STATLANT), Department of Fisheries and Oceans Canada, and the State of North Carolina for both spiny dogfish and unclassified dogfishes. The tables in the Appendix of NEFSC 1998 show which database (General Canvass or Weighout) the landings came from by state for 1962-1988 for unclassified dogfish and for spiny dogfish. Historical records dating back to 1931 indicate that U.S. commercial landings of dogfish in Subareas 5 and 6 were less than 100 mt in most years prior to 1960 (NEFSC 1990). Total commercial landings of spiny dogfish in NAFO Subareas 2-7 by all fisheries increased rapidly from the late 1960s to a peak of about 25,000 mt in 1974 (Table 2.1, Figure 2.1). Substantial harvests of dogfish by foreign trawling fleets began in 1966 in Subareas 5 and 6 and continued through 1977. After 1978, landings by foreign fleets were curtailed, and landings by U.S. and Canadian vessels increased. A sharp intensification of the U.S. commercial fishery began in 1990; estimated landings in 1996, in excess of 27,000 mt, were about five times greater than the 1980-1989 average. Landings between 1997 and 1999 averaged about 20,000 mt. Landings in 2001 and 2002 dropped dramatically due to restrictions imposed by federal and interstate fisheries management plans. Total landings further declined for the next couple of years, until the ASMFC increased the state quotas for 2006-2008 and landings increased slightly. When the stock was declared rebuilt in 2009, landings increased in response to increases in the allowed biological catch (ABC) and trip limits. Landings from 2011-2016 averaged nearly 10,000 mt but have been lower from 2017-2019.

United States Landings

U.S. commercial landings of dogfish from NAFO Subareas 2-6 were around 500 mt in the early 1960s, dropped to levels as low as 70 mt during 1963-1975 while averaging about 90 mt, and remained below 1,000 mt until the late 1970s (Table 2.1). Landings increased to about 4,800 mt in 1979 and remained fairly steady for the next ten years at an annual average of about 4,500 mt. Landings increased sharply to 14,900 mt in 1990, dropped slightly in 1991, but continued a rapid expansion from 18,987 mt in 1992 to over 28,000 mt in 1996. Landings in 1996 were the highest recorded. Landings declined in 1997 and 1998 to around 20,000 mt. In 1999, the last full year unaffected by spiny dogfish regulations, the landings declined to 14,860 mt. U.S. landings dropped to about 981 mt in 2004 in response to quota restrictions. The U.S. landings trend followed the total landings trend described above and in 2019 the landings were 7,910 mt.

The primary gears used by U.S. fishermen to catch spiny dogfish have been otter trawls and sink gill nets (Table 2.2, Figure 2.2). The latter accounted for over 50% of the total U.S. landings during the 1960s, while the former was the predominant gear through the 1970s and into the early 1980s. During the peak period of exploitation in the 1990s, sink gill nets were the dominant gear. Over the last nine years the landings by line gear have averaged almost 2,000 mt, otter trawls have averaged only 500 mt and sink gill nets averaged nearly 6,000 mt.

Since 1979, the bulk of the landings have occurred in Massachusetts (Sosebee 2022b). Landings at the height of the fishery (1991-2000) averaged nearly 20,000 mt. Other states with significant landings include New Jersey, Maryland, and Virginia. Landings in North Carolina peaked in 1996 at 5,992 mt, about half of the Massachusetts landings, but dropped sharply to about 1,300 mt between 1997 and 2000. North Carolina landings in 2001-2002 were negligible. In 2001 and 2002, virtually all of the landings were taken north of Rhode Island since the fishing year is May-April and the fish have migrated north in May. As the quotas increased, so did the landings in most states.

The temporal and spatial pattern of dogfish landings were closely tied to the north-south migration patterns of the stock. Peak landings from May through October coincide with residency of dogfish along the southern flank of Georges Bank, the Gulf of Maine, and the near shore waters around Massachusetts. As the population migrates to the south in late fall and early winter, landings increase in the southern states, especially North Carolina. U.S. dogfish landings have been reported in all months of the year, but most have traditionally occurred from June through September (Sosebee 2022b). During the peak years of the domestic fishery, substantial harvest was also taken during autumn and winter months. When the directed fishery was severely curtailed in 2001, landings by statistical area indicate that most landings during the 1980s originated from statistical area 514 (Massachusetts Bay; Figure 2.3; Sosebee 2022b) and continue to occur in this statistical area. Following the intensification of the fishery in 1990, statistical areas 537 (Southern New England) and 621 (off Delmarva and southern New Jersey) produced substantial quantities. In 1992 and 1993, large landings were reported from statistical areas 631 and 635 (North Carolina). When the directed fishery was reduced, the landings remained around Massachusetts (513, 514 and 521). In more recent years, landings have increased in more southern areas such as 614, 625, and 631.

The spatial distribution of commercial fishing landings and trips were assessed from vessel trip report data by year blocks and gear type (Jones 2022). In recent years (2010-2021), commercial landings from otter trawls were greatest from coastal Gulf of Maine and northern Georges Bank (514, 521, 522), southern New England (537, 539) and the northern Mid-Atlantic Bight (612, 613, 617). Recent gillnet catch was also spread across the northeast US, but much more in coastal waters and extending farther south (621, 625, 631). Long line catches have been more restricted to coastal waters off Massachusetts, and portions of the shelf break in southern New England. For more information on the spatial distribution of commercial effort on spiny dogfish, see Jones (2022). Overtime, the spatial distribution of the commercial fishery has been found to increasingly overlap with the center of abundance, and that increasing availability of the stock to the fishery has been more pronounced in the fall than spring (Sagarese et al. 2015).

Foreign Landings

A substantial foreign harvest of dogfish occurred mainly during 1966-1977 in Subareas 5 and 6, of which were taken primarily by the former USSR. Foreign landings averaged 13,000 mt per year during this time, and reached peaks of approximately 24,000 mt in 1972 and 1974 (Table 2.1). In addition to the former USSR, other countries that reported significant amounts of landings included Poland, the former German Democratic Republic, Japan, and Canada. Since 1978, foreign landings have averaged only about 900 mt annually and, except for those taken by Japan and Poland, have come primarily from Subareas 4 and 3. Canadian landings were low until 1979 when 1,300 mt were landed, and averaged 233 mt until 1990. Canadian landings increased about nine-fold between 1996 and 2001, and from 3,755 mt in 2001 to an average around 2055 mt from 2003-2008. Spiny dogfish taken by the distant water fleets were caught almost entirely by otter trawl, whereas Canadian landings were mainly harvested by gill nets and longlines. In the last ten years, the landings from Canada have been substantially reduced to an annual average of 42 mt.

Commercial Discards

Discard estimates were re-calculated as part of this assessment. The ratio-estimator used in this assessment is based on the methodology described in Rago et al. (2005) and updated in Wigley et al. (2007). It relies on a discard/kept (d/k) ratio, where the kept component is defined as the total landings of all species within a 'fishery.' A fishery is defined as a homogeneous group of vessels with respect to gear type (longline, otter trawl, sink gill net, and scallop dredge), quarter, region (New England, Mid-Atlantic), and by mesh size for otter trawls (<= 5.49", > = 5.5"). All trips were included if they occurred within this stratification, regardless of whether they caught spiny dogfish.

The discard ratio (R_h) for dogfish in stratum h is the sum of discard weight over all trips divided by sum of kept weights over all trips:

$$\hat{R}_h = \frac{\sum_{i=1}^{n_h} d_{ih}}{\sum_{i=1}^{n_h} k_{ih}} \qquad (1)$$

where d_{ih} is the discards for dogfish within trip i in stratum h and k_{ih} is the kept component of the catch for all species. R_h is the discard rate in stratum h. The stratum weighted discard to kept ratio is obtained from the weighted sum of discard ratios over all strata:

$$\hat{R} = \sum_{h=1}^{H} \left(\frac{N_h}{\sum_{h=1}^{H} N_h} \right) \hat{R}_h \qquad (2)$$

The total discards within a strata is simply the product of the estimated discard ratio R and the total landings for the fishery defined as stratum h, i.e., $D_h = R_h K_h$.

Cells (area/quarter/gear/mesh) with less than or equal to three trips were imputed using the sum of discards divided by the sum of kept. The order of imputation was half year within region, annual within region and then across region. For longline gear, there were many missing years. To estimate these, the sum of discards divided by the sum of kept over 1993-2003 was used (for 1993-2001). In 2002, there were two longline trips with a large amount of discards that gave an anomalously high value of total discards. For this year, those trips were omitted to derive the d/r ratio. For scallop dredge (1989-1991) and longline (1989-1990) trips, the average d/k ratio for the first three years for scallop dredge (1992-1994) and for longline (1991-1993) was used to derive the discards. Discards and number of trips by half year and gear type are shown in Tables 2.3-2.5 along with coefficients of variation (CVs) by gear type. The discard mortalities vary by gear type (Table 9 in NEFSC 2006).

Commercial discard estimates over the time series were generally higher than those estimated in the last assessment particularly in the early part of the time series when more imputation was required (NEFSC 2006; Sosebee 2022b). Additionally, some commercial data were revised since the previous assessment, which caused some years to previously have higher discards than this revised version. Discards declined from 1993 through 2000, were stable until 2010, and slowly declined to the lowest value in the time series in 2019 (Figure 2.4).

Size and Sex Composition of Commercial Landings and Discards

The sex of commercial landings was not recorded routinely until 1982 and discards until 1991. For details on the commercial landings sampling program, see Burns et al. (1983). The estimated sex composition of the landings from previous assessments was based on pooled samples over the entire year. For this assessment, the Working Group estimated the sex and size composition by gear type and by half year. The details are given in Sosebee 2022b.

Recreational Landings and Discards

Recreational landings and discards were obtained from the Marine Recreational Information Program (MRIP) http://www.st.nmfs.noaa.gov/recreational-fisheries/access-data/run-a-data-query/index. Descriptions of the program are in Van Voorhees et al. (1992) and Papacostas and Foster (2021). Of note, recreational catch since 2018 uses a mail-based survey for total effort to improve response rates and reduce bias, and catch before 2018 is calibrated from effort estimates from a telephone-based survey (Breidt et al. 2017).

The MRIP estimates are partitioned into three categories of numbers caught: A, B1, and B2. Type A catches represent landed fish enumerated by the interviewer, while B1 are landed catches reported by the angler. Type B2 catches are those fish caught and returned to the water. Biological information on recreationally caught dogfish is generally scant and the data are not collected by sex.

Recreational landings in number ranged between 1,736 and 806,857 over the entire time period with no observable trend (Table 2.6, Figure 2.5). The total discards are a larger fraction of the catch ranging from 128,652 to over 7 million fish in 2014, with the largest discards occurring from 2004-2019. Recreational discard mortality was assumed to be 20%, which is at the high end of published studies of other fish (NEFSC 2006). This makes the range of dead discards 25,730 to over 1 million fish.

Size and Sex Composition of Recreational Landings and Discards

The previous assessments assumed an average weight of 2.5 kg per fish based on limited length information to convert numbers of fish to metric tons. This assessment is using a different method based on the average length composition (Details in Sosebee 2022b). The range of landings in mt is 4 to 2,837 mt (Figure 2.6) with the majority of the time series < 500 mt and less than 250 mt from 2005-2019. Discards increased between 2000 and 2009, peaked at just over 2,700 mt in 2014, and averaged almost 600 mt between 2016 and 2019.

Discard Mortality

The Working Group reviewed the literature to determine if new research has been conducted to inform inferences in discard mortality in the commercial and recreational fisheries. Several papers were examined by the Working Group (Rago et al. 1998; Mandleman and Farrington 2007a; Mandleman and Farrington 2007b; Rulifson 2007; Courtney and Mathers 2019; Courtney et al. 2021). The Working Group did not find new research quantifying discard mortalities rates for Atlantic spiny dogfish in either the commercial or recreational fisheries As such, the Working Group used discard mortality rates that were previously used during the NEFSC (2006) assessment. A single discard mortality rate was assigned for the recreational fishery as 0.20, whereas those for commercial fisheries were designated by gear type: otter trawl (0.50), sink gillnet (0.30), scallop dredge (0.70), longline (0.10; Sosebee 2022b).

Overall Sex Composition

The number of females landed by all gears combined increased between 1989 and 1996 to about 10.5 million fish (Figure 2.7). The same increase occurred with males but on a much smaller scale (average of 2 million fish between 1996 and 1999). Since 2005, females have averaged around 3 million fish while males averaged around 500,000 fish. The sex ratio of the discards was closer to 0.5 over the time series (Figure 2.8). Overall catch of females averaged nearly 12 million fish from 1989-1999 and decreased to just over 4 million fish from 2005-2019 (Figure 2.9). Males decreased from an average of 5 million fish to 1.7 million over the same time periods.

Commercial Trawl Catch Per Unit Effort

Evaluating CPUE information in the stock assessment process can provide additional information on stocks' interannual changes, particularly when the spatiotemporal patterns of existing fisheries-independent surveys do not adequately capture the species spatial and temporal distribution (Cadrin et al. 2020). CPUE metrics can use various metrics of effort to standardize the catch rates to evaluate the performance of a fishery. During the Stakeholder Session Meeting, harvesters inquired about the utility of deriving CPUE indices to evaluate the fishery and stock (Appendix A). As part of the assessment, CPUE indices were derived by combining existing data from two of the region's fine-scale fishery dependent data sets: NEFSC Study Fleet Program and Northeast Fisheries Observer Program (Jones et al. 2022). Integrating bottom trawl gear observations from these datasets, which represent the largest sample of records by gear type and avoid issues related to targeting, both nominal and model-based CPUE annual indices were derived to understand the efficacy of CPUE in tracking changes in the stock. The model-based approaches used Generalized Additive Models (GAMs) with a suite of ecological and economic covariates that were hypothesized to influence CPUE. From 2007 through 2021, the nominal CPUE index was highlighted by variability over time, with either a stable or slightly declining trend. The model-based indices accounting for significant covariates produced smaller confidence intervals around the annual indices than the nominal index, but a stronger decline from the early 2010s through 2021 (Figure 2.10). Correlations between the CPUE indices and NEFSC bottom trawl survey indices varied in the relationship and significance; spring bottom trawl indices tended to be negatively correlated with the CPUE indices, and positively correlated with the fall indices (Jones et al. 2022).

During the Stakeholder Session, harvesters noted recent lower yields and attributed them to not being able to get away from smaller fish, and the challenges of needing to find the marketable big fish within small fish schools (Appendix A). Others noted that the past year was one of the first where medium to large females were found. Another consideration for the changes in CPUE indices over time are gear and management changes. While gillnet gear was not included in these CPUEs, for example, one person during the Stakeholder Session noted that when approximately four years ago, gillnet gear changed from 6.5" to 7" mesh, catch decreased due to extrusion through the mesh (Appendix A).

Table 2.1. Total spiny dogfish commercial landings (mt, live) in NAFO Subareas 2 to 7, 1962-2019 by country.

Year	United States	Canada	Distant Water Fleets	Total Landings
1962	235	0	0	235
1963	610	0	1	611
1964	730	0	16	746
1965	488	9	198	695
1966	578	39	9,389	10,006
1967	278	0	2,436	2,714
1968	158	0	4,404	4,562
1969	113	0	9,190	9,303
1970	106	19	5,640	5,765
1971	73	4	11,566	11,643
1972	69	3	23,991	24,063
1973	89	20	18,793	18,902
1974	127	36	24,513	24,676
1975	147	1	22,523	22,671
1976	550	3	16,788	17,341
1977	931	1	7,199	8,131
1978	828	84	622	1,534
1979	4,753	1,331	187	6,271
1980	4,085	660	599	5,344
1981	6,865	564	974	8,403
1982	5,411	389	364	6,164
1983	4,897	0	464	5,361
1984	4,450	2	391	4,843
1985	4,028	13	1,012	5,053
1986	2,748	20	368	3,136
1987	2,703	281	139	3,123
1988	3,105	1	647	3,753
1989	4,492	167	256	4,915

1990	14,729	1,309	393	16,431
1991	13,104	307	234	13,645
1992	16,427	868	67	17,362
1993	20,777	1,435	27	22,239
1994	18,305	1,820	2	20,127
1995	21,588	956	14	22,558
1996	26,926	431	236	27,593
1997	18,351	446	214	19,011
1998	20,628	1,055	607	22,290
1999	14,855	2,091	554	17,500
2000	9,257	2,741	402	12,400
2001	2,294	3,820	677	6,791
2002	2,199	3,584	474	6,257
2003	1,170	1,302	643	3,115
2004	981	2,362	330	3,673
2005	1,146	2,270	330	3,746
2006	2,248	2,439	10	4,697
2007	3,008	2,384	31	5,423
2008	4,135	1,572	131	5,838
2009	5,392	113	82	5,587
2010	5,440	6	127	5,573
2011	9,479	125	143	9,747
2012	10,595	65	137	10,797
2013	7,312	5	61	7,378
2014	10,649	54	31	10,734
2015	8,663	1	23	8,687
2016	12,097	32	24	12,153
2017	8,735	54	0	8,789
2018	6,878	45	0	6,923
2019	7,910	36	1	7,947

Table 2.2. United States spiny dogfish commercial landings (mt, live) by gear type, 1962-2019. Other gear includes seines, dredges, pots, and unknown.

Year	Line Trawl	Otter Trawl	Sink Gill Net	Other	Total
1962	18.7	78.3	129.4	8.4	234.9
1963	49.8	85.5	435.5	38.8	609.6
1964	12.5	75.4	619.0	23.4	730.4
1965	55.1	52.3	358.4	22.2	488.0
1966	84.7	95.2	358.0	40.1	578.1
1967	23.9	110.8	98.0	44.9	277.5
1968	2.5	78.0	54.3	23.2	158.0
1969	1.9	88.4	6.4	16.7	113.4
1970	1.8	80.5	12.4	11.0	105.7

1971	0.0	53.0	4.1	16.2	73.3
1972	0.6	53.5	0.7	14.4	69.2
1973	0.5	76.7	6.3	5.8	89.4
1974	1.9	79.2	11.3	34.9	127.3
1975	0.3	89.4	14.4	42.8	146.9
1976	5.2	71.6	438.3	34.5	549.6
1977	2.8	102.6	798.9	27.2	931.4
1978	3.4	121.4	687.1	16.6	828.4
1979	17.7	3517.6	1199.8	17.6	4752.7
1980	12.1	3370.1	638.2	64.7	4085.1
1981	1.0	6287.1	568.1	8.7	6865.0
1982	2.9	5065.6	320.1	22.0	5410.6
1983	0.2	3367.5	1523.7	5.1	4896.5
1984	0.9	2486.0	1955.6	7.9	4450.4
1985	158.7	2844.4	1017.4	7.6	4028.0
1986	2.6	1258.1	1470.3	16.7	2747.6
1987	7.8	1848.1	814.6	32.8	2703.4
1988	4.7	1589.5	1502.1	9.0	3105.2
1989	144.5	486.5	3859.8	1.3	4492.0
1990	17.7	7010.8	7698.3	1.7	14728.5
1991	31.5	5199.5	7849.7	23.0	13103.6
1992	28.9	4978.9	11388.6	30.7	16427.1
1993	259.7	5087.8	15417.1	11.9	20776.5
1994	853.5	2844.2	14467.3	139.7	18304.6
1995	1725.5	2194.6	17402.4	265.5	21588.0
1996	1650.1	3136.7	22051.4	87.4	26925.6
1997	1423.4	1786.4	15080.9	60.6	18351.2
1998	1503.5	2656.7	16427.8	39.7	20627.6
1999	1760.6	2269.7	10597.2	227.1	14854.6
2000	1835.0	3175.3	4235.5	10.9	9256.7
2001	1328.4	239.8	717.1	8.3	2293.6
2002	1074.4	236.6	885.0	2.9	2198.9
2003	664.7	38.0	409.5	57.8	1170.0
2004	45.0	150.6	760.7	24.7	981.0
2005	149.1	251.5	694.0	51.2	1145.7
2006	263.1	469.4	1349.3	166.4	2248.2
2007	484.7	201.7	1891.0	430.0	3007.6
2008	533.9	269.7	2928.2	403.0	4134.8
2009	595.3	809.1	3792.3	195.7	5392.3
2010	754.2	666.6	3880.7	138.8	5440.2
2011	1006.2	1082.8	7049.5	341.0	9479.4
2012	2298.6	809.6	7065.3	421.9	10595.3
2013	943.8	550.0	5566.0	252.5	7312.3
2014	2194.2	531.6	7650.2	272.8	10648.7
2015	1897.7	390.8	6261.4	113.0	8662.9
2016	3376.3	445.4	8114.1	161.0	12096.9

2017	2045.2	466.7	6015.9	207.2	8734.9
2018	1836.3	288.2	4514.5	239.3	6878.3
2019	1445.8	220.5	5887.4	356.7	7910.3

Table 2.3. Discard estimates of spiny dogfish in the large mesh (LM >= 5.5 inches) and small mesh (SM <= 5.49 inches) otter trawl (OT) fleets from the Northeast Fisheries Observer Program from 1989-2019 split out by the first (H1) and last (H2) half of each year.

Year	LM H1 trips	LM H1 discards	LM H2 trips	LM H2 discards	LM Total trips	LM Total discards	LM Total CV	SM H1 trips	SM H1 discards	SM H2 trips	SM H2 discards	SM Total trips	SM Total discards	SM Total CV	Overall OT CV
1989	31	8433.6	30	6568.3	61	15001.9	29.9	45	9423.8	75	3979.0	120	13402.8	39.5	24.4
1990	26	6965.5	28	18270.1	54	25235.6	38.4	41	7553.3	43	6974.6	84	14527.9	29.1	26.6
1991	31	4279.4	51	9232.2	82	13511.5	20.6	61	3117.5	113	3860.8	174	6978.3	23.5	15.7
1992	64	40401.9	18	14873.9	82	55275.8	30.5	52	6231.3	52	3374.6	104	9605.9	44.7	26.8
1993	26	4875.3	30	7872.1	56	12747.4	31.6	27	3466.1	20	4278.2	47	7744.3	19.5	21.0
1994	42	4903.1	15	528.7	57	5431.9	26.5	13	645.6	20	6563.1	33	7208.7	55.8	33.8
1995	56	8574.5	67	4253.1	123	12827.7	37.1	26	971.7	77	6977.6	103	7949.3	27.7	25.2
1996	32	2118.7	30	1037.7	62	3156.4	36.4	36	6979.0	94	410.7	130	7389.7	22.3	19.0
1997	23	2342.5	15	539.8	38	2882.3	34.1	48	2337.7	22	272.7	70	2610.4	36.6	24.9
1998	21	1806.4	5	641.9	26	2448.4	22.0	15	2794.2	23	1966.0	38	4760.2	29.8	21.1
1999	17	1749.3	32	3104.8	49	4854.1	30.2	22	170.5	32	3021.7	54	3192.1	31.0	22.0
2000	77	1802.0	52	320.9	129	2122.9	26.0	29	203.5	27	594.6	56	798.1	36.7	21.4
2001	71	1492.1	136	1307.6	207	2799.6	23.8	38	300.1	36	714.1	74	1014.2	19.3	18.2
2002	47	1932.4	212	1510.5	259	3443.0	22.7	27	209.6	70	1483.6	97	1693.2	10.1	15.5
2003	196	972.6	207	1224.6	403	2197.1	14.5	67	632.5	80	1135.4	147	1767.9	27.4	14.6
2004	227	855.2	413	1816.1	640	2671.2	12.8	149	1309.5	281	1238.8	430	2548.3	24.0	13.4
2005	670	1014.5	773	1719.6	1443	2734.1	20.5	181	684.1	244	1427.7	425	2111.9	18.4	14.1
2006	415	870.1	275	3344.0	690	4214.1	33.6	126	1183.8	110	1063.4	236	2247.2	17.8	22.8
2007	332	2441.7	449	2356.5	781	4798.2	19.6	126	1924.8	168	2195.8	294	4120.7	18.0	13.4
2008	412	1058.4	473	1413.7	885	2472.0	11.4	106	1208.9	107	797.5	213	2006.4	24.2	12.5
2009	479	2163.5	567	1100.6	1046	3264.2	15.1	199	3389.6	306	1395.5	505	4785.1	14.4	10.5
2010	523	2435.1	807	1390.9	1330	3825.9	8.4	313	1062.9	294	640.2	607	1703.1	16.6	7.7
2011	898	1990.2	953	2144.8	1851	4135.0	9.0	255	1816.7	302	593.1	557	2409.8	19.7	9.2
2012	977	2653.6	743	1681.1	1720	4334.7	7.8	185	1520.8	201	843.2	386	2364.0	21.3	9.0
2013	789	2169.3	557	3172.5	1346	5341.9	8.0	279	931.9	358	648.4	637	1580.4	16.6	7.3
2014	706	3435.7	761	1816.4	1467	5252.1	10.3	321	2250.6	441	736.1	762	2986.7	11.1	7.7
2015	609	1754.0	519	1296.5	1128	3050.4	11.6	280	1592.4	369	489.6	649	2082.1	14.3	9.0
2016	455	1684.2	463	1348.6	918	3032.8	9.4	374	1080.3	629	967.5	1003	2047.8	13.9	7.9
2017	444	1686.4	521	935.7	965	2622.1	9.7	681	2096.7	971	732.0	1652	2828.7	11.5	7.6
2018	486	1009.3	468	1175.5	954	2184.8	12.6	441	1088.2	788	656.4	1229	1744.6	15.2	9.7
2019	595	1037.5	758	1650.1	1353	2687.7	7.1	484	1912.9	632	837.6	1116	2750.5	9.7	6.0

Table 2.4. Discard estimates of spiny dogfish in the sink gill net (SGN) and longline (LL) fleets from the Northeast Fisheries Observer Program from 1989-2019 split out by the first (H1) and last (H2) half of each year.

Year	SGN H1 trips	SGN H1 discards	SGN H2 trips	SGN H2 discards	SGN Total trips	SGN Total discards	SGN Total CV	LL H1 trips	LL H1 discards	LL H2 trips	LL H2 discards	LL Total trips	LL Total discards	LL Total CV
1989	1	3042.0	106	4995.7	107	8037.7	14.0		707.6		429.0		1136.7	
1990	75	1501.4	78	2447.9	153	3949.2	28.0		566.4		445.1		1011.5	
1991	194	5277.6	763	8983.0	957	14260.7	8.6	1	529.6	17	414.9	18	944.5	4.3
1992	497	1844.5	690	3734.9	1187	5579.4	10.1	32	833.3		643.8	32	1477.1	9.5
1993	348	1637.4	422	5478.9	770	7116.2	19.5	3	3333.4	1	2209.1	4	5542.5	
1994	188	343.8	216	1058.2	404	1402.1	23.5	2	2612.0		2201.4	2	4813.4	
1995	298	1119.8	239	3124.8	537	4244.7	31.1	1	2359.5		2384.3	1	4743.8	
1996	254	916.4	168	1587.1	422	2503.5	21.3		2215.1		2067.9		4283.0	
1997	257	1066.2	132	1010.4	389	2076.6	24.8		2401.4		2310.6		4712.0	
1998	267	552.9	136	942.2	403	1495.1	24.5		1995.8	1	2408.7	1	4404.5	
1999	88	1243.9	101	647.0	189	1890.8	26.9		1845.0		1893.7		3738.7	
2000	118	2003.2	108	2710.2	226	4713.4	29.1		1105.8		2082.4		3188.2	
2001	98	1810.4	69	4905.7	167	6716.0	30.2		1578.0		1761.1		3339.1	
2002	67	1522.7	106	3830.1	173	5352.8	20.9		1677.0	9	1012.9	9	2689.9	95.2
2003	162	1110.6	330	4137.9	492	5248.5	12.4	17	6.9	2	9.9	19	16.8	7.9
2004	289	899.4	800	3202.0	1089	4101.5	7.7	9	117.8	113	474.4	122	592.3	10.6
2005	260	1265.9	744	2168.2	1004	3434.2	12.8	88	231.5	204	242.2	292	473.7	12.5
2006	136	930.1	115	2040.1	251	2970.2	19.3	46	471.7	56	661.9	102	1133.5	21.1
2007	100	3076.8	234	1943.6	334	5020.4	22.9	24	142.8	69	1798.9	93	1941.7	39.7
2008	115	2068.1	194	2769.8	309	4837.9	18.2	27	114.7	52	150.5	79	265.2	11.6
2009	190	1098.9	226	4143.7	416	5242.5	14.1	35	129.5	55	599.0	90	728.5	19.8
2010	419	1002.8	1460	1383.2	1879	2386.0	9.3	72	228.8	120	168.7	192	397.5	23.8
2011	733	747.5	1326	2092.5	2059	2840.1	5.4	77	80.5	41	248.5	118	329.0	17.2
2012	755	1112.1	933	1894.8	1688	3007.0	6.4	107	57.3	112	113.2	219	170.6	7.9
2013	233	1177.3	601	1898.9	834	3076.2	9.5	32	37.2	4	55.0	36	92.2	18.9
2014	410	946.9	962	1458.2	1372	2405.2	9.4	26	10.4	18	6.7	44	17.2	51.4
2015	315	758.5	750	916.0	1065	1674.5	23.7	8	23.9	4	27.8	12	51.7	30.1
2016	443	1213.2	543	728.8	986	1942.0	23.0	15	38.9	9	236.0	24	274.9	24.0
2017	485	323.1	622	558.0	1107	881.2	13.7	27	23.2	35	176.7	62	199.9	24.6
2018	374	606.0	456	505.7	830	1111.6	18.4	23	2.4	52	98.3	75	100.7	17.9
2019	586	414.0	584	504.3	1170	918.3	17.5	29	5.9	37	83.6	66	89.4	22.5

Table 2.5. Discard estimates of spiny dogfish in the scallop dredge fleet from the Northeast Fisheries Observer Program from 1989-2019.

Year	Half 1 trips	Half 1 discards	Half 2 trips	Half 2 discards	Total trips	Total discards	Total CV
1989		584.6		293.9		878.6	
1990		556.7		357.0		913.7	
1991		633.6		282.9		916.6	
1992	8	364.4	10	334.1	18	698.5	63.6
1993	14	219.4	8	8.1	22	227.5	40.0
1994	11	350.1	12	271.0	23	621.1	41.0
1995	15	223.0	12	142.2	27	365.2	18.4
1996	22	96.1	18	43.5	40	139.7	31.7
1997	19	117.0	10	81.1	29	198.1	20.6
1998	9	44.4	17	71.2	26	115.6	11.7
1999	15	13.7	56	9.6	71	23.2	40.9
2000	38	17.1	218	26.3	256	43.4	40.4
2001	58	6.3	48	19.2	106	25.5	30.5
2002	34	36.8	66	37.7	100	74.5	18.3
2003	50	63.2	74	51.9	124	115.1	21.7
2004	85	67.6	212	28.4	297	96.0	13.1
2005	128	32.5	206	24.4	334	56.9	17.9
2006	45	75.7	183	95.4	228	171.1	23.2
2007	158	158.2	202	72.5	360	230.7	11.2
2008	385	172.4	257	86.0	642	258.4	11.8
2009	373	334.3	117	123.0	490	457.3	12.1
2010	145	134.6	194	59.2	339	193.8	10.9
2011	177	122.5	216	103.6	393	226.1	16.7
2012	237	337.3	186	87.5	423	424.8	8.8
2013	245	82.5	234	47.9	479	130.4	9.7
2014	233	86.9	250	21.1	483	108.0	10.4
2015	288	26.7	245	14.3	533	41.0	14.5
2016	362	80.4	271	39.2	633	119.7	14.0
2017	377	57.3	269	17.3	646	74.6	12.0
2018	275	71.0	282	63.6	557	134.6	14.4
2019	281	54.4	282	79.2	563	133.5	17.1

Table 2.6. Summary of spiny dogfish landings and discards based on Marine Recreational Information Program estimates. Discard mortality is assumed to be 20%.

Year	Observed Harvest (A)	PSE	Reported Harvest (B1)	PSE	Released Alive (B2)	PSE	Total Landings A+B1 (number)	Dead Discards B2 (number)
1981	1,540	56.5	805,317	65.9	128,652	26.2	806,857	25,730
1982	13,193	55.5	9,398	33.6	161,147	43.4	22,591	32,229
1983	14,579	50.4	29,826	48.4	294,107	21.1	44,405	58,821
1984	17,680	73.1	23,124	40.7	994,439	67.6	40,804	198,888
1985	24,512	86.4	34,792	55	167,371	32.5	59,304	33,474
1986	13,036	33	81,888	40.6	564,352	24.7	94,924	112,870
1987	64,431	78.1	64,119	50.6	373,458	42	128,550	74,692
1988	56,212	40.4	87,845	37.7	545,672	23.6	144,057	109,134
1989	49,649	57.6	72,777	28.3	794,579	28.5	122,426	158,916
1990	55,501	41.6	71,655	35.2	753,649	20.3	127,156	150,730
1991	81,441	29.6	53,394	35.9	1,040,163	18.4	134,835	208,033
1992	123,555	48.6	32,165	27.4	523,665	16	155,720	104,733
1993	38,093	34.3	40,403	42.4	778,604	19.7	78,496	155,721
1994	13,890	40.4	44,574	58.6	593,746	22.4	58,464	118,749
1995	19,030	30.4	16,562	47.2	356,311	25.3	35,592	71,262
1996	6,753	44	4,365	68.8	186,192	19.4	11,118	37,238
1997	31,872	48.1	12,055	70.1	487,269	20.3	43,927	97,454
1998	21,530	41.4	44,432	94.1	417,596	22.4	65,962	83,519
1999	21,757	63.3	13,231	74.5	362,473	19.7	34,988	72,495
2000	1,640	44	96	85.7	335,904	24.6	1,736	67,181
2001	6,751	56.3	3,352	68.5	1,153,341	12.5	10,103	230,668
2002	3,000	37.6	140,033	66.1	997,419	15	143,033	199,484
2003	15,581	42	8,584	56.6	1,584,326	14.1	24,165	316,865
2004	75,946	49.1	71,732	50.2	2,705,518	13.8	147,678	541,104
2005	8,811	41.4	10,001	42.8	1,983,774	19.3	18,812	396,755
2006	7,980	40.1	23,195	61.2	2,336,176	13.9	31,175	467,235
2007	3,319	62	48,365	63.3	2,413,174	14	51,684	482,635
2008	25,731	36.9	68,959	48.3	2,216,029	13.3	94,690	443,206
2009	9,216	42.2	33,972	39	2,885,331	14.8	43,188	577,066
2010	5,112	42	10,637	66.5	1,936,270	19.9	15,749	387,254
2011	16,750	39.9	17,716	54.7	2,372,432	15.8	34,466	474,486
2012	6,629	68.7	12,719	81.7	1,726,341	27.6	19,348	345,268
2013	20,326	56.2	55,131	73	4,803,736	19	75,457	960,747
2014	5,159	56.6	39,952	25.5	7,008,107	43	45,111	1,401,621
2015	9,173	56.7	16,379	62.9	1,711,330	22.3	25,552	342,266
2016	35,052	80.7	43,877	62.6	3,630,248	26.1	78,929	726,050
2017	19,524	60.8	35,806	37.4	1,435,399	20.9	55,330	287,080
2018	4604	69.8	16,864	53.1	1490265	19.5	21,468	298,053
2019	17352	52	6899	60.2	2318948	17.6	24,251	463,790

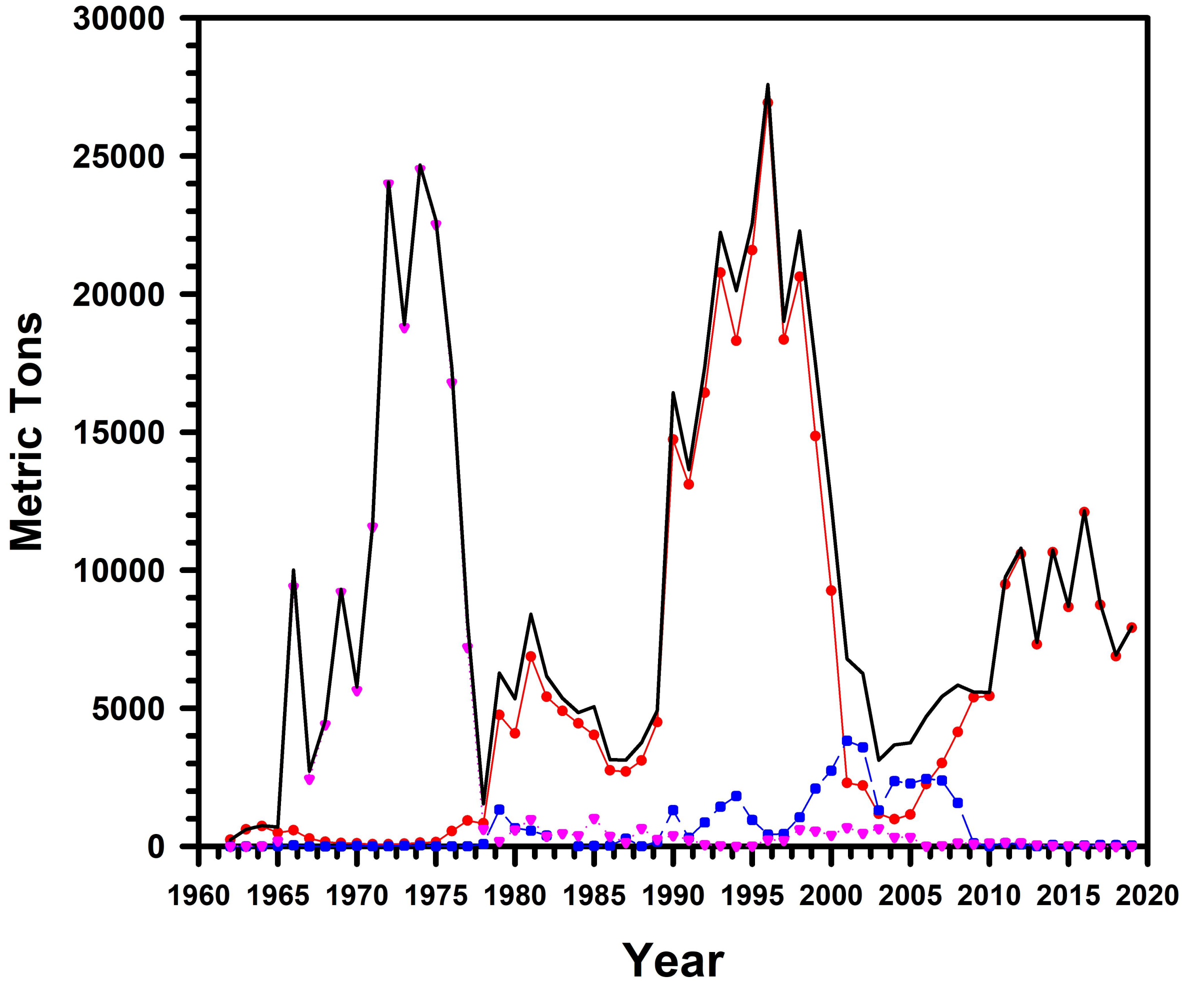


Figure 2.1. Commercial landings (metric tons) from the United States (red circles), Canada (blue squares) and other foreign (pink triangles) fleets and total landings (black solid line) in NAFO Subareas 2-7 from 1962-2019.

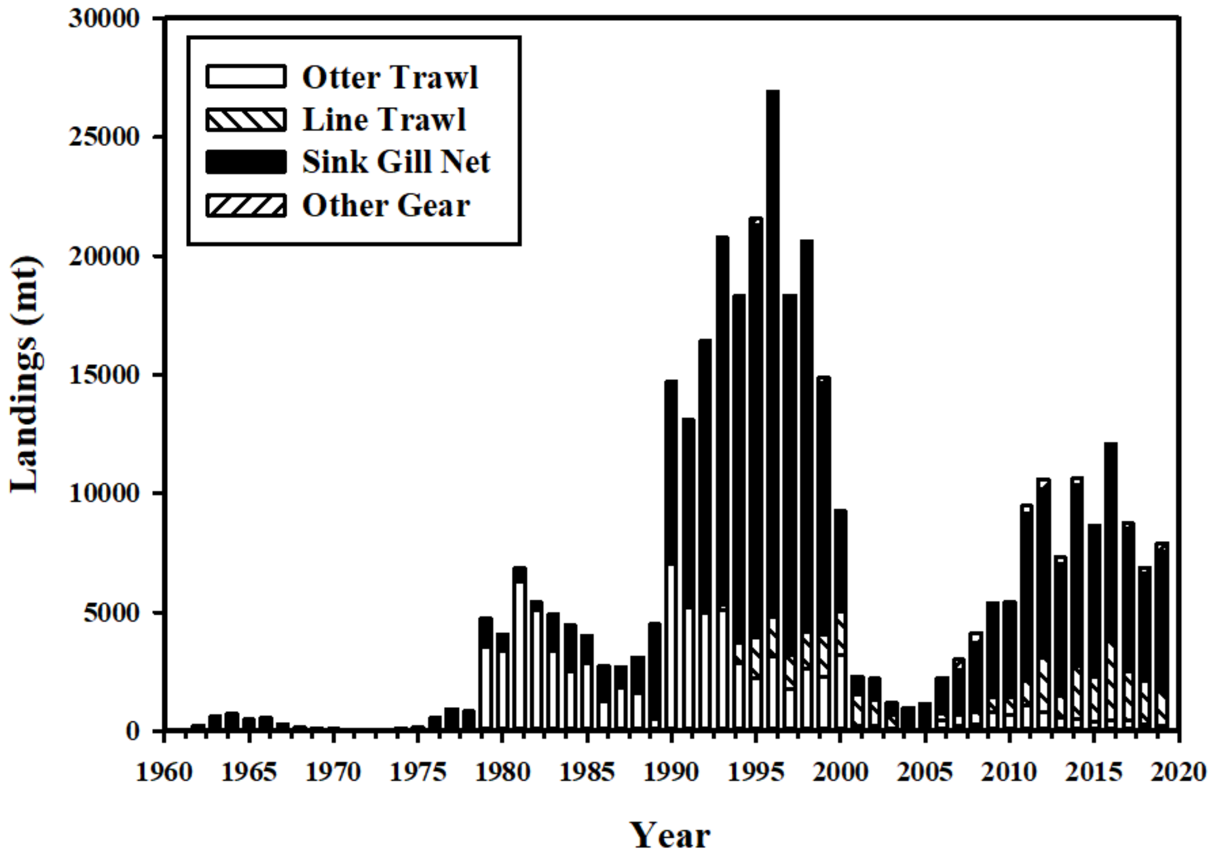


Figure 2.2. U.S. landings (metric tons) of spiny dogfish from NAFO subareas 2-7 by gear type, 1962-2019.

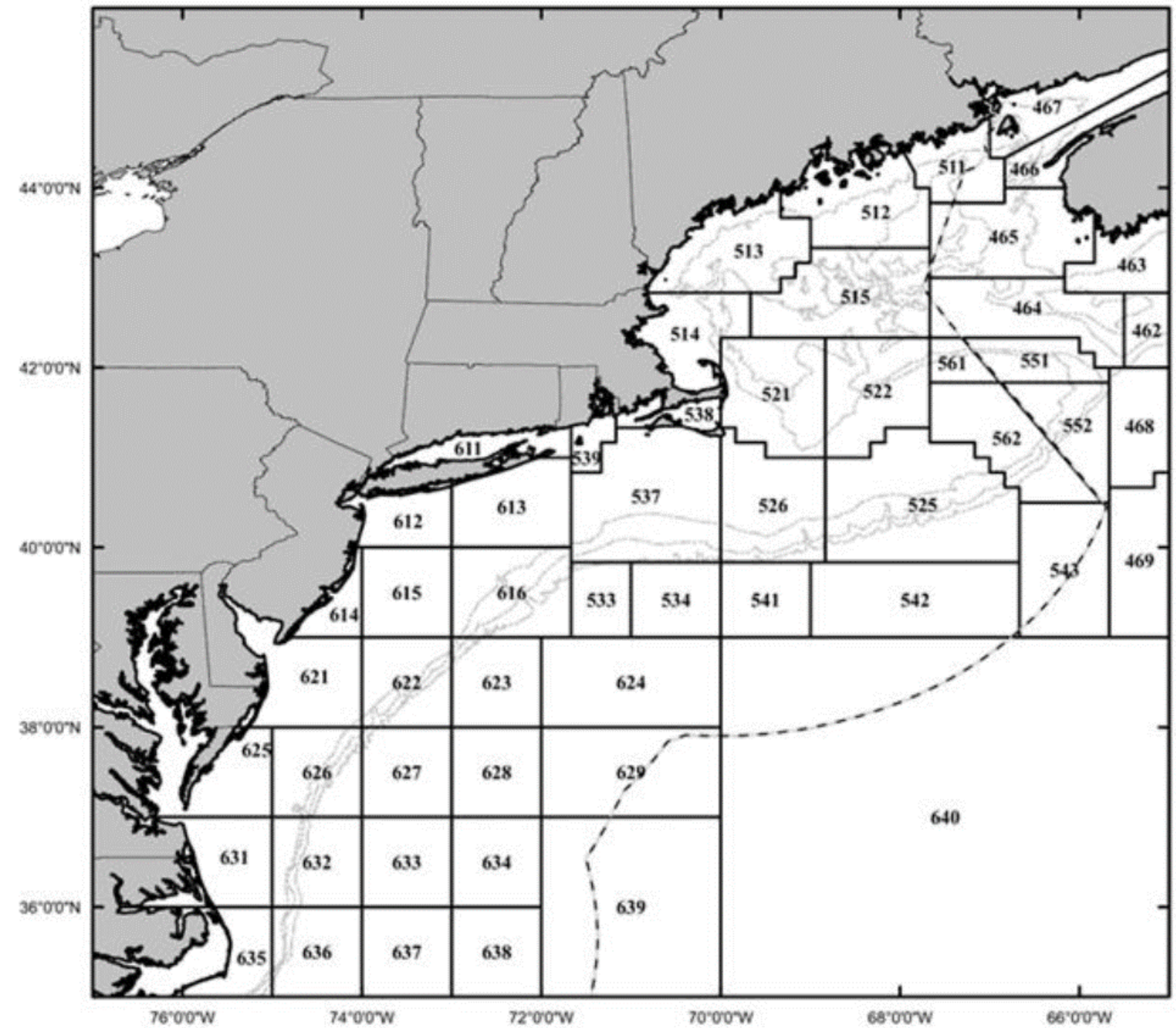


Figure 2.3 Map of fishing statistical areas as defined by the NOAA Fisheries.

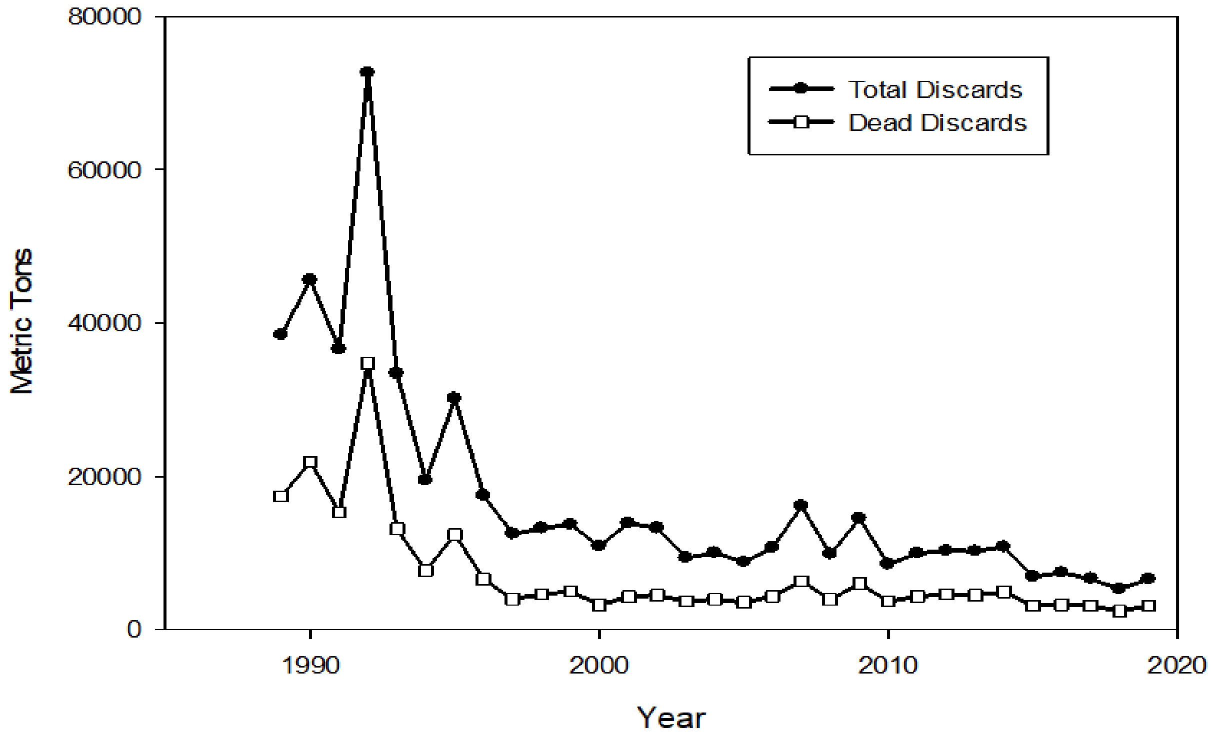


Figure 2.4. Total discards (closed circles) and dead discards (open squares) estimated for spiny dogfish using the methodology developed in this report from 1989-2019.

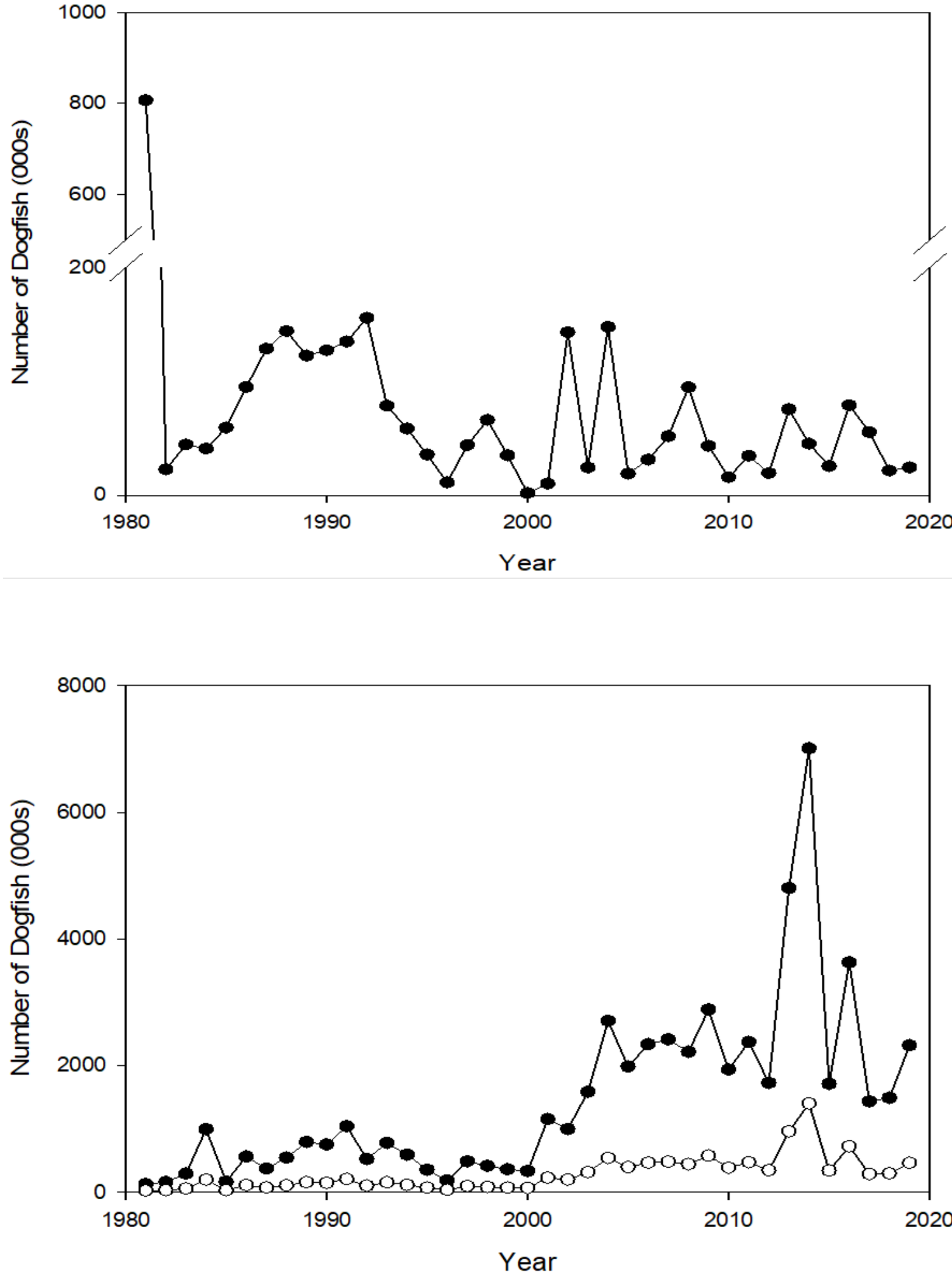
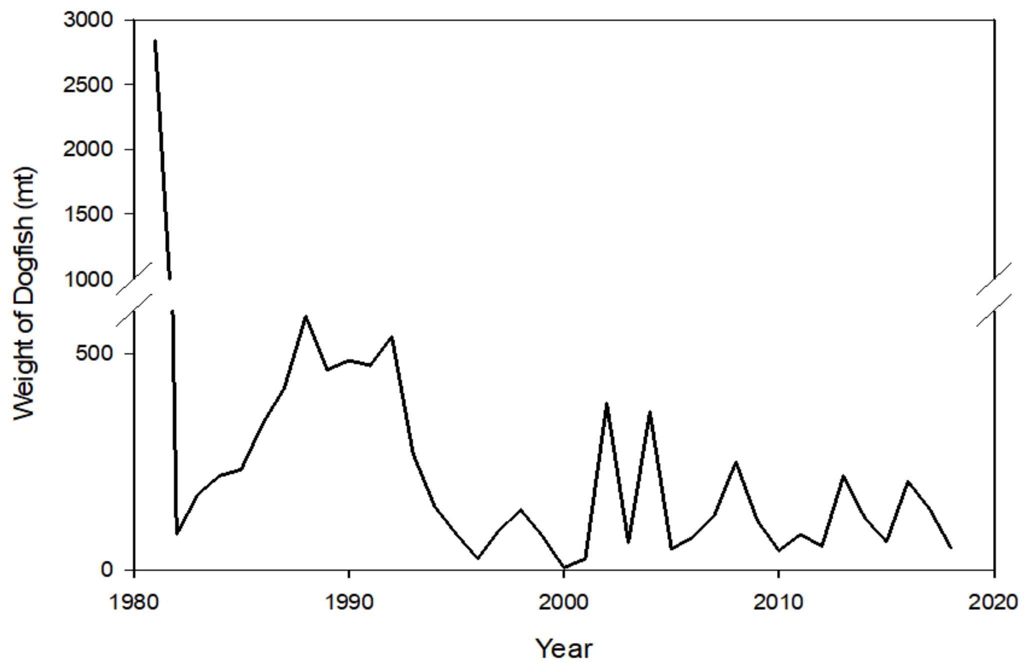


Figure 2.5. Estimates of recreational landings (top panel) and discards (bottom panel, total = black circles, dead = open circles) from 1981-2019 in number.



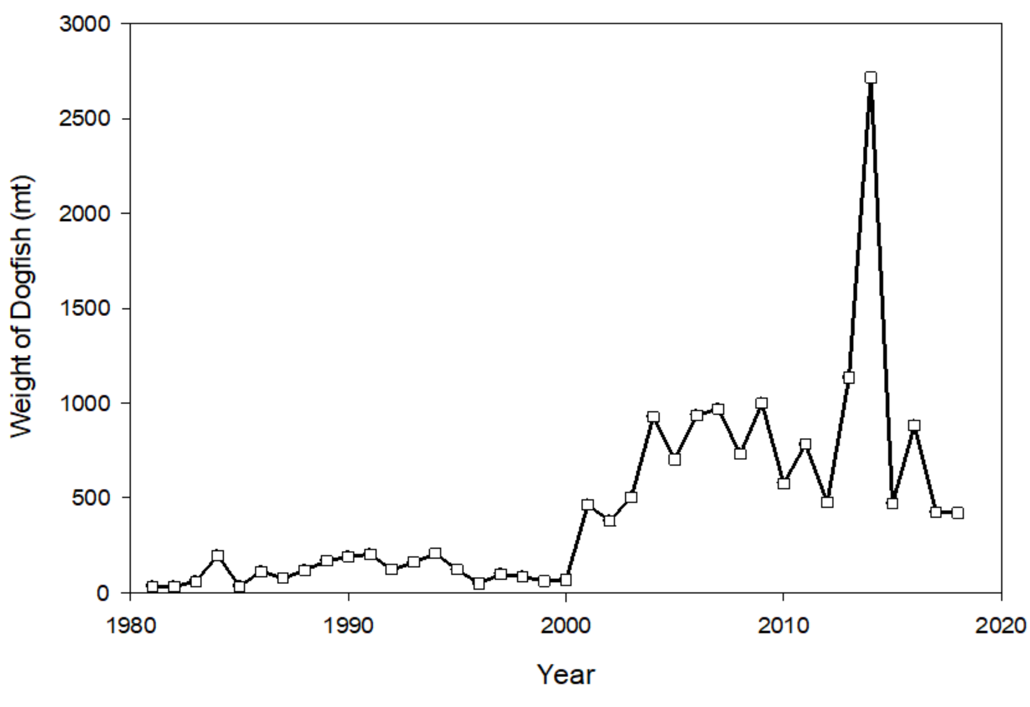


Figure 2.6. Estimates of recreational landings (top panel) and dead discards (bottom panel) from the new length-based method.

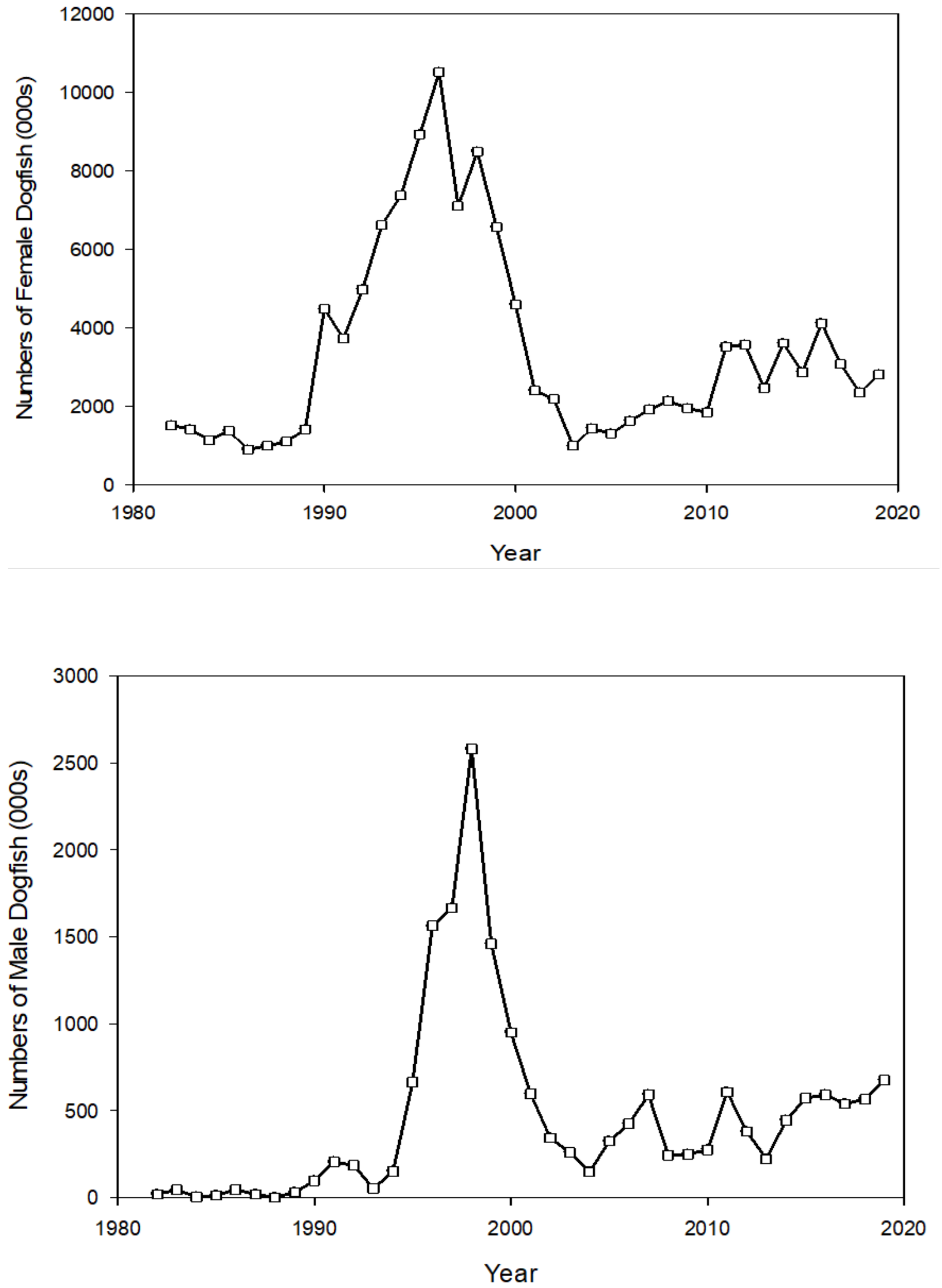


Figure 2.7. Estimates of total landings of females (top panel) and males (bottom panel) in 000s of fish from the new length-based method.

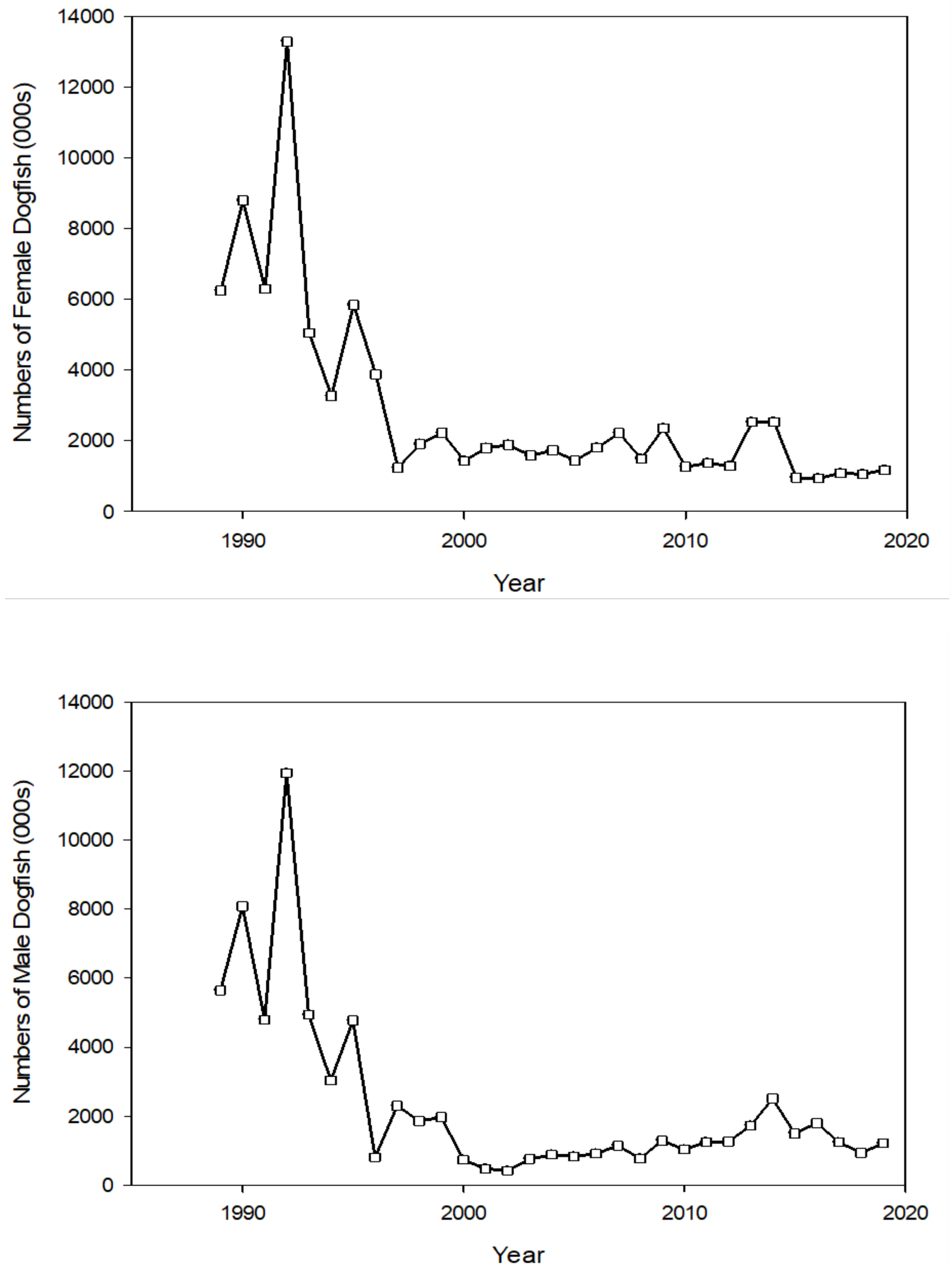


Figure 2.8. Estimates of dead discards of females (top panel) and males (bottom panel) in 000s of fish from the new length-based method.

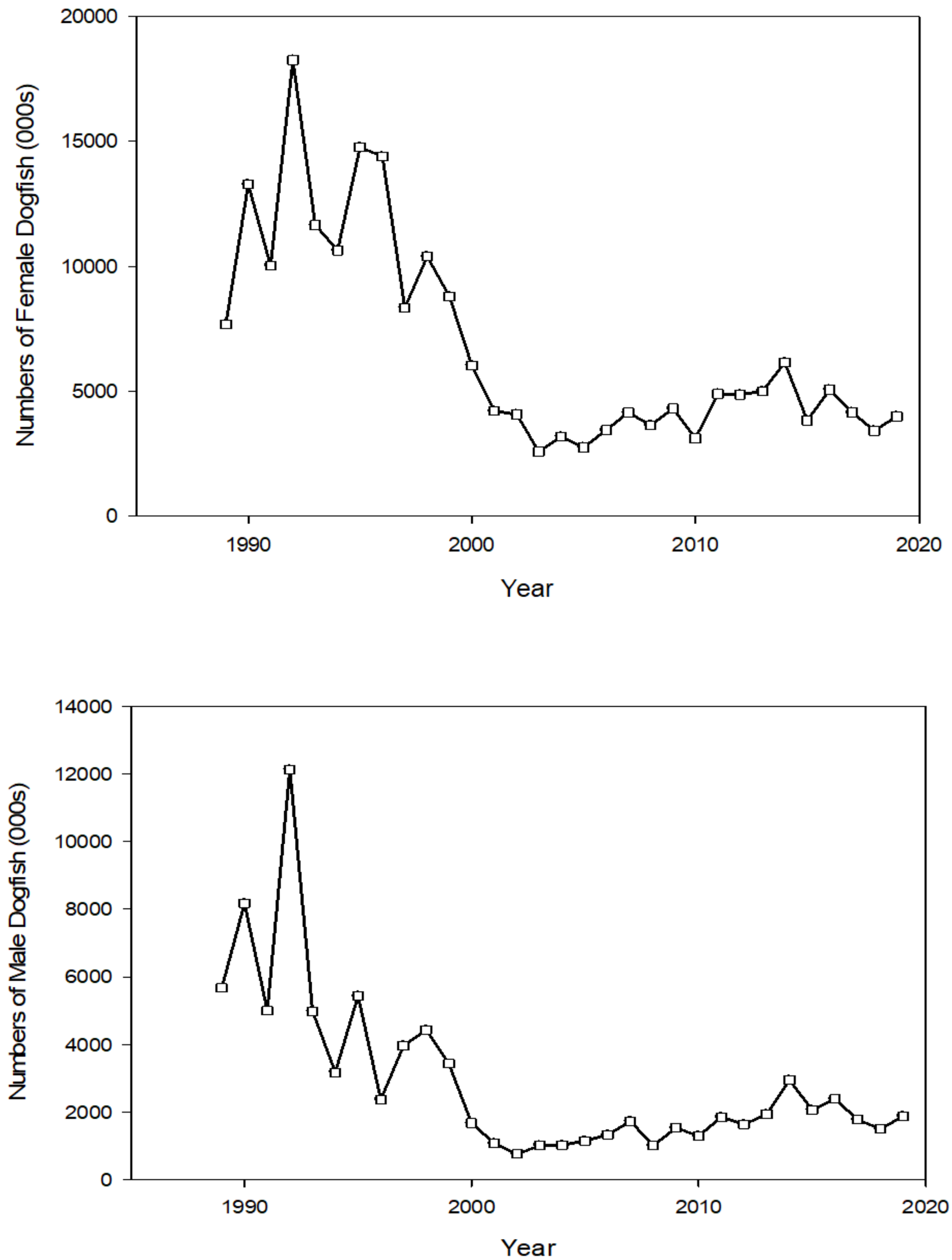


Figure 2.9. Estimates of total catch (landings plus dead discards) of females (top panel) and males (bottom panel) in 000s of fish from the new length-based method.

Figure 2.10. Catch rate (CPUE) trends through time for the nominal and standardized methods. The mean survey index (fall and spring combined) is shown in orange as well. The survey catches represent a combination of both male and female dogfish (similar to the CPUEs). The ribbon associated with each blue series approximates a confidence interval. Values are derived from the coefficient values for each year term in each model (Maunder and Punt 2004).

TOR3: SURVEY DATA

"*Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.*"

Fishery independent surveys considered for use in this research track assessment included NEFSC, state, and Canadian trawl surveys and the NEFSC bottom longline survey in the Gulf of Maine. The state surveys considered are more temporally and/or spatially limited when compared with the NEFSC Bottom Trawl Survey, which has the greatest spatial coverage. NEFSC Spring Bottom Trawl Survey data were used to estimate stock biomass in previous assessments and, once management measures went into place, to update biological reference points in between assessments. Studies comparing the seasonal relative abundance and distribution from NEFSC and Canadian trawl surveys indicated that the spring trawl surveys provide the best representation of spiny dogfish abundance in the northwest Atlantic (NEFSC 1994; Campana et al. 2007). Additionally, VAST estimates of encounter probability from 1980-2021, using four biannual trawl surveys, indicated higher encounter rates throughout the surveys' combined range during the spring (Figures 3.1 and 3.2; Hansell and McManus 2022). For these reasons, the Working Group recommended the NEFSC Spring Bottom Trawl Survey for use in the base run of this assessment and all other indices were reviewed for potential use in sensitivity runs.

NEFSC Surveys

Fall and Spring Bottom Trawl Surveys

The NEFSC has conducted both the fall and spring multispecies bottom trawl surveys annually since 1968 as a random stratified survey with coverage from the Gulf of Maine to Cape Hatteras, North Carolina (Figures 3.3 and 3.4). Exploratory analyses of survey data indicated inconsistent sampling in Gulf of Maine stratum 35 (Figure 3.3), including the splitting of the stratum into two sections in 1985 with sampling only occurring in the southern portion of the stratum. This stratum was eliminated from index development. Two vessels, the RV *Albatross IV* and the FRV *Henry B. Bigelow*, have conducted the majority of the surveys with the former vessel used prior to 2009 and the latter vessel used from 2009 to present. When the survey platform changed in 2009, stations less than 18 m in depth were excluded, eliminating many of the shallow inshore stations. Inshore strata retained for index development, given consistent sampling across platforms, were strata 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44-46, 56, 59-61, and 64-66 (Figure 3.4). Survey timing remained relatively consistent across years during the spring survey, but in the fall it tended to begin earlier in the season as the time series progressed (Figure 3.5). Sex was recorded for spiny dogfish caught during the survey starting in 1980. For details on changes in survey coverage, vessels, timing, design, and gear throughout the history of the survey see Johnston and Sosebee (2014).

The Working Group recommended application of the seasonal vessel calibration factors from Miller et al. (2010) to account for the vessel/gear change in 2009. Other available calibration factors were not applicable during this assessment process because the factors were not found to be significant or did not apply to the temporal or spatial scale of the survey used for this assessment. Relative abundance indices using mean numbers and weight per tow were developed for both the spring and fall survey by sex and combined sex from 1980 to 2021 (Tables 3.1 - 3.4). Design based total biomass estimates were developed for both the fall and spring surveys (Figures 3.6 and 3.7)

The Working Group recommended using the spring index in the base run and the fall index as a sensitivity. Both indices were also recommended for use in spatiotemporal habitat modeling (VAST) to explore distribution shifts and to develop an integrated survey index.

Winter Bottom Trawl Survey

The NEFSC initiated an offshore winter bottom trawl survey in 1992 to target flatfish and provide better estimates of their abundance than produced from the spring and fall surveys (Terceiro 2003). The winter bottom trawl survey ended in 2007 based on the new vessel (FRV *Henry B. Bigelow*) and gear changes planned for the spring and fall surveys likely improving flatfish catches (Johnston and Sosebee 2014). This survey was conducted in February and timing was consistent across years (Figure 3.8). Survey coverage generally ranged from Georges Bank to the mid-Atlantic, with consistent coverage only occurring off southern New England and the Mid-Atlantic (Figure 3.3, strata 1-12 and 61-76). Two different vessels were used to conduct this survey, but not during consistent time frames and no conversion factors were developed for the two vessel/gear combinations. For additional information on the survey design, coverage, and vessels see Johnston and Sosebee (2014). Both flatfish and elasmobranch (including spiny dogfish) catchability for this survey were high (NEFSC 2000, 2003). Stratified mean number per tow estimates for spiny dogfish declined across the time series in the regions consistently sampled (Figure 3.9).

The Working Group did not recommend the use of the winter survey index for this assessment due to the short time series, limited consistent spatial coverage, and lack of a conversion factor or consistent time frames for the different vessel/gear combinations.

Gulf of Maine Bottom Longline Survey

The NEFSC Gulf of Maine Bottom Longline Survey was initiated in 2014 and has occurred in the spring and fall concurrently with the NEFSC Bottom Trawl Survey. The NEFSC Bottom Trawl Survey cannot efficiently sample very complex, rough-bottom areas. This bottom longline survey was designed to increase sampling of several data-poor groundfish stocks that are associated specifically with rough-bottom habitat (McElroy et al. 2019). Survey coverage included six offshore strata in the Gulf of Maine: 26, 27, 37, and portions of 28, 29, and 36, all with sub-stratification by bottom type (Figure 3.10, Nieland and McElroy 2022). For more details on the gear and survey design see McElroy et al. (2019). Stratified mean numbers and weight per set were developed for spiny dogfish by season, bottom type, and sex for the survey from 2014 through 2021 (Figures 3.11 and 3.12). No significant differences were found between longline catches by sex or combined sex with bottom type based on an ANOVA test ($P<.05$; Nieland and McElroy 2022). Additionally, visual and regression analyses comparing combined bottom type longline and trawl indices from the same strata by sex and season indicated general agreement among survey trends with number derived indices showing better agreement (Nieland and McElroy 2022).

Lengths of spiny dogfish caught in the bottom longline and trawl surveys (only for the six strata covered by the longline survey) were compared by proportions at length (Figure 3.13). A Kolmogorov-Smirnov test was performed to determine if the proportions at length from the longline and trawl surveys by sex, season, and year came from the same distribution. The proportions at length for females during the spring surveys in 2017, 2018, and all years combined and the fall surveys in 2014, 2018, and all years combined had significantly different length distributions ($P<.05$, Nieland and McElroy 2022).

The Working Group recommended sensitivity runs using the stratified mean numbers per set index for the spring and fall longline surveys with combined bottom types.

U.S. State and Interstate Fishery Independent Surveys

ASMFC Northern Shrimp Trawl Survey

The ASMFC Northern Shrimp Trawl Survey is a random stratified bottom trawl survey that began in 1983 with limited sampling in the first year. The survey covers Gulf of Maine waters stratified by depth and area with core coverage in strata 04010, 04030, and 04050-04080 each year except the initial survey year (Figure 3.14; Johnston and Sosebee 2014). The survey takes about two weeks to complete and is conducted during the summer months anytime between July and August with timing trending earlier in the year across the survey time frame (Figure 3.15). For details on survey design and gear see Johnston and Sosebee (2014). Stratified mean numbers and weight per tow indices show an increasing trend with high variability in recent years, primarily driven by males (Figure 3.16).

The Working Group did not recommend this index for use given the timing of the survey, as the NEFSC Spring Bottom Trawl Survey may account for some of these fish before they migrate into the Gulf of Maine. Additionally, the large increase in abundance in the later years could be partially attributed to the gradual shift in survey timing to later in the summer or warming ocean temperatures altering migration timing.

NEAMAP Trawl Survey

The Northeast Area Monitoring and Assessment Program (NEAMAP) Trawl Survey began sampling the coastal ocean from Martha's Vineyard, Massachusetts, to Cape Hatteras, North Carolina, since the fall of 2007 (Figure 3.17). The survey area is stratified by latitudinal/longitudinal region and depth. A four-seam, three-bridle, 400x12 cm bottom trawl is towed for 20 minutes at each sampling site with a target speed-over-ground of 3.0 kts. The net is outfitted with a 2.54 cm knotless nylon liner to retain the early life stages of the various fishes and invertebrates sampled by the trawl. The survey conducts two cruises a year, one in the spring (April-May) and one in the fall (September-November). NEAMAP catches mainly adult spiny dogfish, although some years and seasons also encounter juveniles based on the length frequencies (Figure 3.18. and 3.19). Female and male spiny dogfish were caught more often in the spring (77% and 33% positive tows, respectively) than in the fall (52% and 15% positive tows, respectively).

After reviewing the geometric means provided by NEAMAP, nominal and model based indices were developed for this survey by sex and season (Figures 3.20 and 3.21). Model based indices explored used a variety of generalized models. A full model that predicted catch as a linear function of year, water temperature, salinity, dissolved oxygen, depth, depth stratum, and station was compared with nested submodels using AIC. Based on several diagnostics (AIC, dispersion, percent deviance explained, and resulting coefficients of variation), the model chosen was a negative binomial that included year and station for females in the spring and year and depth strata for males in the spring (Figure 3.21). For females in the fall, the model chosen was a negative binomial that included year, temperature, and depth strata and for males in the fall, year and depth strata (Figure 3.21). Fall nominal and model based indices for both sexes indicate that the survey does not encounter spiny dogfish regularly with only a few peaks in the time series, notably in 2016 (Figures 3.20 and 3.21).

The Working Group recommended that seasonal indices for this survey be used in spatiotemporal habitat modeling (VAST) to explore distribution shifts and to develop an integrated survey index.

MADMF Inshore Bottom Trawl Survey

The Massachusetts Division of Marine Fisheries (MADMF) began a biannual (spring and fall) bottom trawl survey in 1978 in coastal state waters. The survey area is stratified by both bio-geographic region and depth (Figure 3.22). A ¾ Yankee trawl net is used with a 39 ft headrope, 51 ft footrope, 0.25 in codend, 3.5 in cookie sweep, low aspect Tomkiewicz doors (wooden, 325 lb; 72x40 in), 63 ft of ⅜ chain in bottom legs, and 60 ft of ⅜ wire in top legs. The net is towed for 20 minutes at each sampling site with a target speed-over-ground of 2.5 kts. Two vessels have been used to conduct this survey, the F/V *Frances Elizabeth* from 1978 - 1981 and the R/V *Gloria Michelle* from 1982 to present. MADMF catches mainly large juvenile and adult spiny dogfish, although young juveniles are encountered based on the length frequencies (Figure 3.23. and 3.24).

Abundance (mean numbers per tow) and biomass (kg per tow) indices for spiny dogfish from Massachusetts spring and fall inshore bottom trawl surveys were developed for 1980-2021 (Figures 3.25 and 3.26). The spring survey usually occurs before the major influx of dogfish to Massachusetts waters. In the fall, catches tend to be an order of magnitude larger, as much of the dogfish stock is concentrated near the Massachusetts coast. Wide variations in availability result in highly variable survey indices. High variability in this survey is also a reflection of the seasonal use by dogfish of the area surveyed.

The Working Group recommended that seasonal indices for this survey be used in spatiotemporal habitat modeling (VAST) to explore distribution shifts and to develop an integrated survey index.

ME-NH Inshore Groundfish Trawl Survey

The Maine-New Hampshire Inshore Groundfish Trawl Survey is a biannual (spring and fall) random stratified survey by depth (5-20, 21-35, 36-55, and 55+ fathoms) and area based on geologic, oceanographic, geographic, and biologic factors that started in 2000. Sex data was not recorded until 2005. Survey coverage is in the coastal waters along the Maine and New Hampshire coast within the Gulf of Maine (Figure 2.27). For details on survey design and gear see Sherman et al. (2005).

Abundance (mean numbers per tow) and biomass (kg per tow) indices for spiny dogfish from the ME-NH spring and fall inshore bottom trawl surveys were developed for 1980-2021 (Figures 3.28 and 3.29). Similar to what was seen in the MADMF survey, catches were greater in the fall than in the spring, with the exception of a notable peak in mean numbers per tow during the spring in 2016. This peak was not seen in the weight per tow plot and can be explained by the size distribution of the catches during the spring in 2016, which was skewed towards young-of-the-year sized fish (Figure 3.30). There also appears to be an overall declining trend during the fall survey, although there is high interannual variability across the time series. There is no discernable trend during the spring season. These trends are also apparent in the mean catch at length plots (Figures 3.30 and 3.31).

The Working Group recommended that seasonal indices for this survey be used in spatiotemporal habitat modeling (VAST) to explore distribution shifts and to develop an integrated survey index.

Rhode Island Coastal Trawl Surveys

The Rhode Island Department of Environmental Management's Division of Marine Fisheries conducts two coastal bottom trawl surveys in Rhode Island waters, the Monthly (1990-present) and Seasonal (1979 - present) Trawl Surveys. The Monthly Trawl Survey has 13 fixed stations located in Narragansett Bay (12) and in Rhode Island Sound (1) surveyed in the middle of each month (Figure 3.32). The Seasonal Survey occurs during the Spring (April-May) and Fall (September-October) with a combination of fixed (12) and random (14) stations in Narragansett Bay and 18 fixed stations in Rhode Island and Block Island Sound (Figure 3.33). For details on survey design and gear see Parkins and Olszewski (2021).

The majority of spiny dogfish encountered during the Rhode Island bottom trawl surveys are over 75 cm stretched total length with a female to male ratio of approximately 4:1. Abundance (mean numbers per tow) and biomass (kg per tow) indices for spiny dogfish for both spring and fall seasonal and the monthly trawl surveys were developed for 1979-2021 and 1990-2021, respectively (Figures 3.34 and 3.35). Catches were low throughout the time series for each survey with the exception of a peak in the mid-2000s that was associated with high coefficients of variation (Figures 3.34 and 3.35).

The Working Group did not recommend this index for use given the low encounter rates and the limited spatial coverage of the surveys.

Canada DFO Bottom Trawl Surveys

Canada Department of Fisheries and Oceans (DFO) Bottom Trawl surveys were designed to provide abundance trends for fish and invertebrates and use Northwest Atlantic Fisheries Organization (NAFO) Divisions to define area coverage (Figure 3.36).

Scotian Shelf Trawl Survey (NAFO Divisions 4VWX)

The Canada DFO Scotian Shelf survey was initiated in 1970 as a summer survey with coverage in NAFO Divisions 4VWX (Figure 3.36). For information on survey design, gear, and vessels see Fowler and Showell (2009) and DFO (2020). The design based biomass index developed for this survey shows high inter-annual variability with an increasing trend in catches through 2002 followed by a decreasing trend for the remainder of the time series (Figure 3.37). Although sex data was not available at the time of analyses, previous assessment reports indicated that the adult females were not encountered on this summer survey (TRAC 2010) and the male biomass was nearly 2.8 times greater than female biomass estimated from this survey (NEFSC 2006). Length frequency data indicate no young of the year caught during this survey, only larger juveniles and likely adult males (Figure 3.38)

The Working Group did not recommend this index for use in the assessment due to the high inter-annual variability. The Working Group also cited the need to review the catch per set information by sex and combined, which was not possible during the assessment time frame. Spatiotemporal modeling would also be beneficial in the future to investigate potential shifts in distribution or migration timing.

Eastern Georges Bank (NAFO Division 5Ze)

The Canada DFO Eastern Georges Bank survey was initiated in 1987 as a winter (February) survey with coverage in NAFO Division 5Ze (Figure 3.36). For information on survey design, gear, and vessels see Stone and Gross (2012). The design based biomass index developed for this survey shows a steep increase with high interannual variability followed by a sharp decline and in the mid 1990s remaining at low levels until drops to zero in 2003 and basically stays there except for a minor blip in 2008 (Figure 3.39). Comparison to NEFSC Spring Bottom Trawl Survey data in the same region shows a similar trend with the drop in the mid 1990s but with more variability after it drops off (Figure 3.40).

The Working Group did not recommend this index for use in the assessment model given they would need to review the catch per set information by sex and combined sex for this assessment, data which were not available during the assessment time frame and the trend from this survey is already seen within the NEFSC Spring Bottom Trawl Survey data. The Working Group did highlight a future need for spatiotemporal modeling to help determine what is behind the declining trends seen in this region.

Southern Gulf of St Lawrence

The Canada DFO Southern Gulf of St Lawrence survey was initiated in 1971 as an annual survey conducted each September (Figure 3.41). For information on survey design, gear, and vessels see Hurlbut and Clay (1990). Abundance (mean numbers per tow) and biomass (kg per tow) indices (Figures 3.42 and 3.43) and spatiotemporal plots of biomass (Figure 3.44) were developed. There were no spiny dogfish catches during the first 12 years of the survey and then there was a large spike in the late 1980s (Figures 3.42 and 3.43). This was followed by a decline with high inter-annual variability until spiny dogfish disappeared from the survey again in 2003 (Figures 3.42 and 3.43). These trends are similar to what was seen in Canadian and U.S. surveys on eastern Georges Bank (Figures 3.39 and 3.40).

Campana et al. (2007) reported on this abrupt appearance of spiny dogfish in the southern part of the Gulf of St. Lawrence suggesting it is a sink population and that there had been no immigration or recruitment, slowed individual growth due to the colder temperatures, and a gradual reduction in numbers.

The Working Group did not recommend this index for use in the assessment model, but did highlight the need to do some future spatiotemporal modeling to help determine what is behind the declining trends seen in this region.

Spring Grand Banks (NAFO Divisions 3LNOP)

The Canada DFO Spring Grand Banks survey was initiated in 1996 with coverage in NAFO Divisions 3LNOP (Figure 3.36). For information on survey design, gear, and vessels see Rideout and Ings (2020). Abundance and biomass indices for spiny dogfish were developed for the Spring Grand Banks Survey and catch distribution was plotted for the last year in the time series (Figure 3.45). Catches were low across the time series with an increase during the last few years of the survey. The estimates at the end of the time series had large error bars and high inter-annual variability.

The Working Group did not recommend these indices for use in the assessment given the low encounter rates throughout the majority of the time series and the uncertainty in the estimates in recent years. The Working Group did highlight a future need for spatiotemporal modeling to help determine what is behind the increasing trend seen at the end of the time series.

Fall Grand Banks and Labrador (NAFO Divisions 2HJ3KVLNO)

The Canada DFO Fall Grand Banks and Labrador survey was initiated in 1995 with coverage in NAFO Divisions 2HJ3KVLNO (Figure 3.36). For information on survey design, gear, and vessels see Rideout and Ings (2020). Abundance and biomass indices for spiny dogfish were developed for the Fall Grand Banks Survey and catch distribution was plotted for the last year in the time series (Figure 3.46). As seen in the spring survey, catches were low across the time series with an increase during the last few years of the survey. The estimates at the end of the time series had large error bars and high inter-annual variability.

The Working Group did not recommend these indices for use in the assessment due given the low encounter rates throughout the majority of the time series and the uncertainty in the estimates in recent years. The Working Group did highlight a future need for spatiotemporal modeling to help determine what is behind the increasing trend seen at the end of the time series.

Integrated Survey Indices

A model based approach to deriving a spring index of abundance and length composition was pursued by the Working Group with two objectives: account for survey or environmental considerations that may influence catchability, and integrate multiple surveys into a single index to better describe the population. Spatiotemporal models have the ability to account for spatial shifts and can yield more precise/accurate indices (Shelton et al. 2014). Fitting assessments to these models can also lead to less retrospective bias and outperform assessments with design-based indices (Cao et al. 2017). Previous research has shown how diel effects can influence spiny dogfish catch from fisheries-independent surveys (Sagarese et al. 2016), warranting evaluation of a model-based index approach for inclusion in the assessment model.

A VAST model was developed to both include explanatory covariates and integrate survey information. As described in TOR1, the VAST model represents a delta-model that predicts the probability of an encounter and the positive catch rate as two separate generalized linear mixed models. A Bernoulli distribution was assumed for probability of a positive catch and a Poisson distribution for positive catch. Time of day, bottom temperature and depth associated with each tow were explored as covariates. For both the spring and fall model configurations, AIC, and model diagnostics supported including depth as a modulate of density (Hansell and McManus 2022). In deriving the single model-based index of abundance, the VAST model incorporated data from four biannual trawl surveys (Figure 3.47): the Northeast Fisheries Science Center (1980 – 2021); Massachusetts Division of Marine Fisheries (1980 – 2021); Maine/New Hampshire (2005 – 2021); and Northeast Area Monitoring and Assessment Program (2007 – 2021; Hansell and McManus 2022).

In the spring, encounter probability and abundance are high in the mid-Atlantic (Figures 3. 1 and 3.48). In contrast, in the fall encounter probability and abundance are estimated to be lower in the mid-Atlantic and higher in the Gulf of Maine (Figures 3.2 and 3.49). For the spring and fall, VAST estimates of relative abundance for male and female dogfish are similar to the NEFSC designed based estimates (Figures 3.50 - 3.53). In the spring, VAST estimates differ from designed based estimates for the inshore surveys (menh, madmf, and neamap; Figures 3.50

and 3.51). In the fall, VAST estimates were more similar to design based estimates from the inshore surveys (Figures 3.52 and 3.53)

A multivariate VAST model was fit to length to produce standardized length composition data. The model fit to spring inshore (Maine-New Hampshire, Massachusetts Division of Marine Fisheries, and NEAMAP) and offshore surveys (NEFSC). The model failed to converge using 3 cm length bins so length bins were increased to 6 cm bins and the model successfully converged (Figures 3.54). The model estimates of length composition are similar to design based estimates of length composition (Figures 3.55 and 3.56).

The Working Group recommended using the spring VAST index as a sensitivity run in the model.

Table 3.1. Annual NEFSC Fall Bottom Trawl Survey mean numbers per tow and coefficients of variation (CV).

Year	Female	CV	Male	CV	Unsexed	CV	Total	CV
1980	3.83	72.81	1.35	60.38	0.03	46.44	5.21	58.64
1981	38.65	68.16	35.03	76.73	0.02	54.20	73.70	71.91
1982	6.68	38.92	6.69	43.74	0.00		13.37	39.74
1983	18.01	65.81	13.75	56.77	0.00	100.00	31.76	61.69
1984	14.51	38.63	10.59	30.73	0.00		25.11	34.63
1985	20.17	36.33	17.96	48.37	0.08	78.70	38.21	40.71
1986	15.22	29.94	12.43	28.43	0.00		27.65	27.36
1987	16.77	49.40	16.31	48.95	0.00		33.09	49.07
1988	13.60	23.98	10.47	22.72	0.00		24.07	22.44
1989	5.43	29.20	6.53	27.70	0.00		11.95	26.91
1990	11.72	34.74	13.91	44.10	0.00		25.63	35.58
1991	13.77	43.96	19.37	36.25	0.00		33.13	37.14
1992	25.85	33.50	12.47	38.08	0.00		38.32	30.09
1993	4.10	47.84	4.36	34.98	0.00		8.46	35.30
1994	9.06	33.17	11.65	37.56	0.00		20.71	34.17
1995	7.36	27.79	12.92	23.18	0.00		20.28	23.62
1996	19.80	70.86	12.93	58.97	0.00		32.73	66.00
1997	9.57	25.51	15.02	40.51	0.00		24.59	30.72
1998	16.76	42.22	10.50	26.50	0.00		27.26	34.81
1999	8.12	18.46	8.98	12.88	0.17	100.00	17.28	13.74
2000	5.25	22.79	11.37	40.94	0.00		16.63	33.85
2001	21.69	31.65	12.34	33.89	0.00		34.03	27.41
2002	15.48	29.23	15.10	37.73	0.00		30.59	31.72

2003	6.76	31.11	5.65	26.42	0.00		12.41	24.69
2004	16.45	22.10	17.32	19.89	0.01	100.00	33.78	18.85
2005	8.18	31.45	24.41	28.26	0.00		32.59	27.35
2006	20.34	25.97	26.57	25.68	0.00		46.91	22.84
2007	18.29	43.29	22.24	22.93	0.00		40.54	31.30
2008	13.22	19.58	18.11	20.14	0.00		31.33	18.82
2009	19.74	34.87	24.31	20.66	0.00		44.04	24.08
2010	24.07	35.36	24.03	28.50	0.00		48.10	32.37
2011	24.18	41.61	32.20	29.51	0.00		56.38	34.78
2012	60.37	19.62	62.27	20.53	0.00		122.64	20.14
2013	50.17	27.89	59.12	24.20	3.69	83.63	112.97	25.78
2014	28.93	39.32	36.74	34.71	0.00		65.67	36.85
2015	23.51	29.81	16.77	31.79	0.00		40.28	29.11
2016	20.95	40.54	35.61	35.05	12.91	100.00	69.47	35.55
2017								
2018	17.96	24.99	17.87	20.94	0.00		35.84	23.33
2019	26.32	24.90	43.43	26.20	0.00		69.75	24.19
2020								
2021	9.20	25.50	20.37	26.42	0.00		29.57	24.55

Table 3.2. Annual NEFSC Fall Bottom Trawl Survey mean weight (kg) per tow and coefficients of variation (CV).

Year	Female	CV	Male	CV	Unsexed	CV	Total	CV
1980	16.48	83.02	2.43	64.94	0.03	60.99	18.94	75.11
1981	34.64	55.87	12.39	33.47	0.02	81.40	47.04	44.21
1982	9.69	53.79	5.10	35.13	0.00		14.78	44.32
1983	23.43	71.96	13.08	61.85	0.00	100.00	36.50	68.12
1984	25.45	51.56	9.05	32.89	0.00		34.50	45.92
1985	26.93	35.54	13.20	42.90	0.09	73.24	40.22	33.35
1986	24.45	39.38	13.55	28.73	0.00		38.00	33.06
1987	13.14	33.09	10.64	41.29	0.00		23.79	36.18
1988	17.12	26.04	9.59	20.18	0.00		26.71	21.07
1989	4.85	22.31	5.89	30.64	0.00		10.74	23.19
1990	17.15	46.30	14.07	51.91	0.00		31.22	35.93
1991	23.29	46.51	24.37	31.69	0.00		47.66	35.23
1992	41.17	35.02	13.68	51.95	0.00		54.85	32.84
1993	5.66	59.91	4.97	32.73	0.00		10.63	37.63
1994	8.31	31.53	12.88	41.11	0.00		21.19	35.59
1995	5.23	21.46	12.98	22.36	0.00		18.21	20.90
1996	26.62	69.67	14.78	60.49	0.00		41.40	66.20
1997	9.10	20.30	16.27	53.42	0.00		25.37	37.25
1998	25.69	41.77	12.42	25.01	0.00		38.11	34.76
1999	12.06	20.25	12.21	12.63	0.28	100.00	24.55	13.80
2000	8.85	22.02	17.21	41.47	0.00		26.07	33.06
2001	32.57	30.10	15.43	32.60	0.00		48.00	26.51
2002	26.00	28.11	21.22	39.77	0.00		47.22	31.57
2003	13.56	34.33	8.09	26.54	0.00		21.65	26.91

2004	29.25	20.28	22.71	18.06	0.00	100.00	51.96	15.42
2005	14.53	31.33	36.29	29.33	0.00		50.82	28.00
2006	37.61	26.31	36.72	24.40	0.00		74.33	21.55
2007	33.87	40.60	32.55	23.14	0.00		66.43	31.37
2008	19.16	15.82	20.95	19.72	0.00		40.11	16.07
2009	26.63	39.74	28.93	22.02	0.00		55.56	26.87
2010	22.86	30.09	18.31	15.60	0.00		41.17	22.61
2011	13.37	21.80	27.04	22.78	0.00		40.41	22.09
2012	60.43	24.04	54.78	25.26	0.00		115.21	22.92
2013	20.25	19.04	35.29	17.58	4.11	77.67	59.65	17.84
2014	15.77	34.17	27.16	33.16	0.00		42.93	32.00
2015	33.80	48.98	15.28	37.15	0.00		49.08	39.84
2016	22.71	42.78	43.39	36.10	12.00	100.00	78.10	35.77
2017								
2018	13.35	22.21	14.95	20.09	0.00		28.30	20.10
2019	19.12	20.97	44.31	31.89	0.00		63.42	26.91
2020								
2021	5.22	19.32	22.15	34.66	0.00		27.37	31.20

Table 3.3. Annual NEFSC Spring Bottom Trawl Survey mean numbers per tow and coefficients of variation (CV).

Year	Female	CV	Male	CV	Unsexed	CV	Total	CV
1980	13.63	22.38	17.46	24.44	8.29	68.61	39.38	22.71
1981	31.26	20.58	24.79	21.97	0.63	69.32	56.68	19.81
1982	27.09	33.38	23.05	27.50	0.00		50.15	27.95
1983	17.71	22.60	22.91	18.01	0.01	100.00	40.62	19.13
1984	9.31	18.66	12.95	36.27	0.00		22.26	21.96
1985	36.41	29.33	77.83	33.79	0.00	100.00	114.24	26.96
1986	19.14	15.44	9.17	28.67	0.00		28.31	18.91
1987	24.92	24.85	37.80	37.69	0.00		62.72	32.28
1988	35.26	28.18	28.39	38.38	0.02	77.04	63.67	28.12
1989	26.35	19.14	28.35	29.44	0.00		54.70	21.43
1990	43.00	31.12	46.05	52.59	0.00		89.05	41.80
1991	29.57	18.04	31.10	28.51	0.00		60.67	20.66
1992	39.42	24.05	36.02	21.00	0.00		75.44	18.02
1993	27.40	16.56	31.34	45.86	0.00		58.74	30.81
1994	36.80	17.31	51.30	16.66	0.00		88.10	15.85

1995	24.29	22.67	24.65	19.31	0.00		48.94	16.59
1996	42.91	37.92	50.49	26.50	0.00		93.41	29.58
1997	28.26	16.06	28.62	18.06	0.00		56.87	15.89
1998	11.10	20.30	31.37	25.13	0.00		42.47	22.33
1999	20.22	15.46	33.41	18.02	0.00		53.63	16.13
2000	15.00	31.58	21.42	26.54	0.27	100.00	36.69	25.03
2001	10.57	33.94	19.58	27.38	0.00		30.15	28.41
2002	19.66	19.84	31.70	18.61	0.00		51.36	17.95
2003	17.75	13.46	31.42	15.33	0.00		49.17	12.74
2004	10.06	26.93	17.69	26.96	0.00		27.75	25.94
2005	10.23	30.13	36.51	48.97	0.00		46.73	43.60
2006	27.86	21.53	48.70	28.77	0.06	100.00	76.61	23.92
2007	17.11	25.03	27.46	16.84	0.00		44.57	16.53
2008	24.12	14.20	35.87	13.49	0.00		59.99	10.08
2009	24.76	21.09	44.78	17.78	0.00		69.54	18.02
2010	19.48	16.37	36.50	19.88	0.00		55.98	17.30
2011	23.36	23.15	51.69	15.15	0.00		75.05	16.54

2012	47.60	21.65	85.55	37.33	0.00	76.38	133.15	28.39
2013	59.94	43.08	86.03	29.82	0.01	72.43	145.98	35.43
2014								
2015	15.43	28.73	34.95	21.28	0.00	100.00	50.39	20.18
2016	35.48	19.06	60.21	20.89	0.00	100.00	95.70	20.25
2017	18.78	21.39	38.99	14.03	0.00		57.77	16.03
2018	28.71	28.22	42.39	19.68	0.08	56.09	71.18	22.82
2019	39.85	34.91	67.18	12.37	0.00		107.03	17.37
2020								
2021	43.23	18.72	87.42	13.78	0.00		130.65	15.14

Table 3.4. Annual NEFSC Spring Bottom Trawl Survey mean weight (kg) per tow and coefficients of variation (CV).

Year	Female	CV	Male	CV	Unsexed	CV	Total	CV
1980	28.06	18.72	22.13	25.72	18.95	63.03	69.14	21.71
1981	67.85	18.22	28.76	20.37	0.91	35.47	97.53	15.56
1982	83.68	43.71	30.21	28.94	0.00		113.89	36.27
1983	17.82	15.89	20.50	15.96	0.00	100.00	38.33	12.10
1984	23.78	18.77	18.77	38.75	0.00		42.55	19.15
1985	65.73	41.24	97.43	38.37	0.00	100.00	163.16	28.73
1986	38.80	12.98	5.65	32.76	0.00		44.45	14.12
1987	59.66	45.12	39.04	43.89	0.00		98.70	44.28
1988	77.85	40.12	25.94	44.18	0.03	83.60	103.81	35.28
1989	42.18	17.50	33.31	36.97	0.00		75.49	22.11
1990	87.64	26.84	58.40	56.25	0.00		146.04	36.71
1991	52.93	20.47	35.19	33.53	0.00		88.12	22.09
1992	67.39	28.81	42.17	23.61	0.00		109.56	21.41
1993	50.56	17.88	34.34	43.65	0.00		84.90	23.29
1994	34.27	18.36	47.92	17.81	0.00		82.19	14.57
1995	39.08	24.27	33.31	20.97	0.00		72.40	17.81
1996	58.22	36.45	56.61	21.95	0.00		114.83	24.87
1997	43.81	16.11	36.25	18.22	0.00		80.06	15.50
1998	15.57	21.50	42.05	25.84	0.00		57.62	22.48
1999	30.89	12.99	43.60	17.29	0.00		74.49	14.03
2000	28.49	41.08	28.64	25.90	0.40	100.00	57.53	29.39
2001	19.30	37.33	28.37	27.57	0.00		47.67	29.98
2002	34.57	18.68	42.29	17.29	0.00		76.87	16.22
2003	30.34	13.95	43.76	15.41	0.00		74.10	12.28

2004	14.78	17.20	22.40	19.86	0.00		37.18	14.34
2005	17.68	31.64	48.17	50.36	0.00		65.85	43.08
2006	59.14	23.53	68.10	29.47	0.00	100.00	127.24	24.23
2007	35.51	27.39	37.78	16.73	0.00		73.29	18.52
2008	53.26	15.49	50.52	13.76	0.00		103.78	10.21
2009	31.47	16.45	49.47	19.10	0.00		80.94	15.95
2010	33.27	18.67	46.17	22.36	0.00		79.44	18.30
2011	41.04	23.07	65.17	14.12	0.00		106.21	15.04
2012	66.15	18.30	96.32	44.81	0.01	80.31	162.48	29.30
2013	40.45	31.52	71.58	14.84	0.01	82.00	112.04	18.33
2014								
2015	25.18	40.38	45.00	24.30	1.15	99.81	71.33	24.39
2016	48.15	15.30	73.64	19.71	0.01	100.00	121.80	17.12
2017	18.73	20.07	46.67	13.57	0.00		65.40	14.90
2018	35.41	25.01	51.46	15.11	0.01	54.06	86.88	18.61
2019	53.86	38.82	82.52	12.37	0.00		136.38	19.05
2020								
2021	58.40	20.38	117.92	13.69	0.00		176.31	15.72

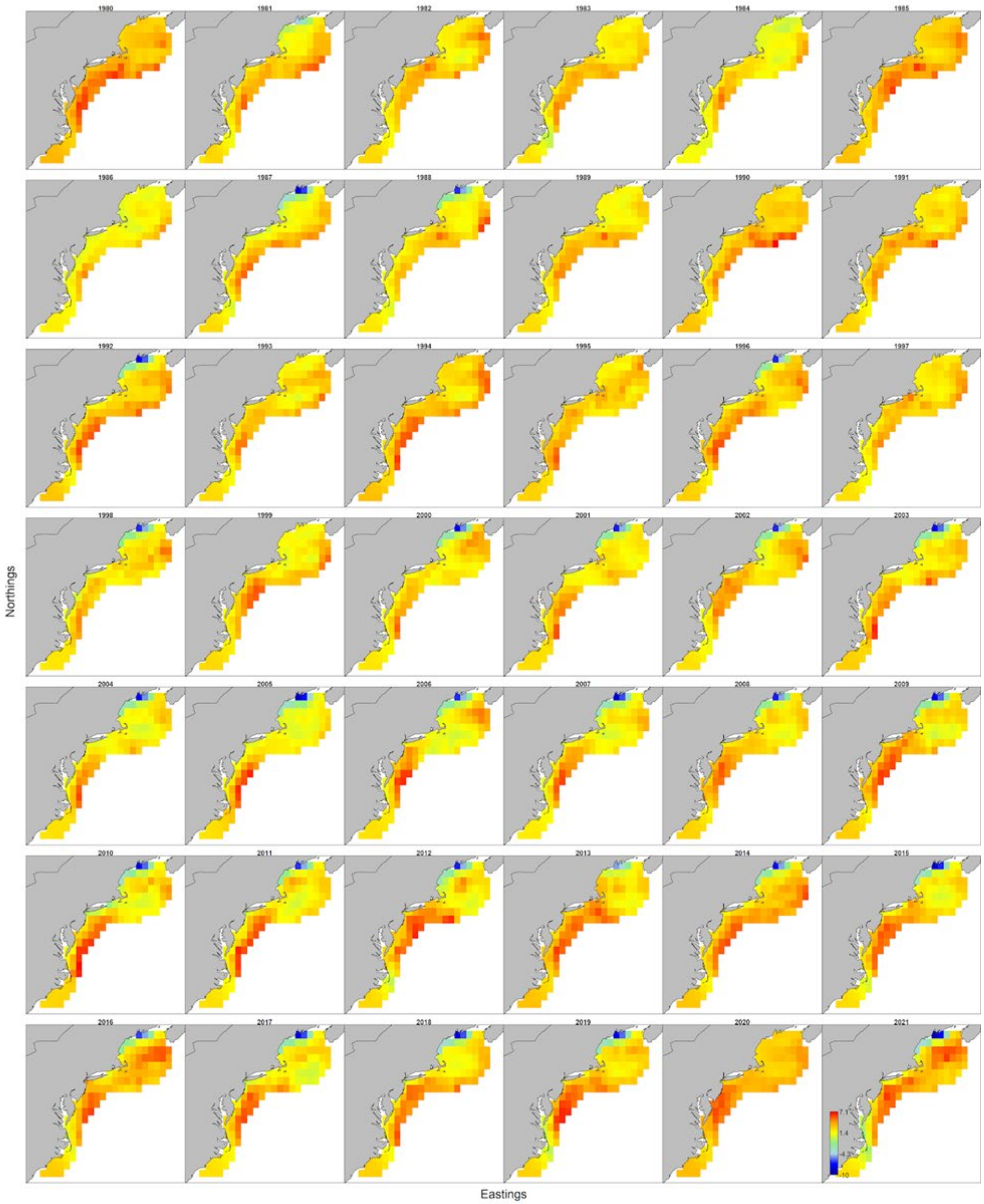


Figure 3.1. VAST estimated encounter probability for spiny dogfish in the spring by year (1980-2021).

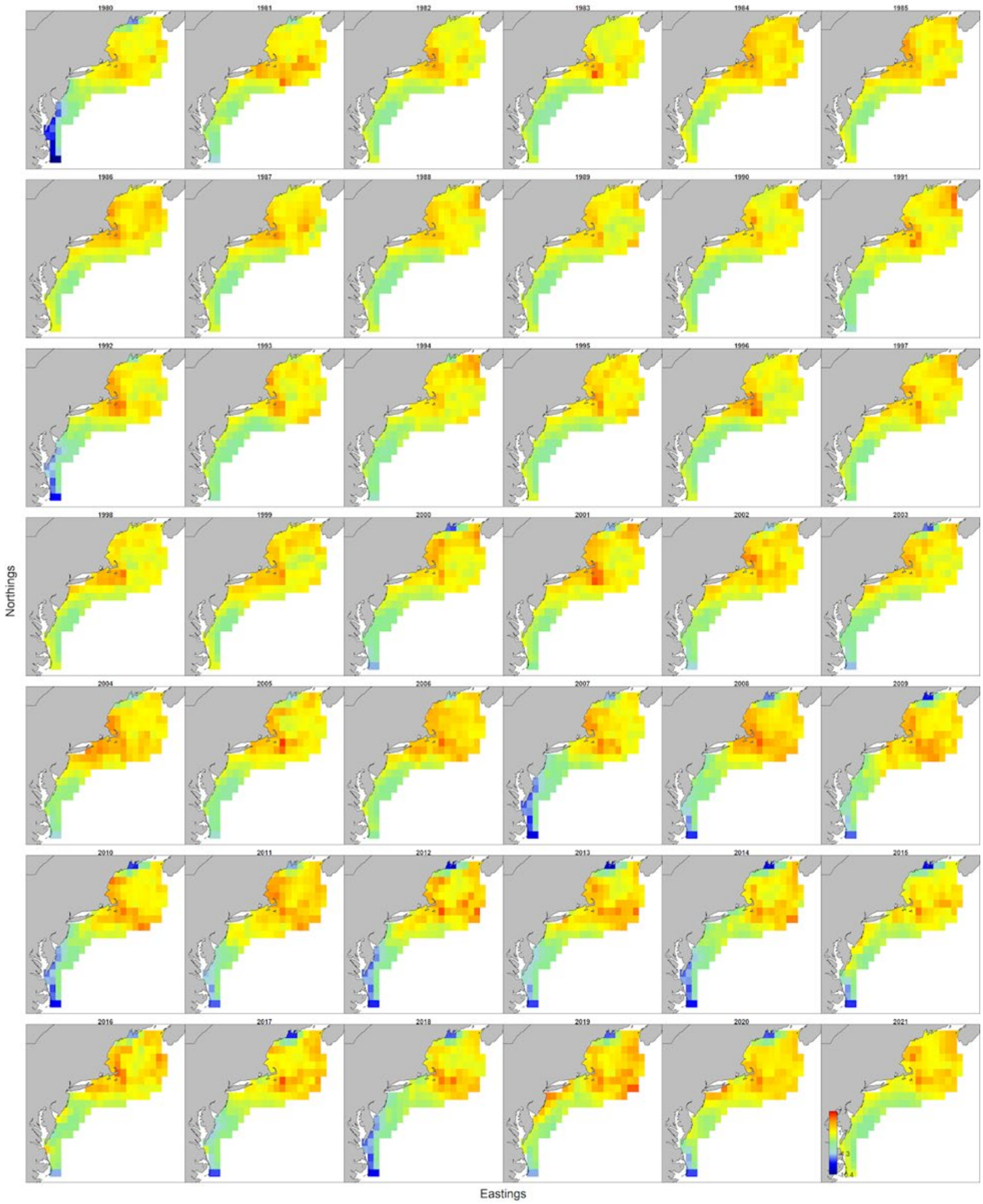


Figure 3.2. VAST estimated encounter probability for spiny dogfish in the fall by year (1980-2021).

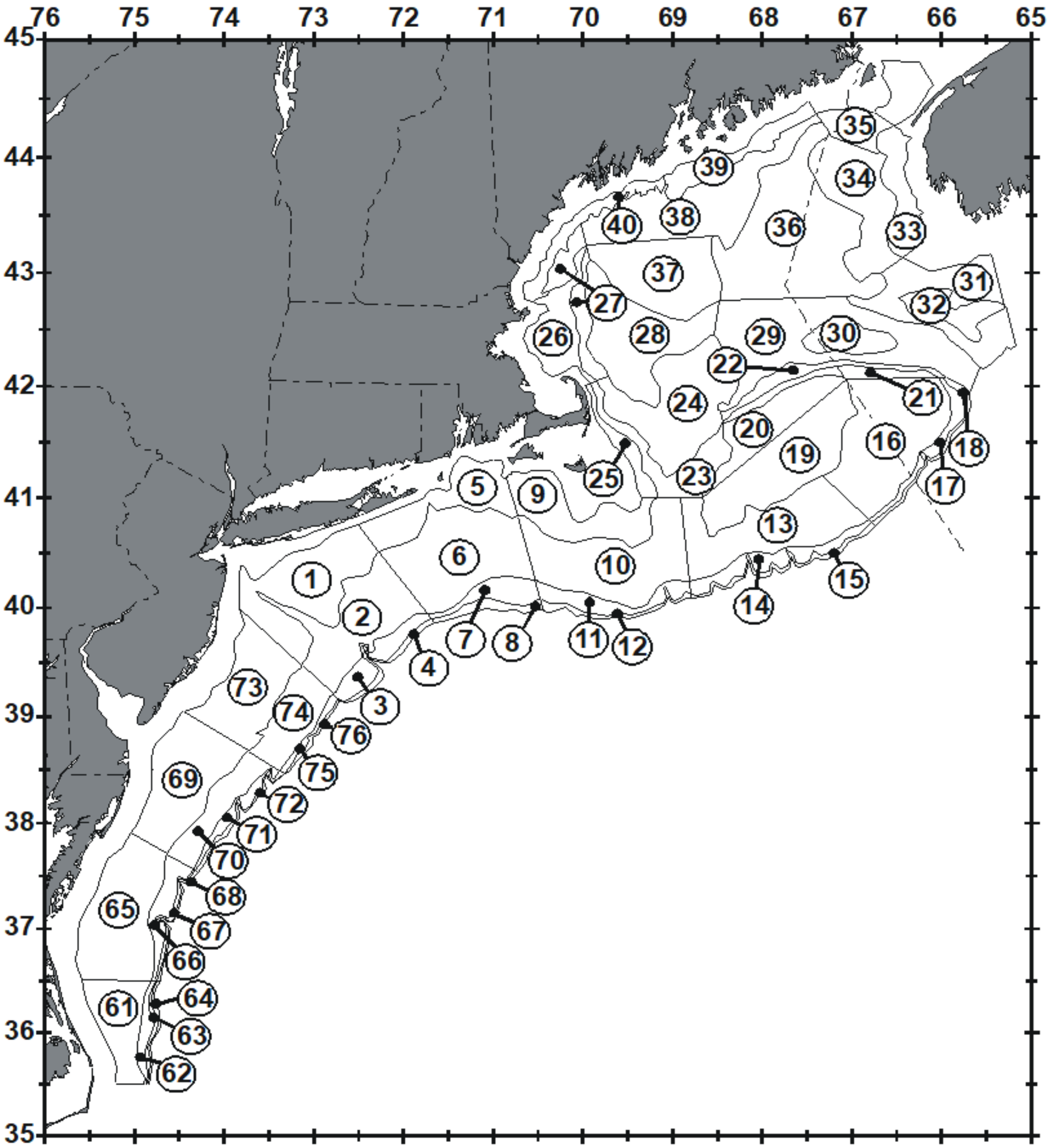


Figure 3.3. NEFSC Fall, Spring, and Winter Bottom Trawl Survey offshore stations.

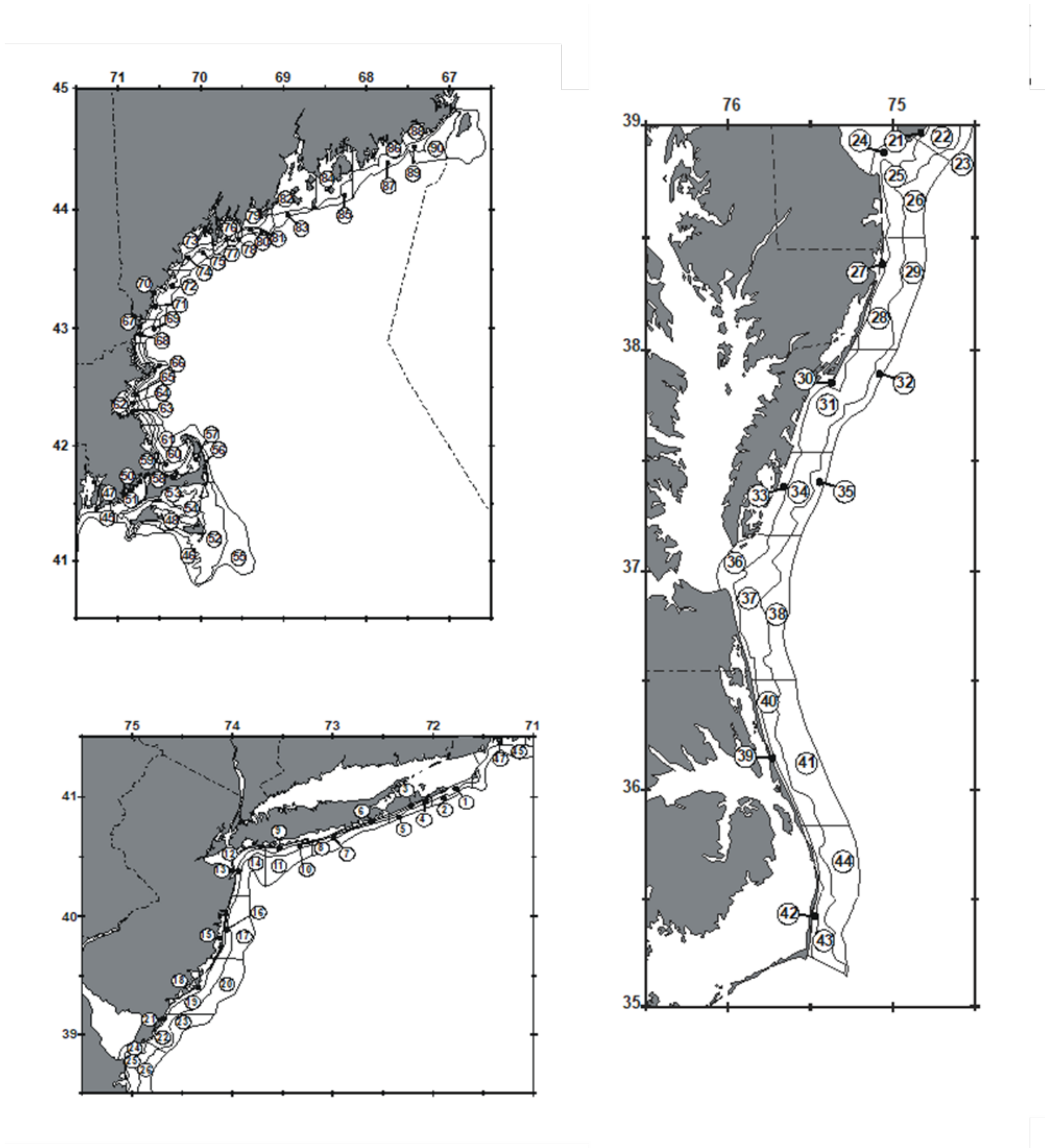


Figure 3.4. NEFSC Fall and Spring Bottom Trawl Survey inshore stations.

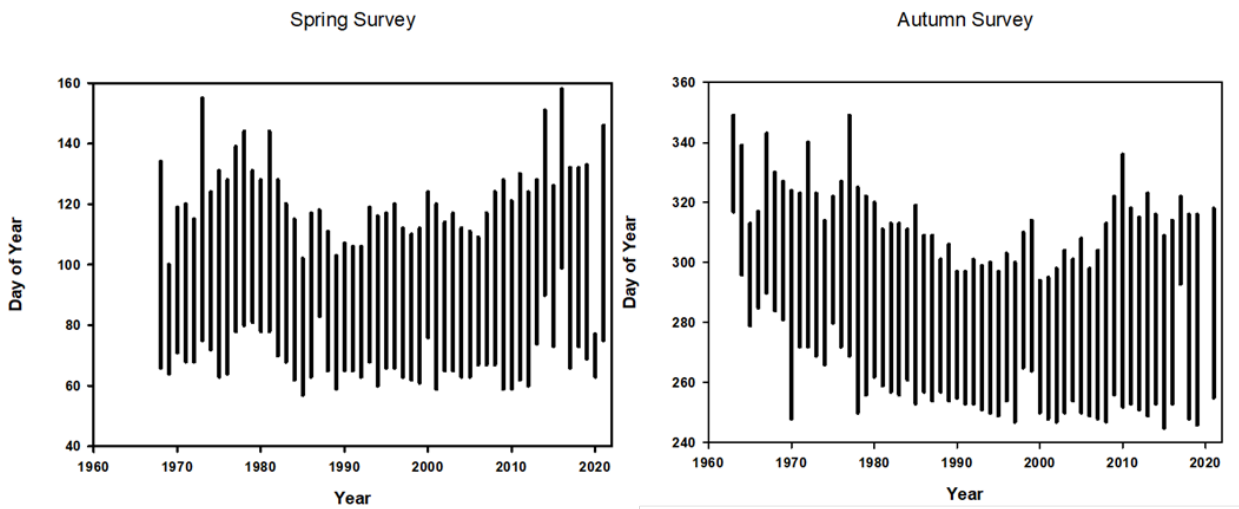


Figure 3.5. NEFSC Spring and Fall Bottom Trawl Survey annual timing.

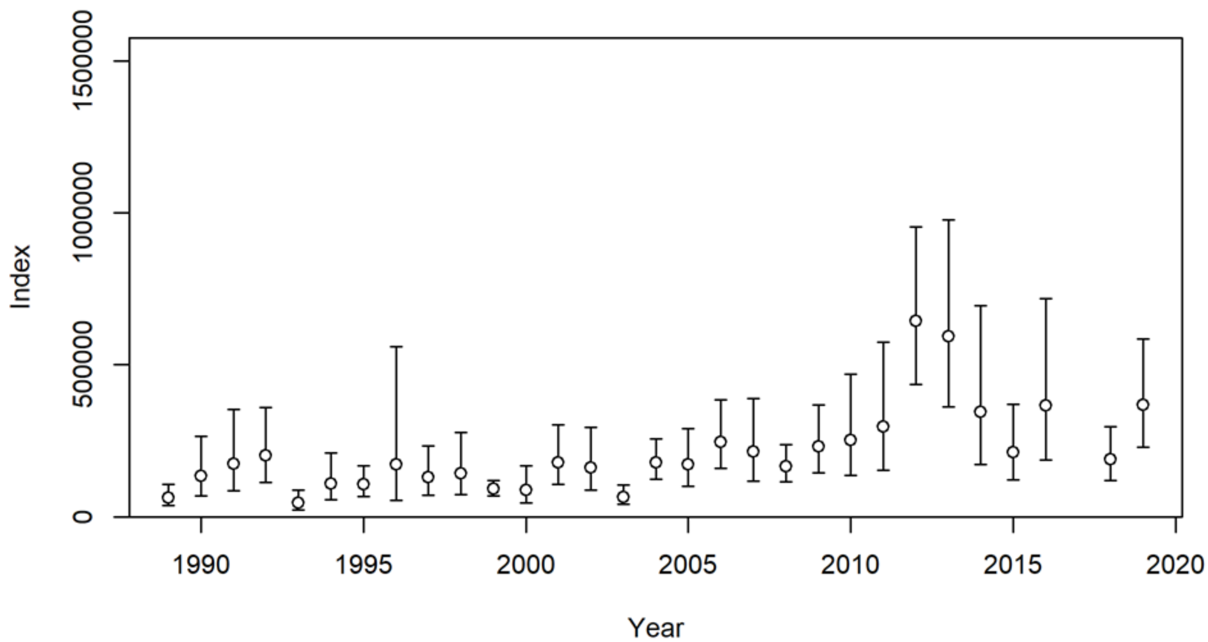


Figure 3.6. Annual NEFSC Fall Bottom Trawl Survey design based estimates of total biomass.

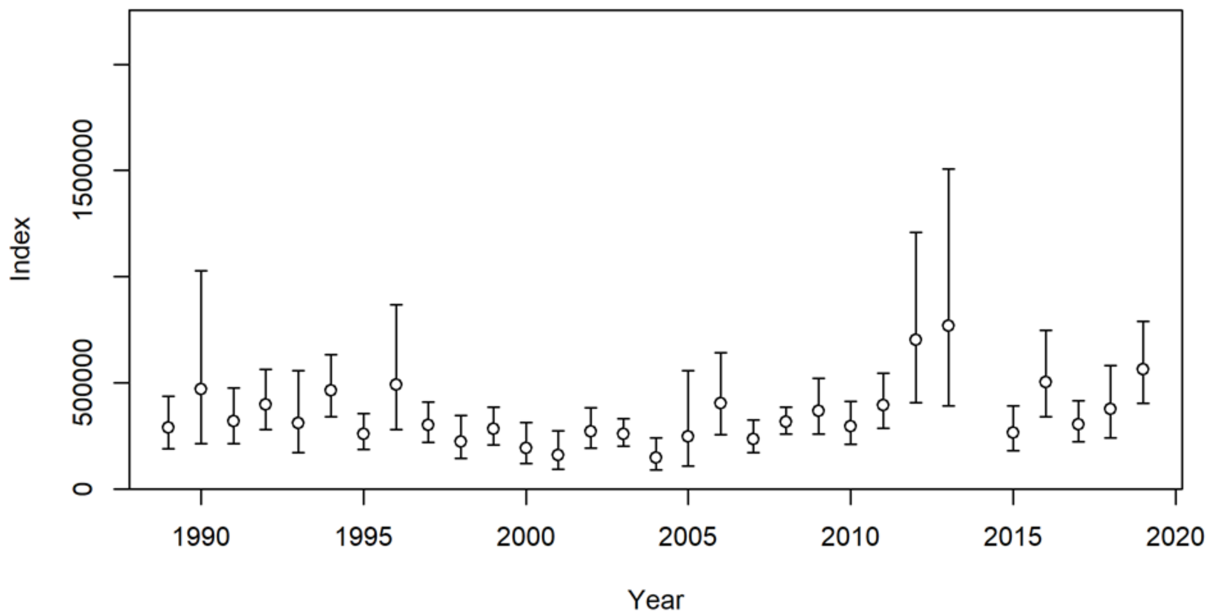


Figure 3.7. Annual NEFSC Spring Bottom Trawl Survey design based estimates of total biomass.

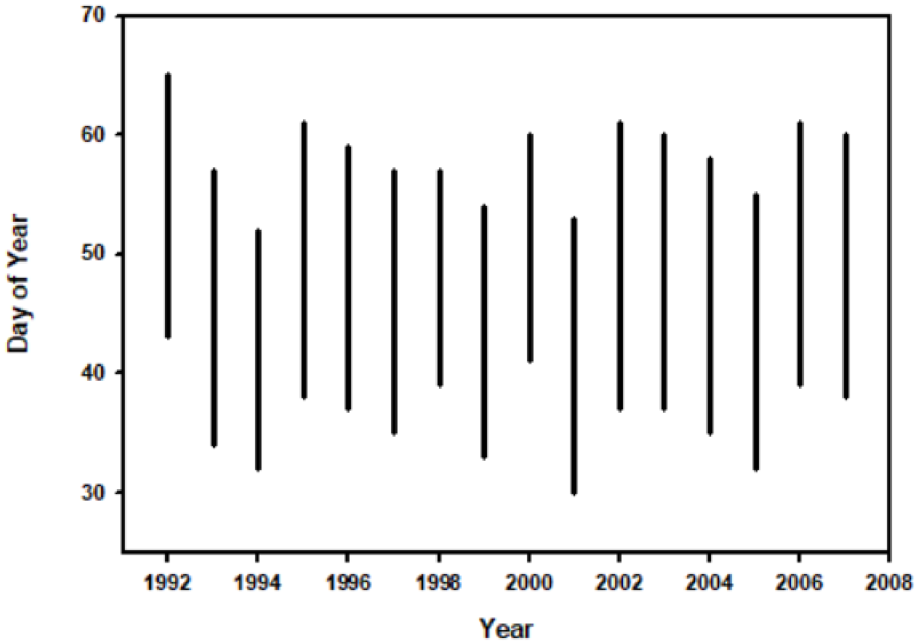
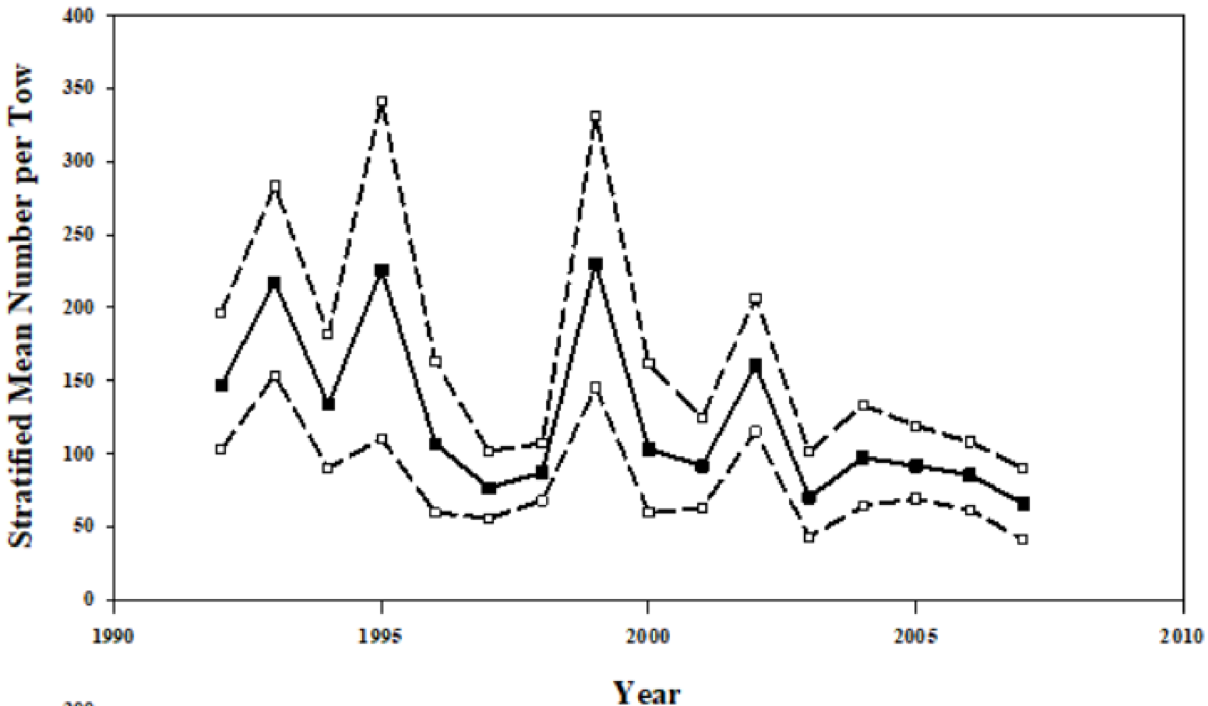


Figure 3.8. NEFSC Winter Bottom Trawl Survey annual timing.



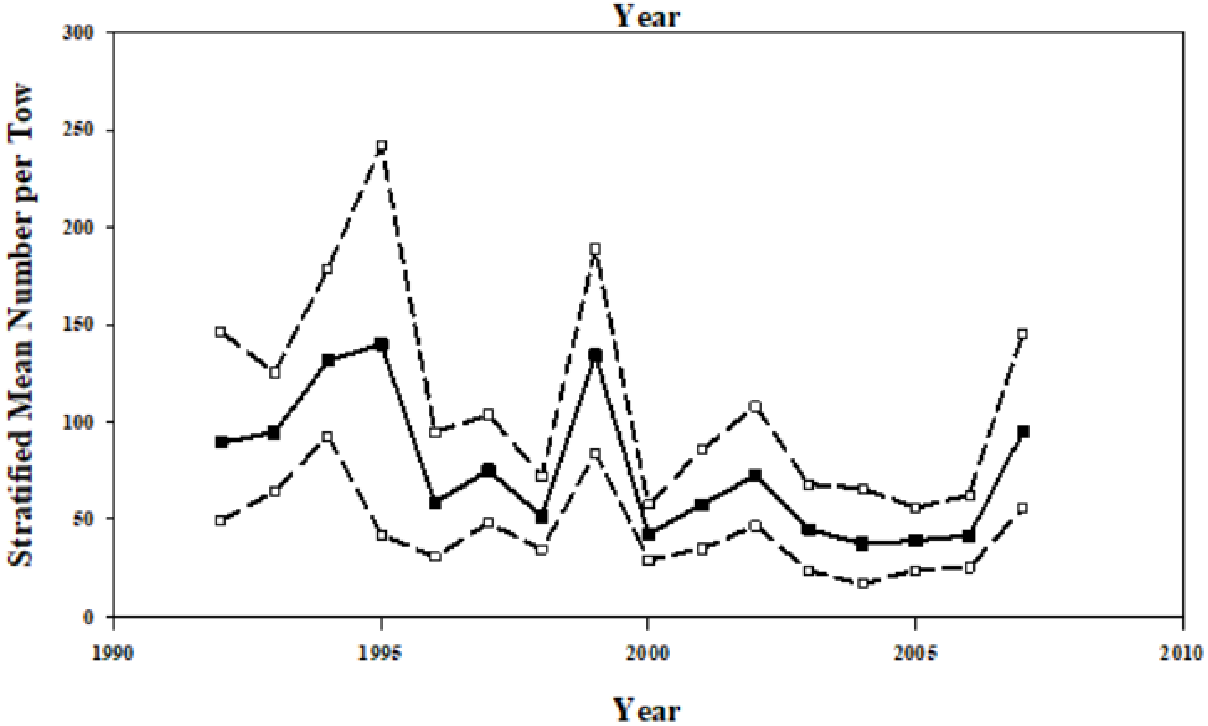


Figure 3.9. Annual NEFSC Winter Bottom Trawl Survey stratified mean number per tow (solid circles and line) and 95% upper and lower confidence limits (open circles and dashed lines) for the strata off southern New England (strata 01010-01120, top) and mid-Atlantic (strata 0610-01760, bottom) regions.

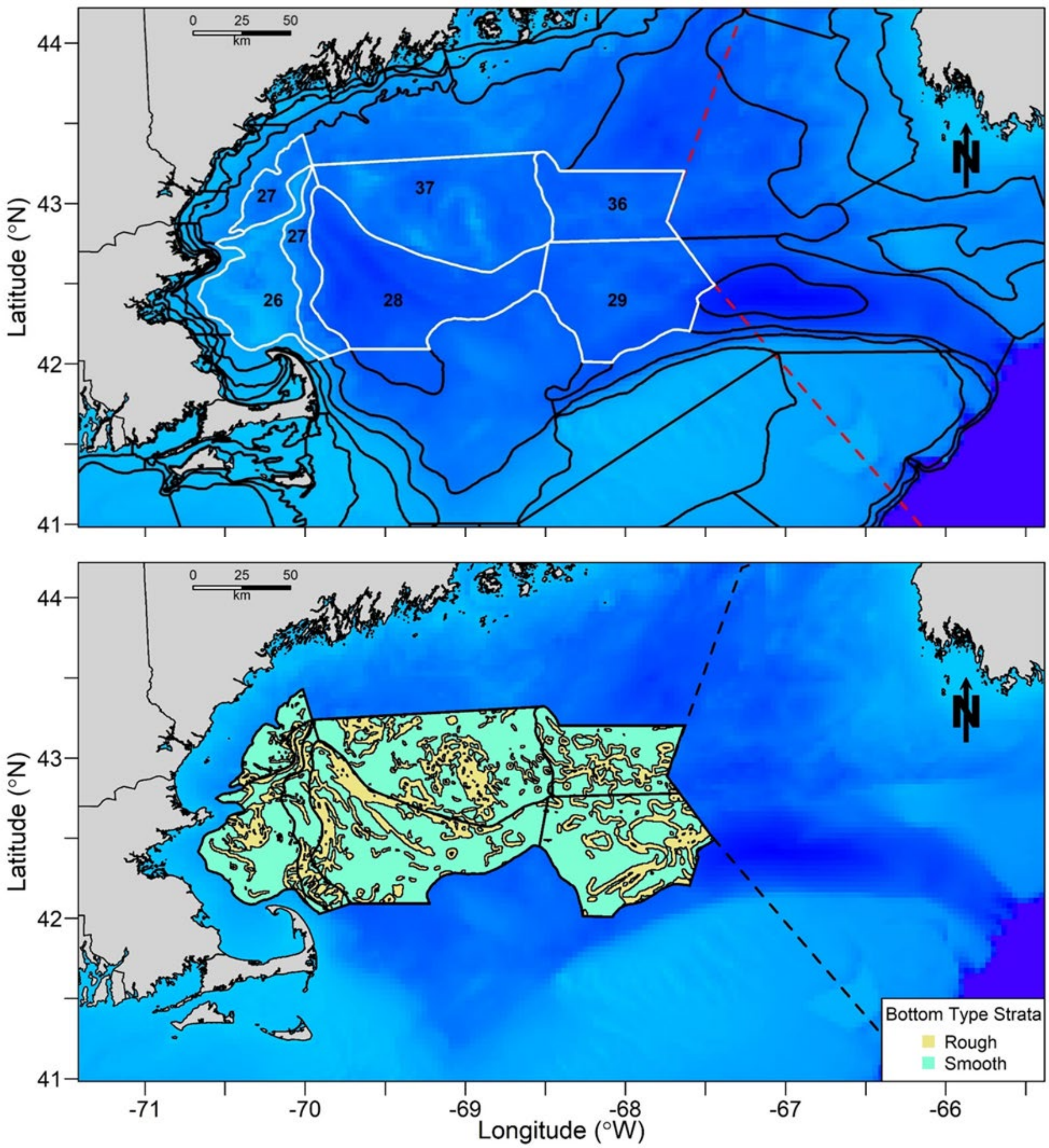


Figure 3.10. Northeast Fisheries Science Center Gulf of Maine Bottom Longline Survey strata (black lines, top panel) and their sub-stratification (bottom panel) by rough (yellow) and smooth (green) bottom types. The dashed line is the Exclusive Economic Zone boundary.

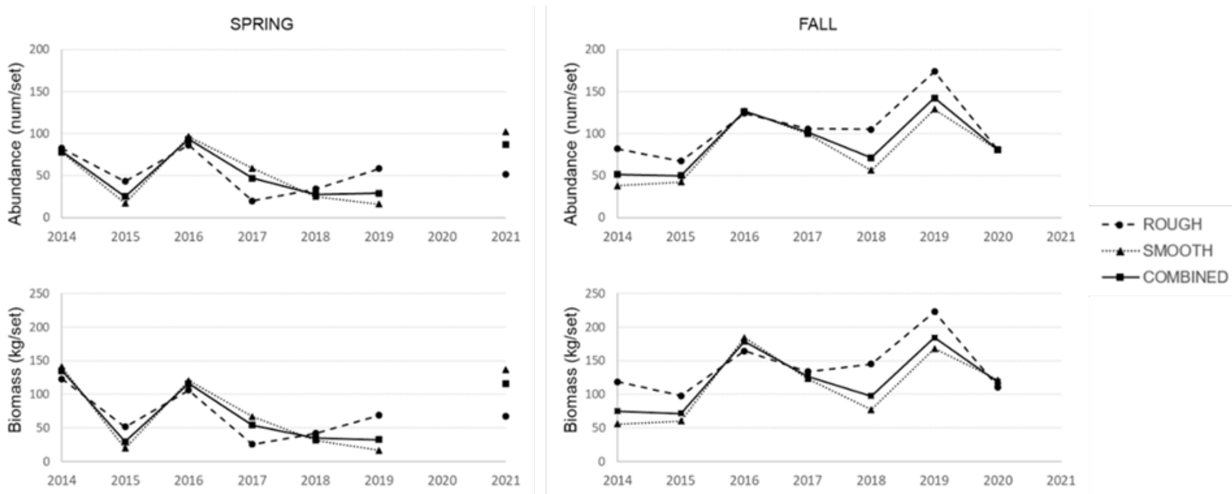


Figure 3.11. NEFSC Gulf of Maine Bottom Longline Survey Stratified mean numbers/set (top row) and kg/set (bottom row) index estimates for male spiny dogfish for the spring (a) and fall (b) by year and bottom type (Rough (circle with dashed line), Smooth (triangle with dotted line), and Combined (square with solid line)).

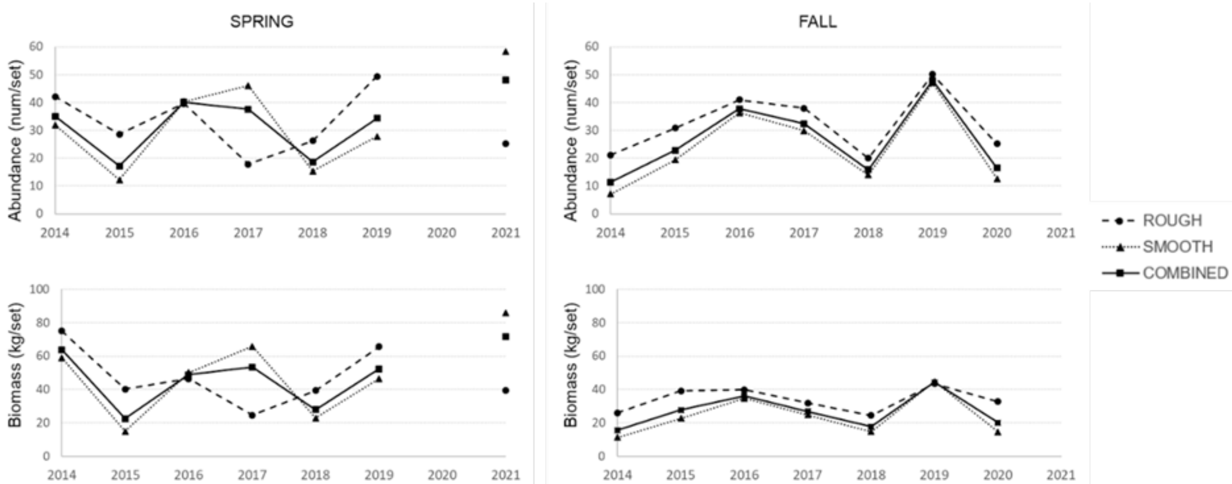


Figure 3.12. NEFSC Gulf of Maine Bottom Longline Survey Stratified mean numbers/set (top row) and kg/set (bottom row) index estimates for female spiny dogfish for the spring (a) and fall (b) by year and bottom type (Rough (circle with dashed line), Smooth (triangle with dotted line), and Combined (square with solid line)).

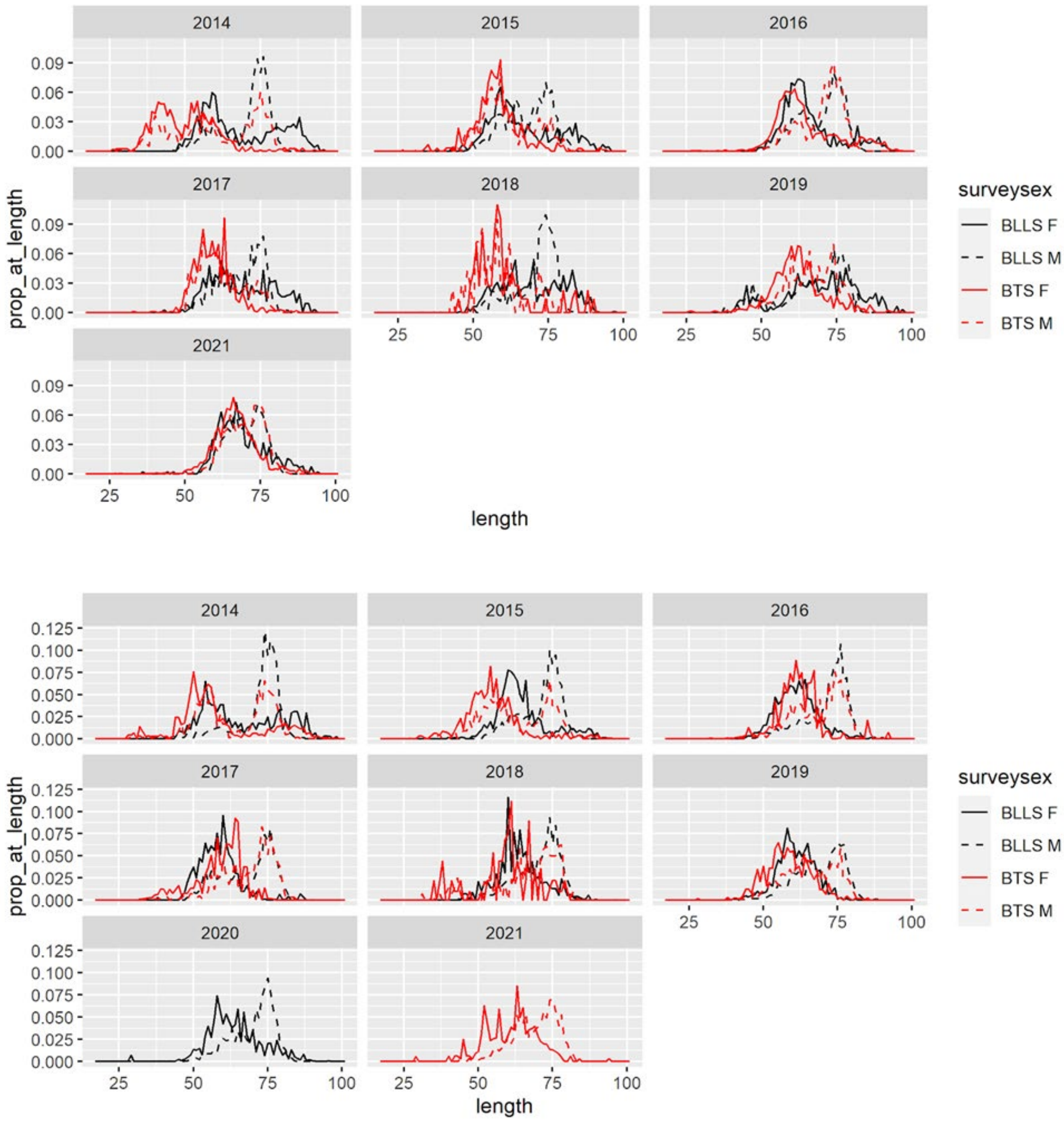


Figure 3.13. Spiny dogfish proportions at length (cm) by season (spring = top, fall = bottom), sex (females = solid lines, males = dashed lines), and year in the Northeast Fisheries Science Center Gulf of Maine Bottom Longline Survey (BLLS; black lines) and the Northeast Fisheries Science Center Bottom Trawl Survey (BTS; red lines) during 2014 – 2021. BTS lengths were only for the 6 strata covered by the BLLS.

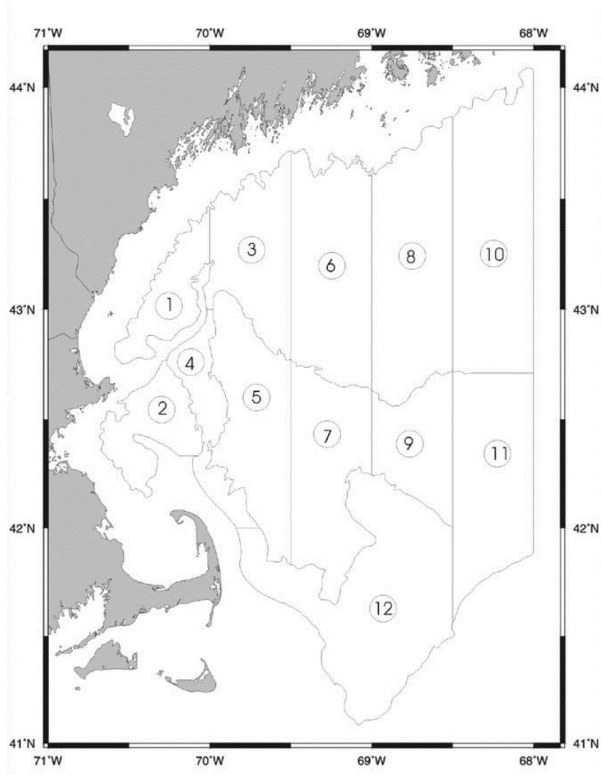


Figure 3.14. Strata used for the ASMFC Northern Shrimp Bottom Trawl Survey.

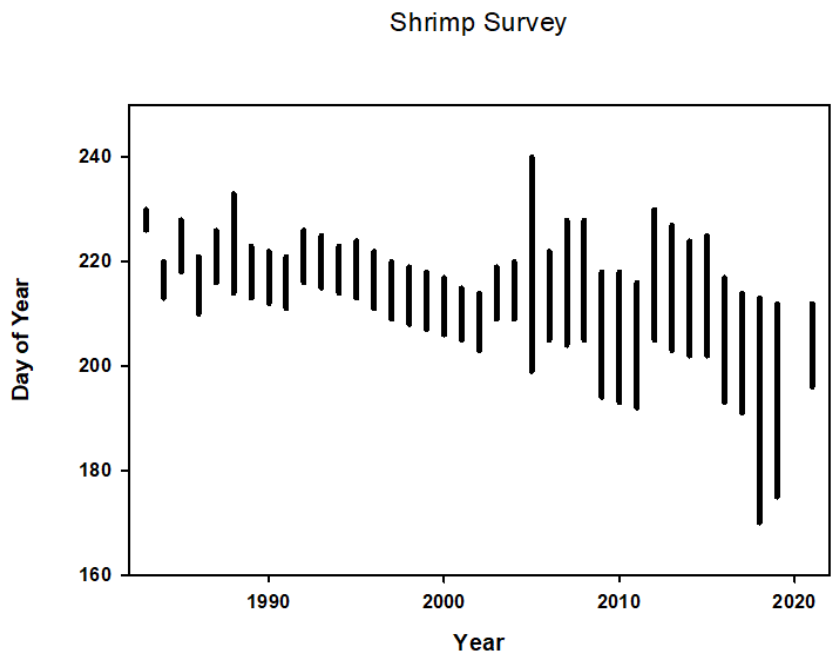


Figure 3.15. Timing of the ASMFC Northern Shrimp Bottom Trawl Survey from 1983-2021.

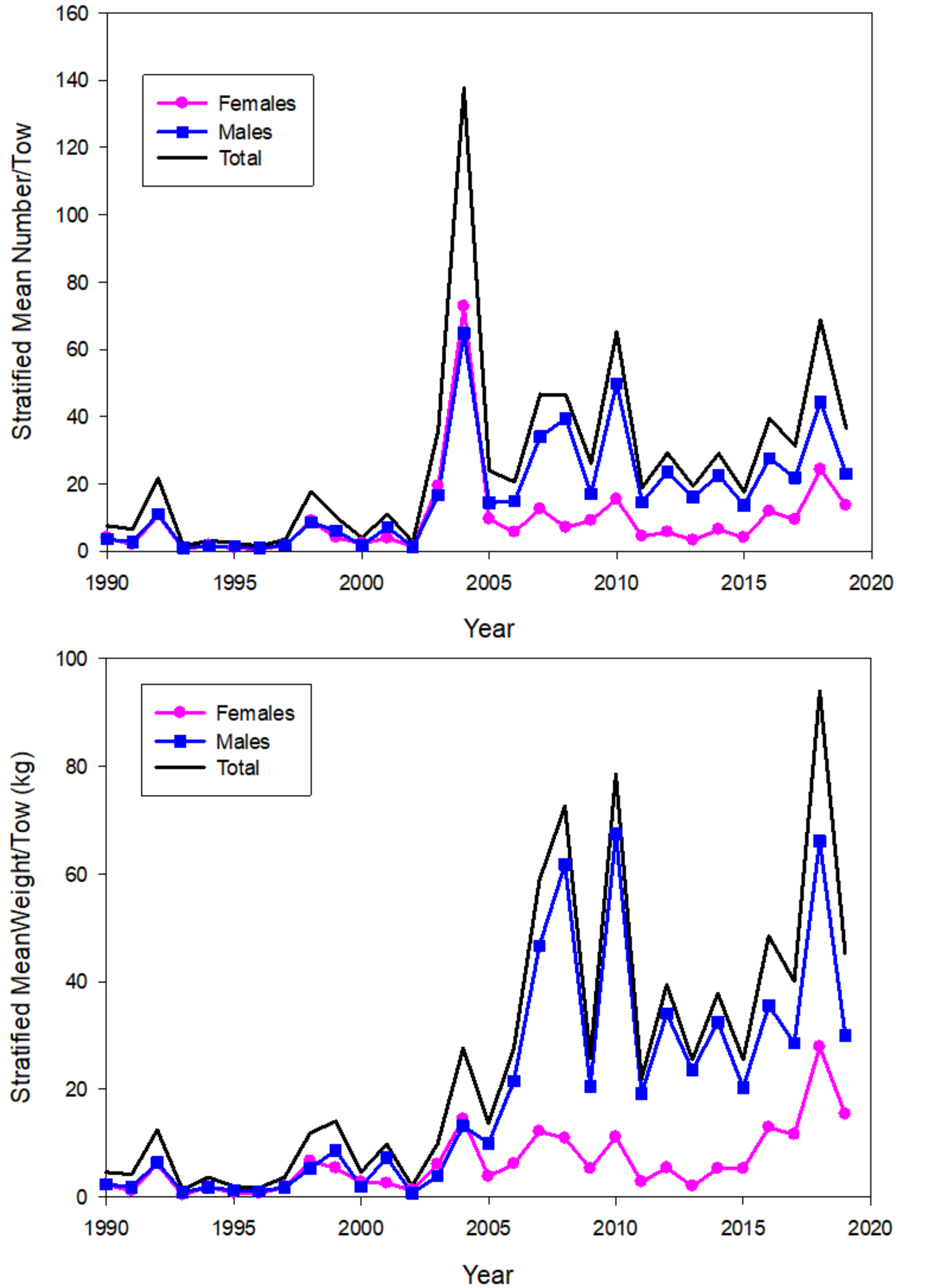


Figure 3.16. Stratified mean numbers (top panel) and weight (bottom panel) per tow from the ASMFC Northern Shrimp Bottom Trawl Survey from 1990-2019.

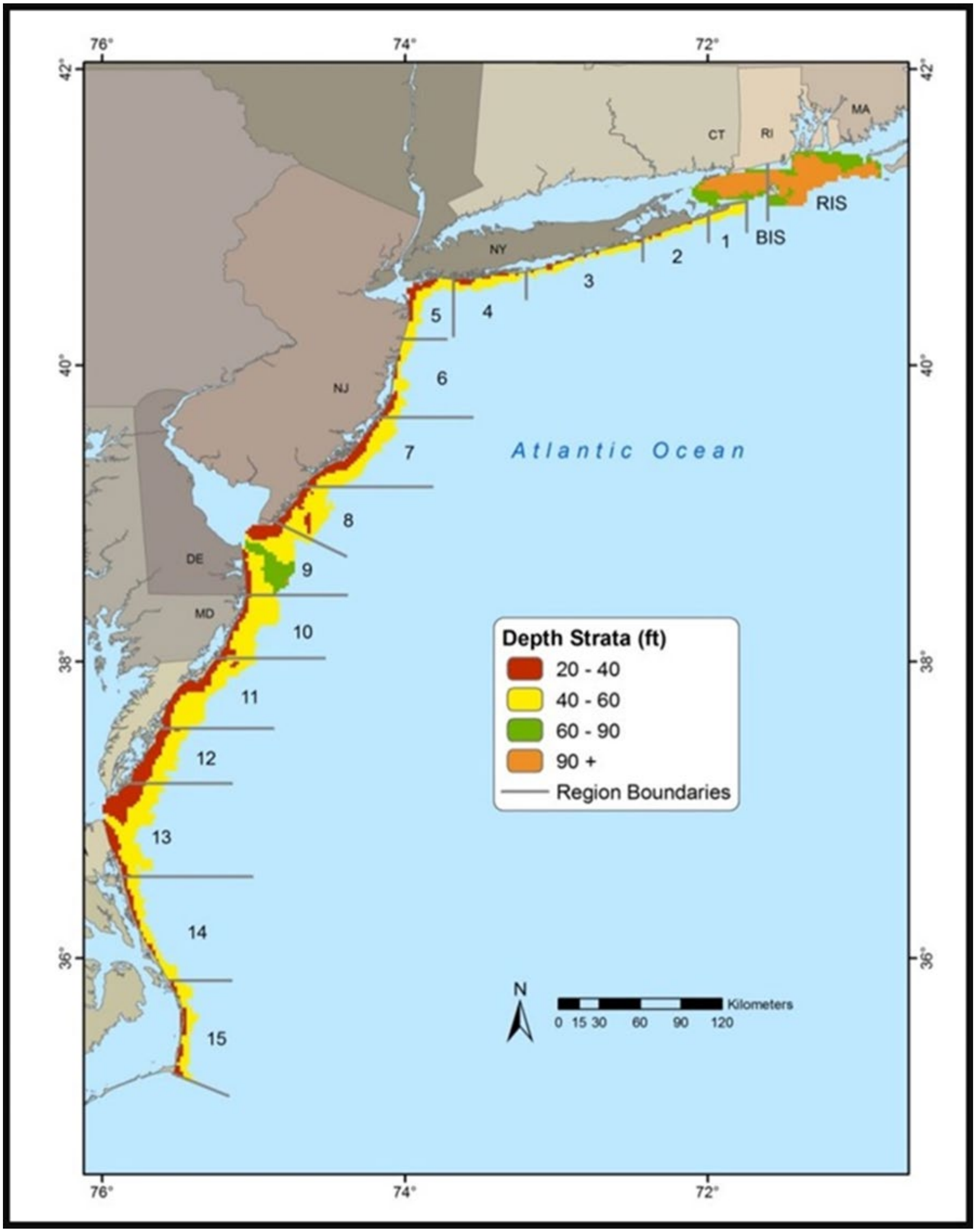


Figure 3.17. Sampling strata used in the NEAMAP survey. Map provided by NEAMAP and available here: http://www.neamap.net/index.html.

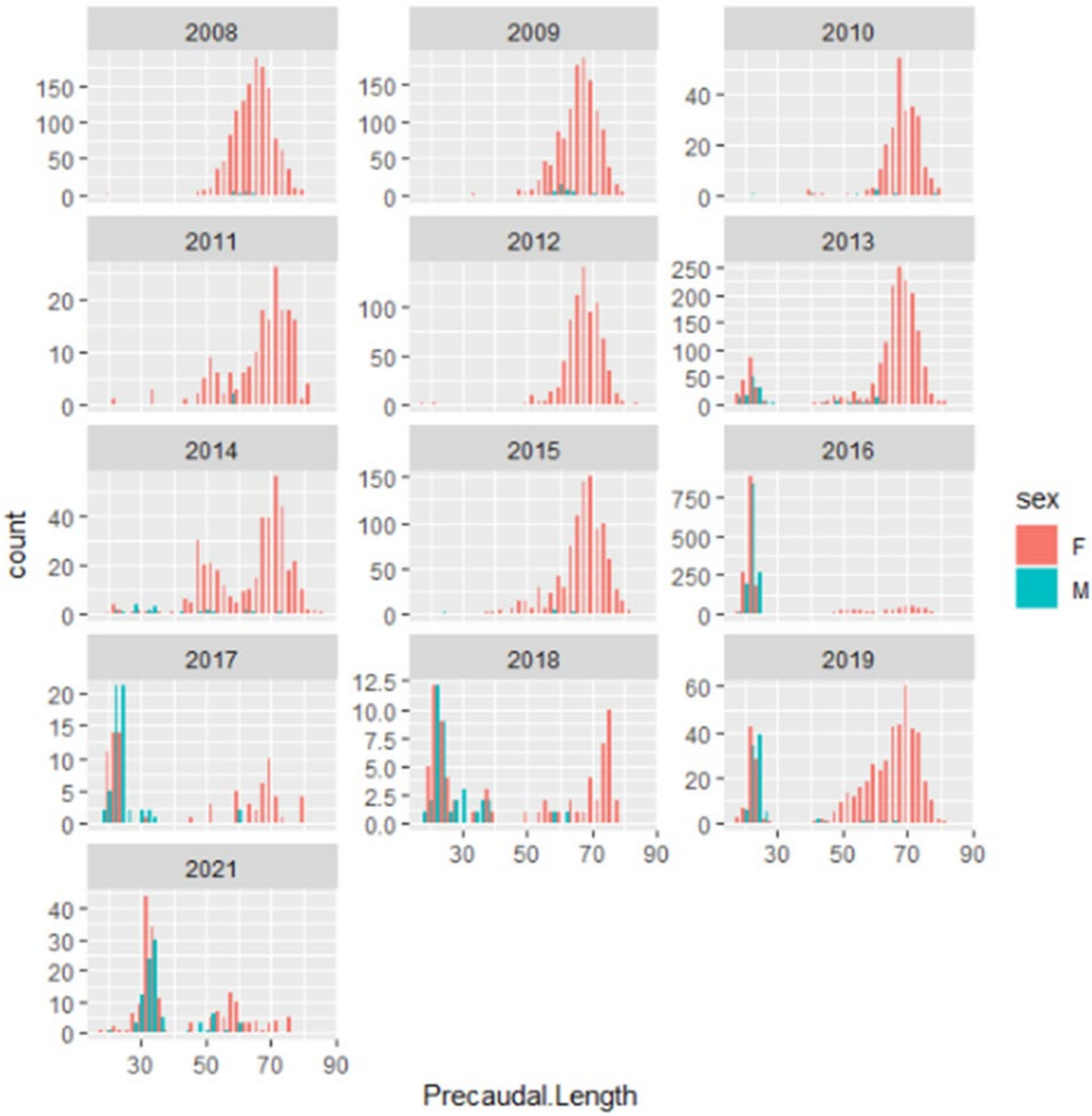


Figure 3.18. NEAMAP precaudal length (cm) frequencies of male and female spiny dogfish by year for the spring tows.

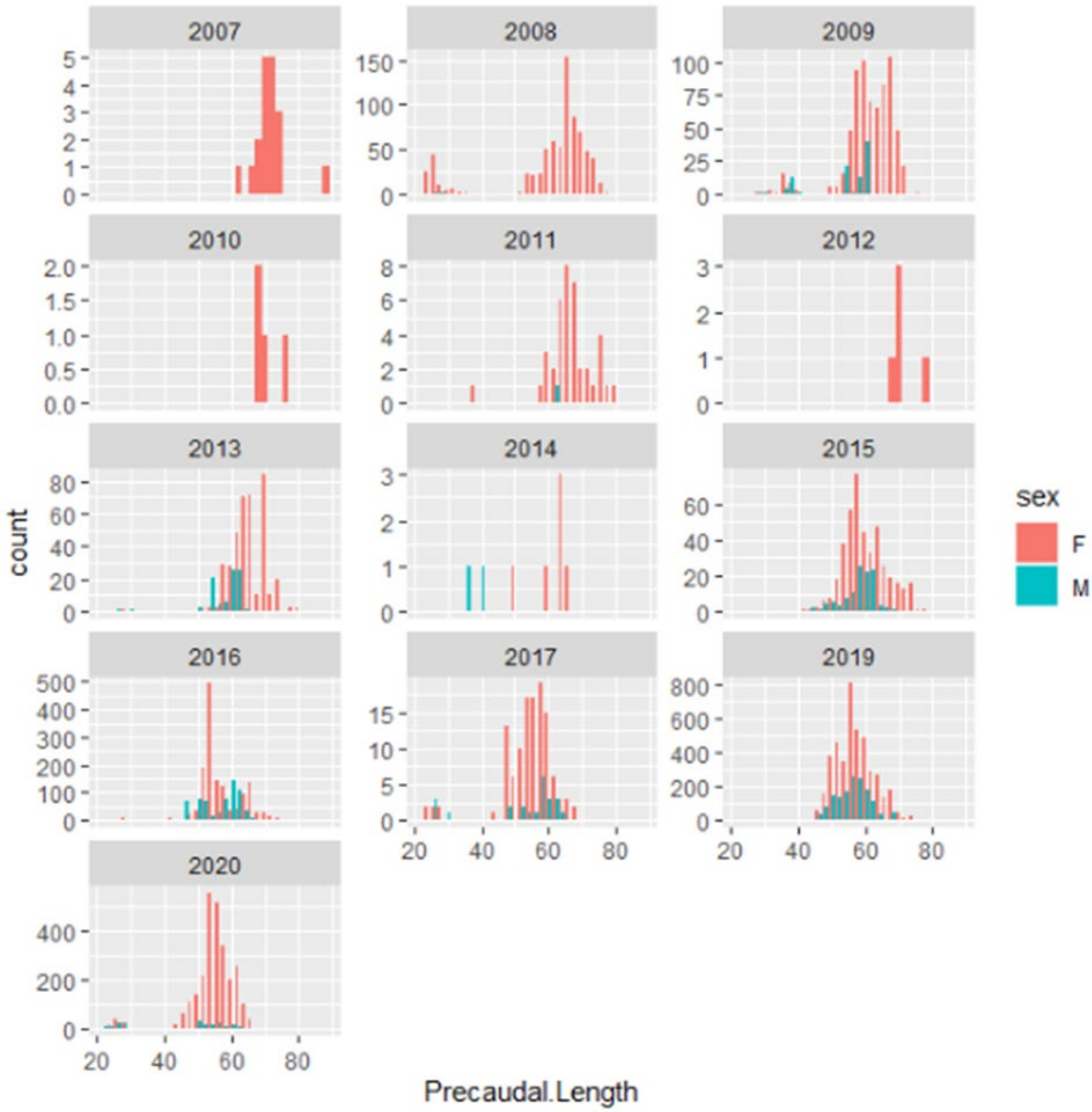


Figure 3.19. NEAMAP precaudal length (cm) frequencies of male and female spiny dogfish by year for the fall tows.

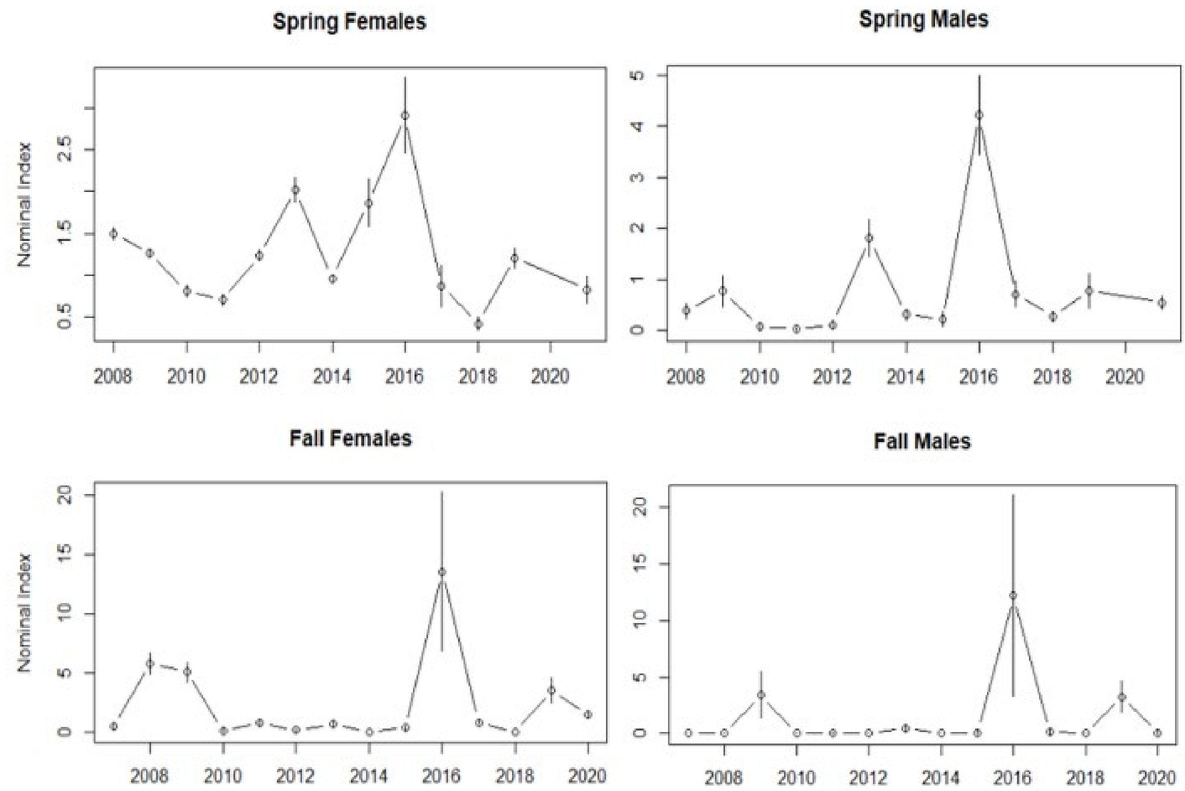


Figure 3.20. NEAMAP nominal (arithmetic mean) indices by season and sex with 95% confidence intervals.

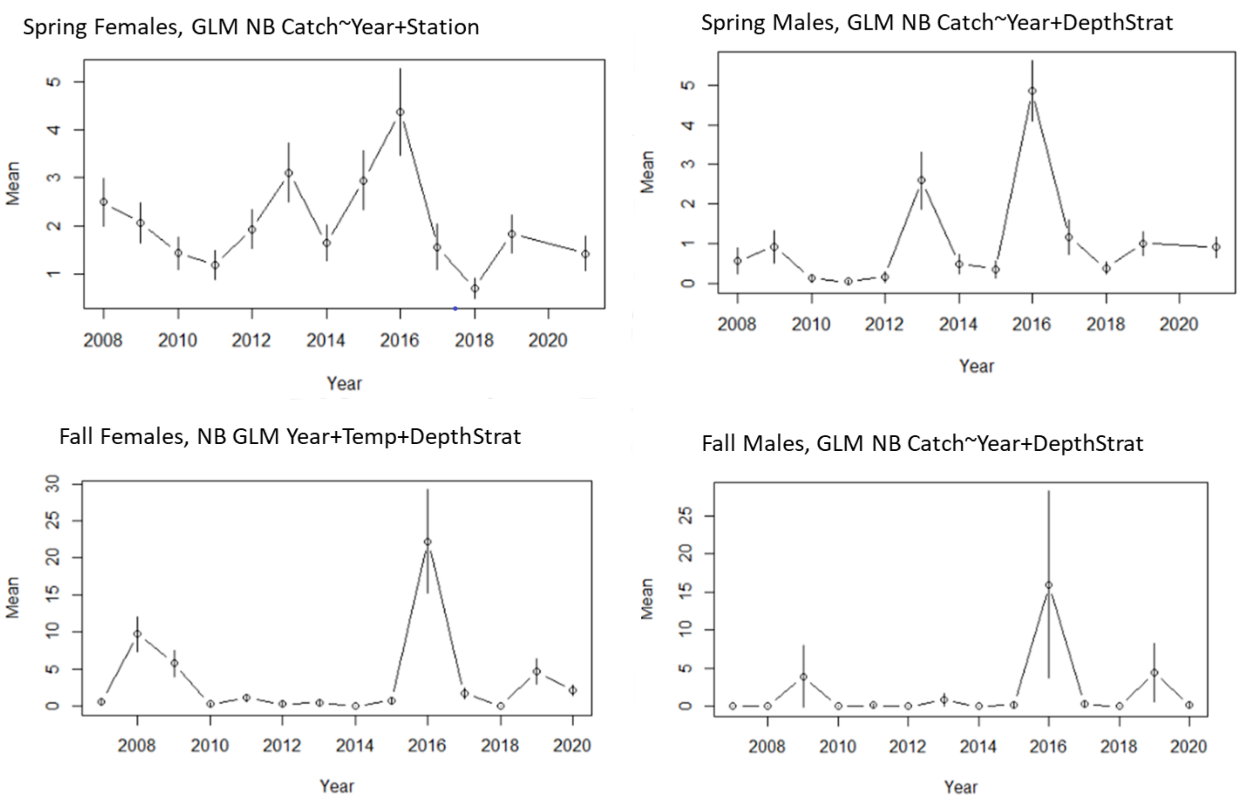


Figure 3.21. NEAMAP standardized indices of abundance by season and sex with 95% confidence intervals.

MADMF TRAWL SURVEY

STATION ASSIGNMENT BY STRATUM

REGION	STRATUM	DEPTH (METERS)	DEPTH (FEET)	SUBSTRATUM A	SUBSTRATUM B	# STATIONS ASSIGNED
5	31	0 - 9.1	0 - 30	2	1	3
	32	9.2 - 18.3	31- 60	1	2	3
	33	18.4 - 27.4	61 - 90	2	2	4
	34	27.5 - 36.6	91 - 120	2	2	4
	35	36.7 - 54.9	121 - 180	2	3	5
	36	>=55.0	>180	1	1	2
4	25	0 - 9.1	0 - 30	2	2	4
	26	9.2 - 18.3	31- 60	3	2	5
	27	18.4 - 27.4	61 - 90	3	2	5
	28	27.5 - 36.6	91 - 120	3	2	5
	29	36.7 - 54.9	121 - 180	3	2	5
	30	>=55.0	>180	0	2	2
3	17	0 - 9.1	0 - 30	2	3	5
	18	9.2 - 18.3	31- 60	2	3	5
	19	18.4 - 27.4	61 - 90	1	1	2
	20	27.5 - 36.6	91 - 120	2	0	2
	21	36.7 - 54.9	121 - 180	2	0	2
2	15	0 - 9.1	0 - 30	10		10
	16	9.2 - 18.3	31- 60	11		11
1	11	0 - 9.1	0 - 30	3	2	5
	12	9.2 - 18.3	31- 60	5	2	7
	13	18.4 - 27.4	61 - 90	3	2	5
	14	27.5 - 36.6	91 - 120	1	1	2
TOTAL STATION ASSIGNMENT						103

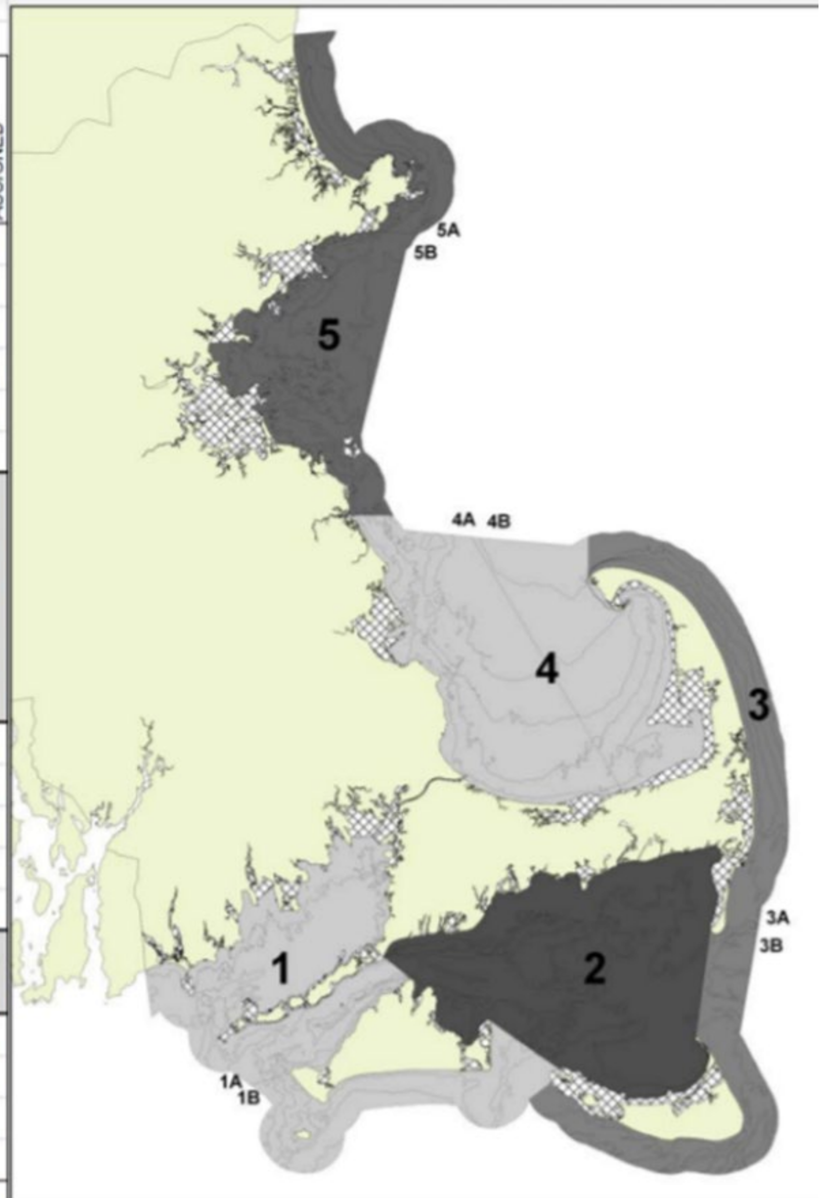


Figure 3.22 MADMF biannual bottom trawl survey strata.

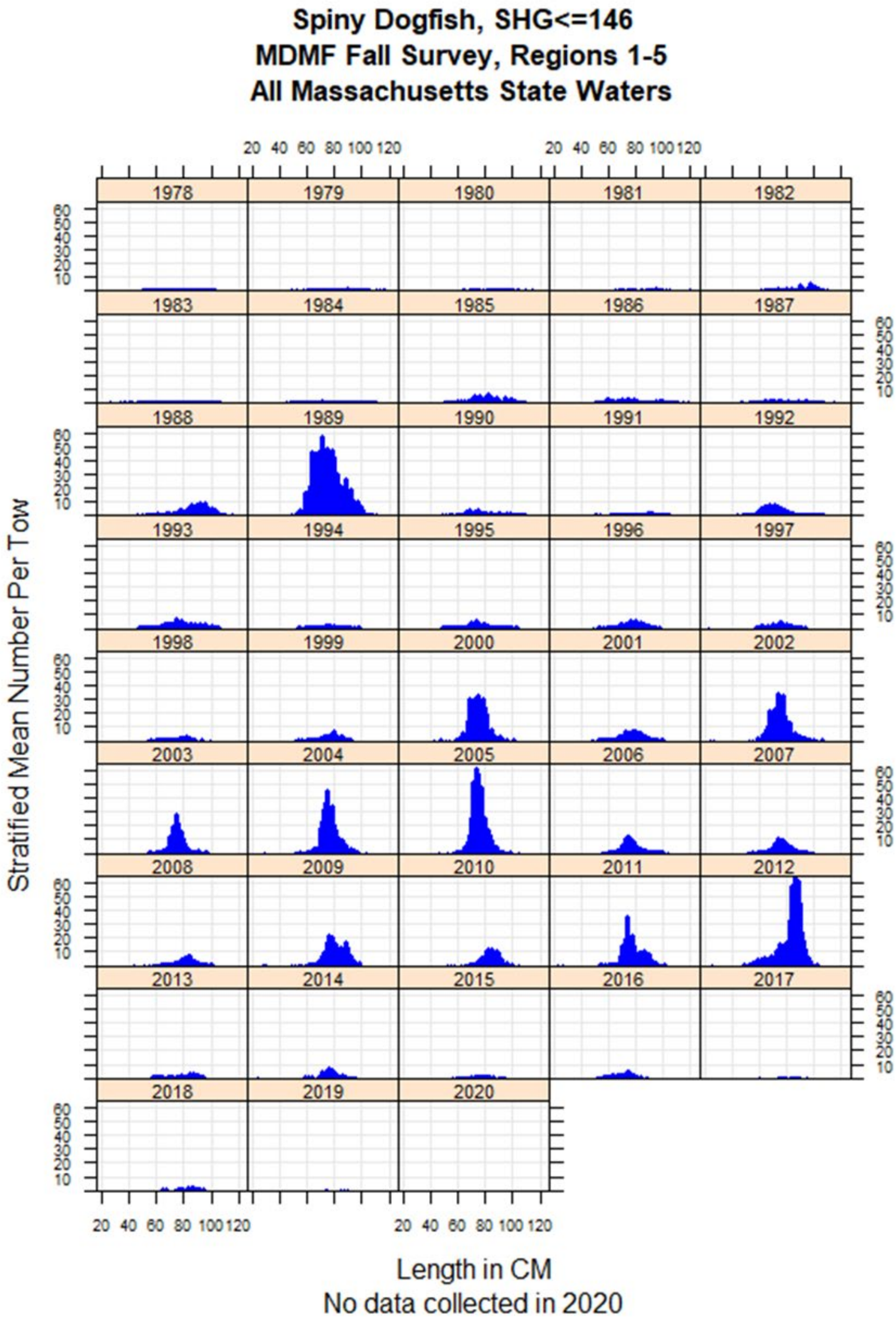


Figure 3.23. MADMF fall bottom trawl survey length composition by year.

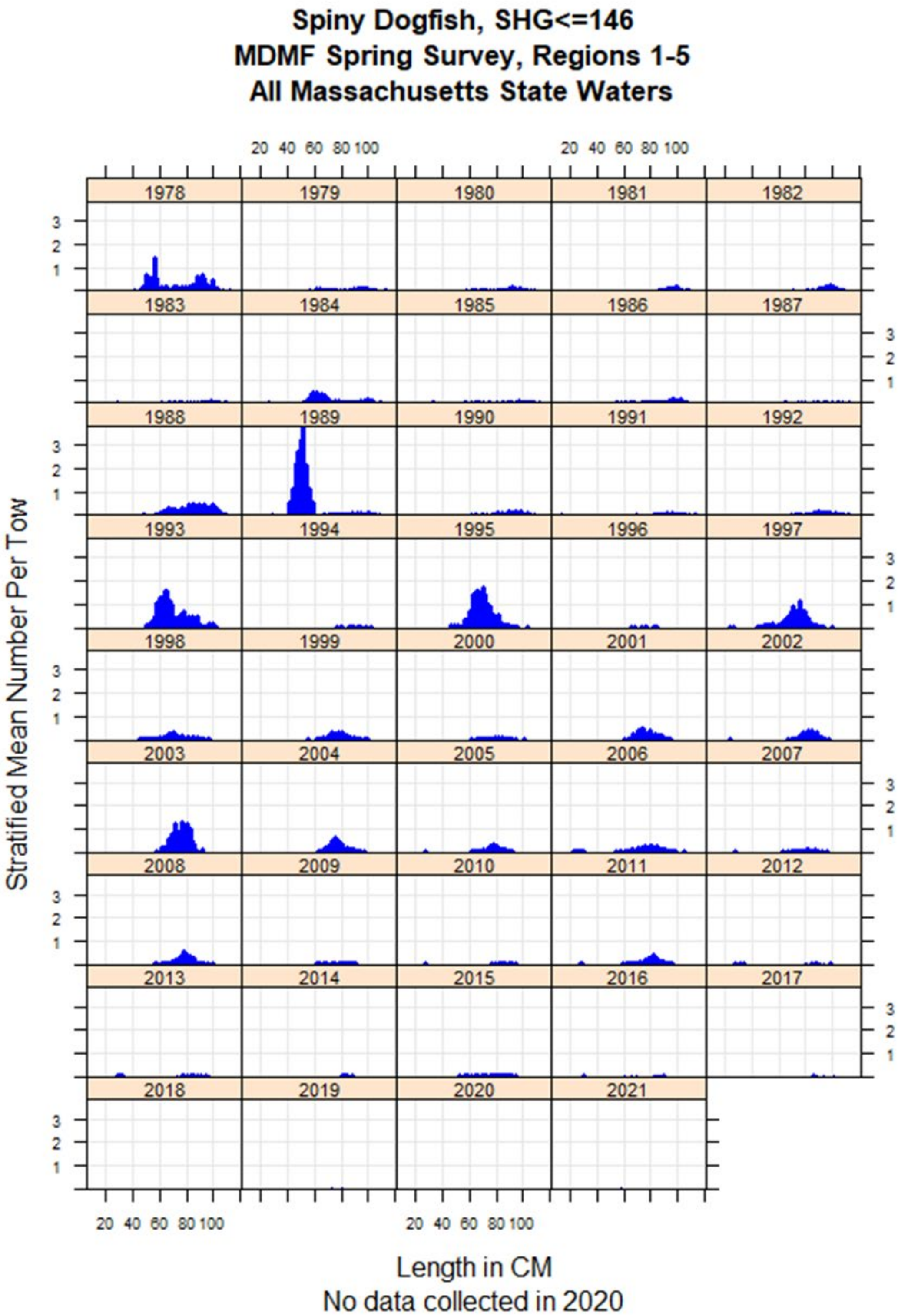


Figure 3.24. MDMF spring bottom trawl survey length composition by year.

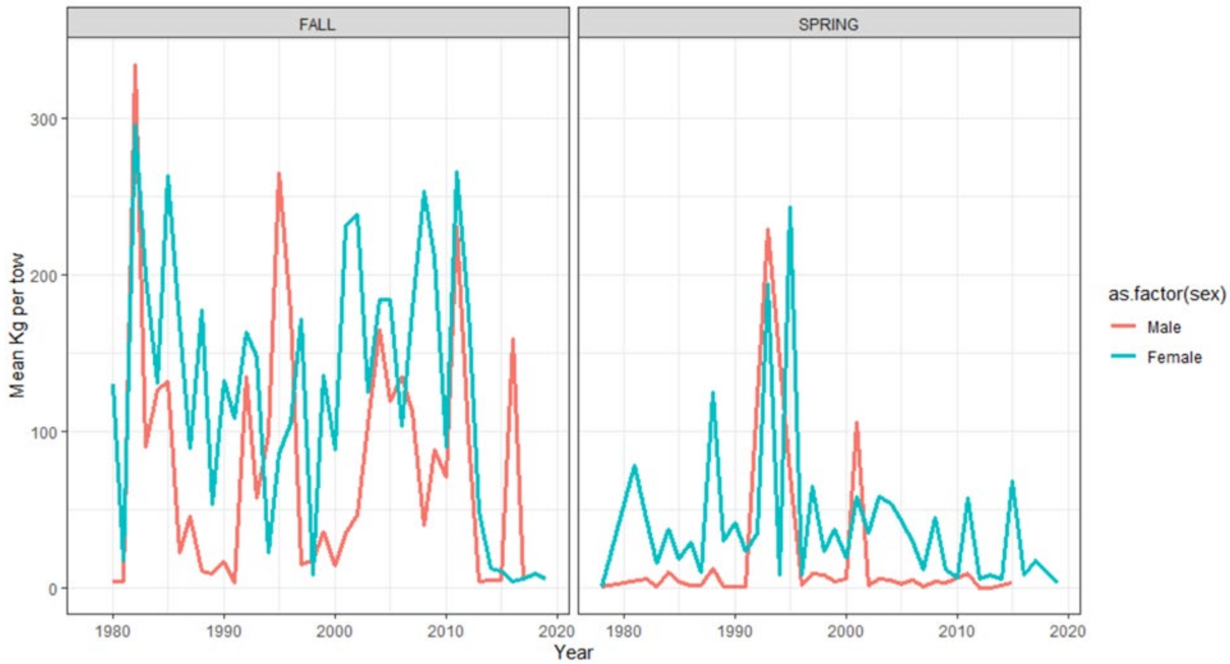


Figure 3.25. MADMF fall and spring survey mean kg/tow.

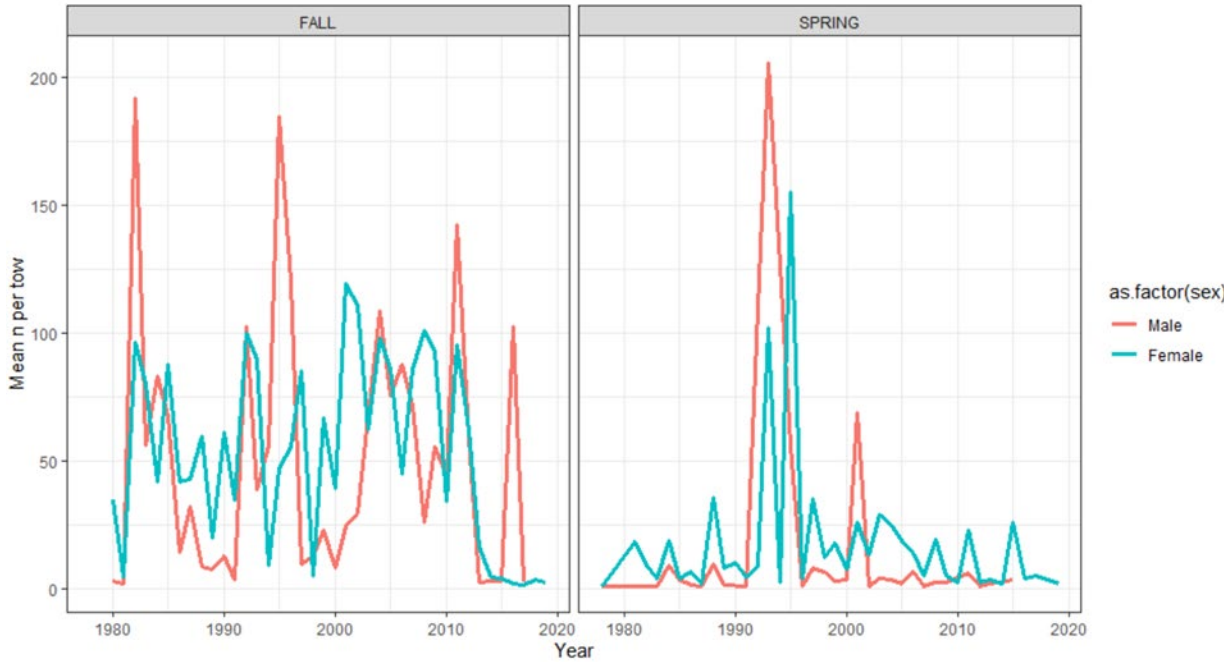


Figure 3.26. MADMF fall and spring survey mean numbers/tow.

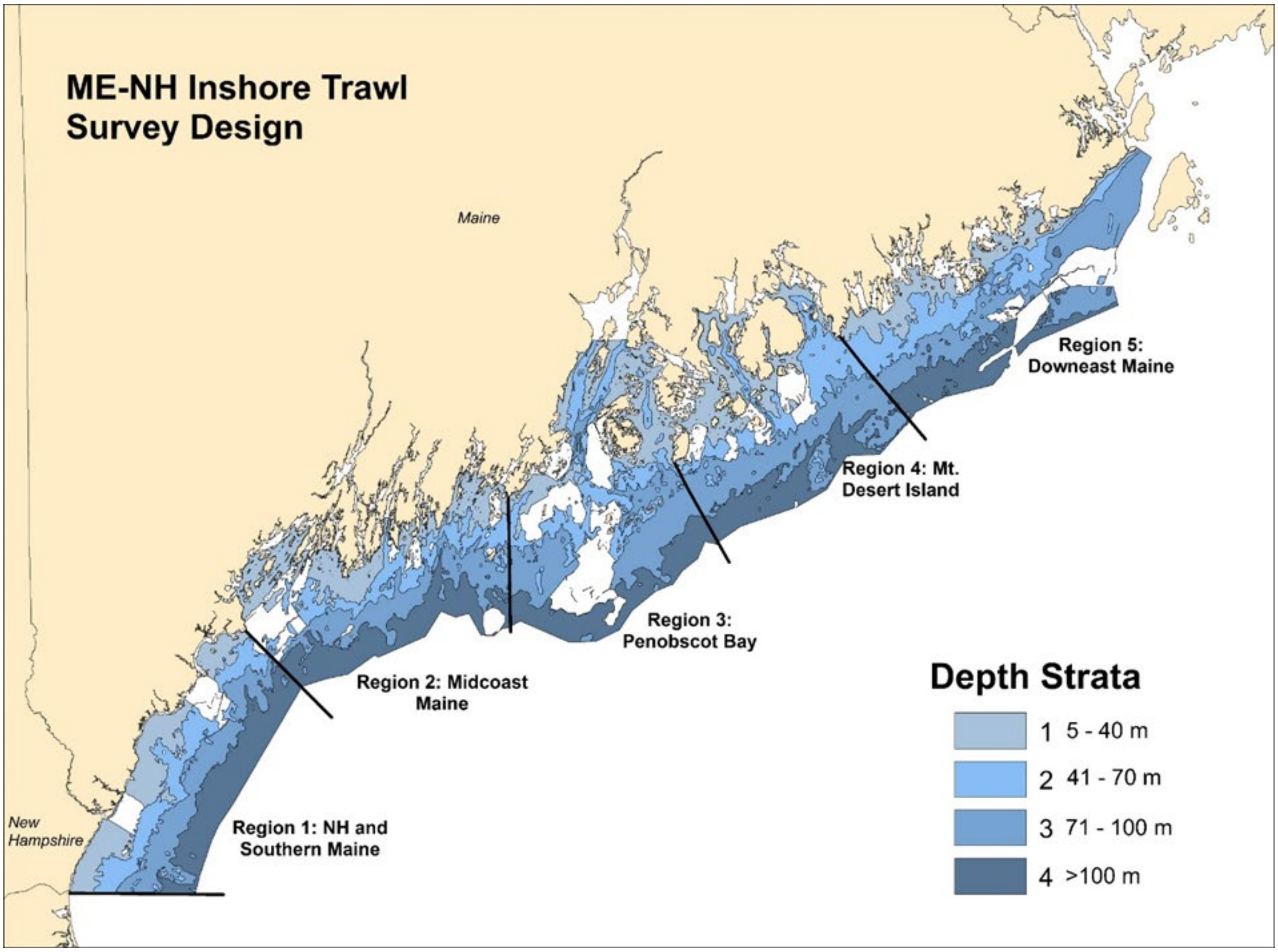


Figure 3.27. ME-NH Inshore Trawl Survey Strata.

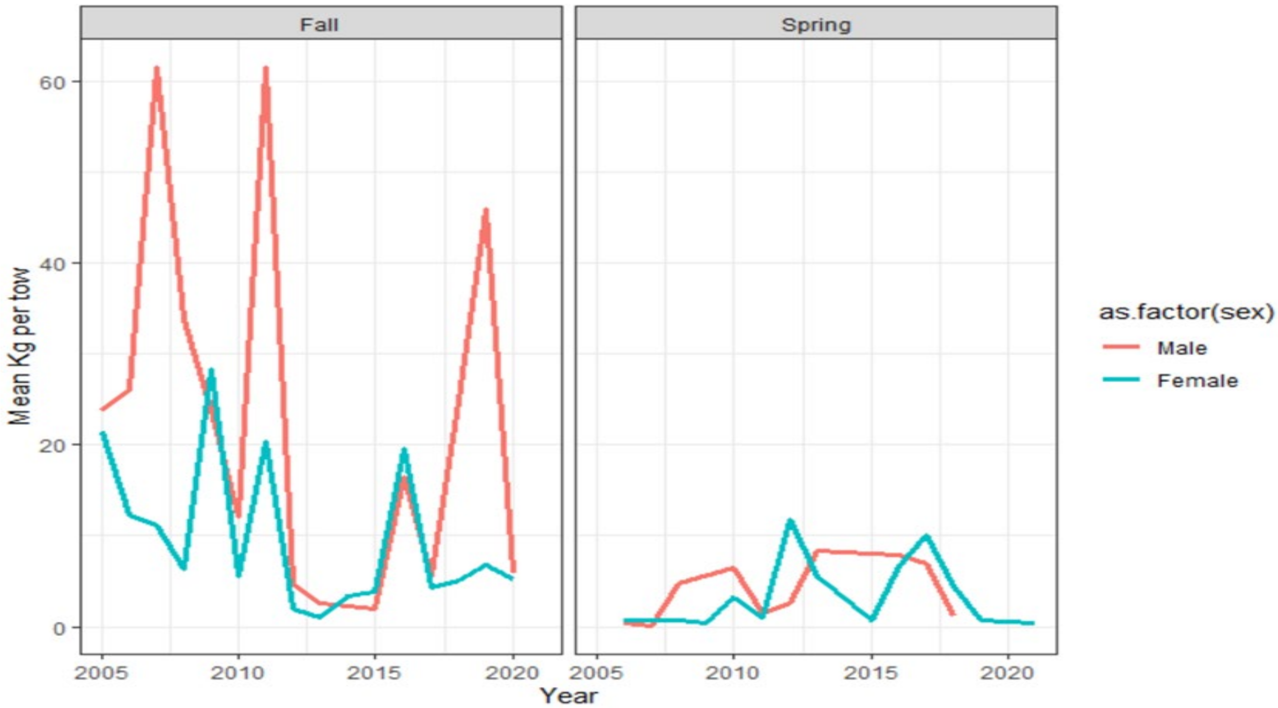


Figure 3.28. ME-NH Inshore Trawl fall and spring survey mean kg/tow.

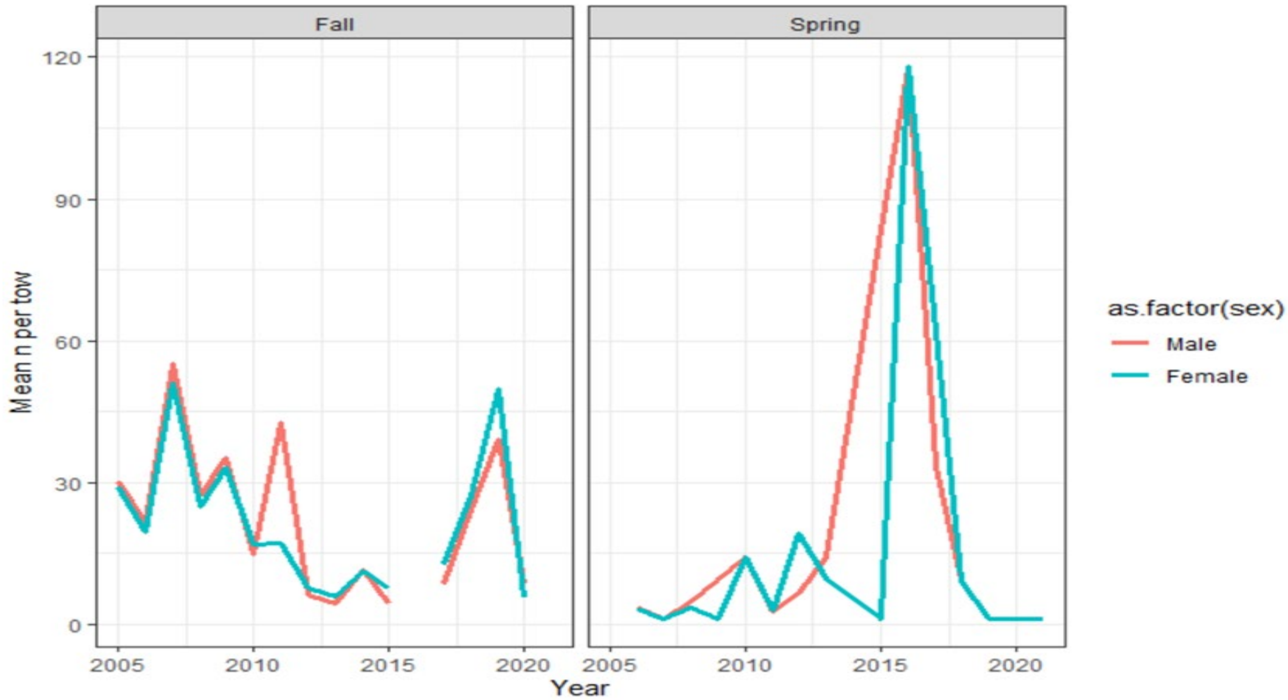


Figure 3.29. ME-NH Inshore Trawl fall and spring survey mean numbers/tow.

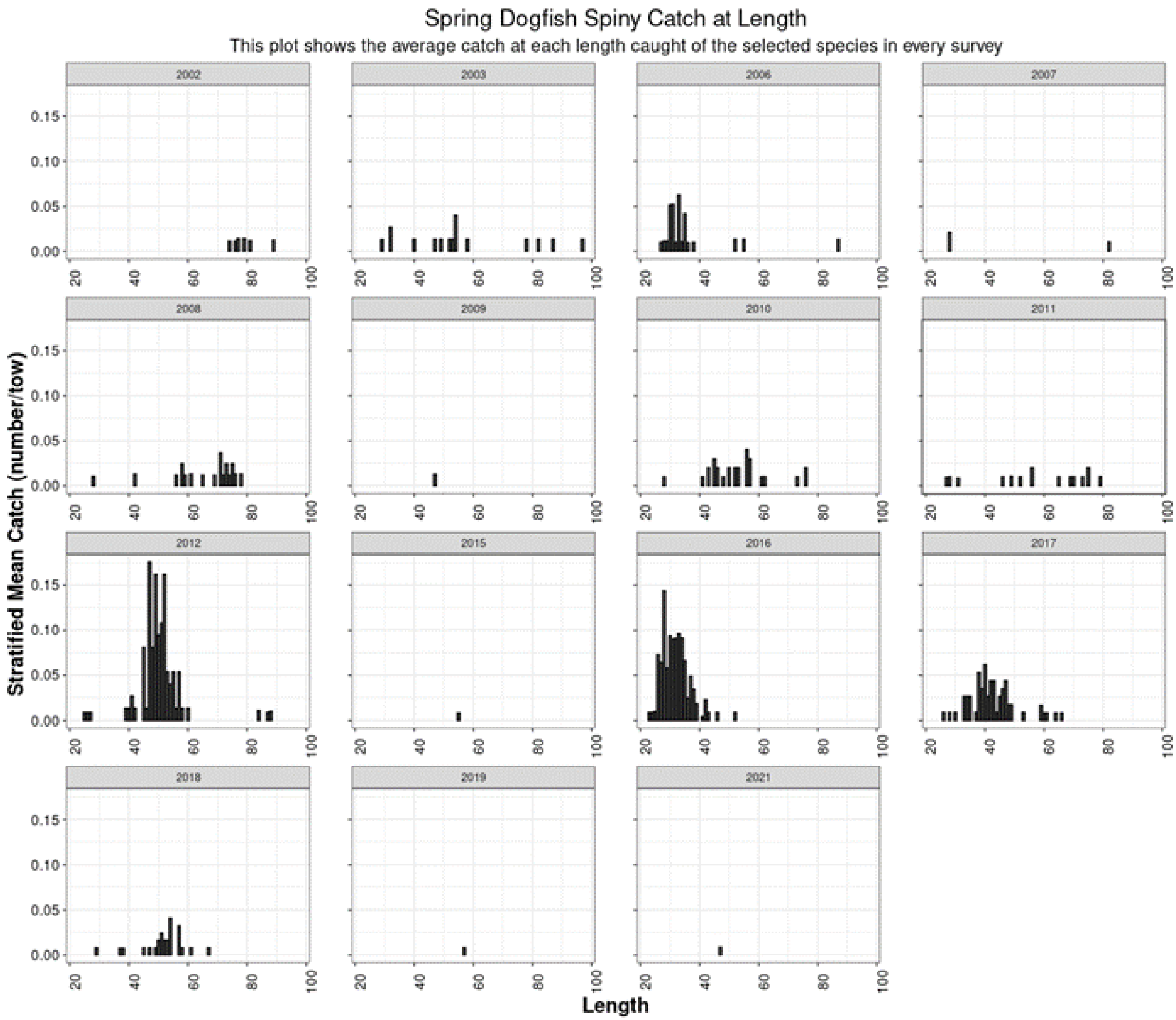


Figure 3.30. ME-NH spring bottom trawl survey length composition by year.

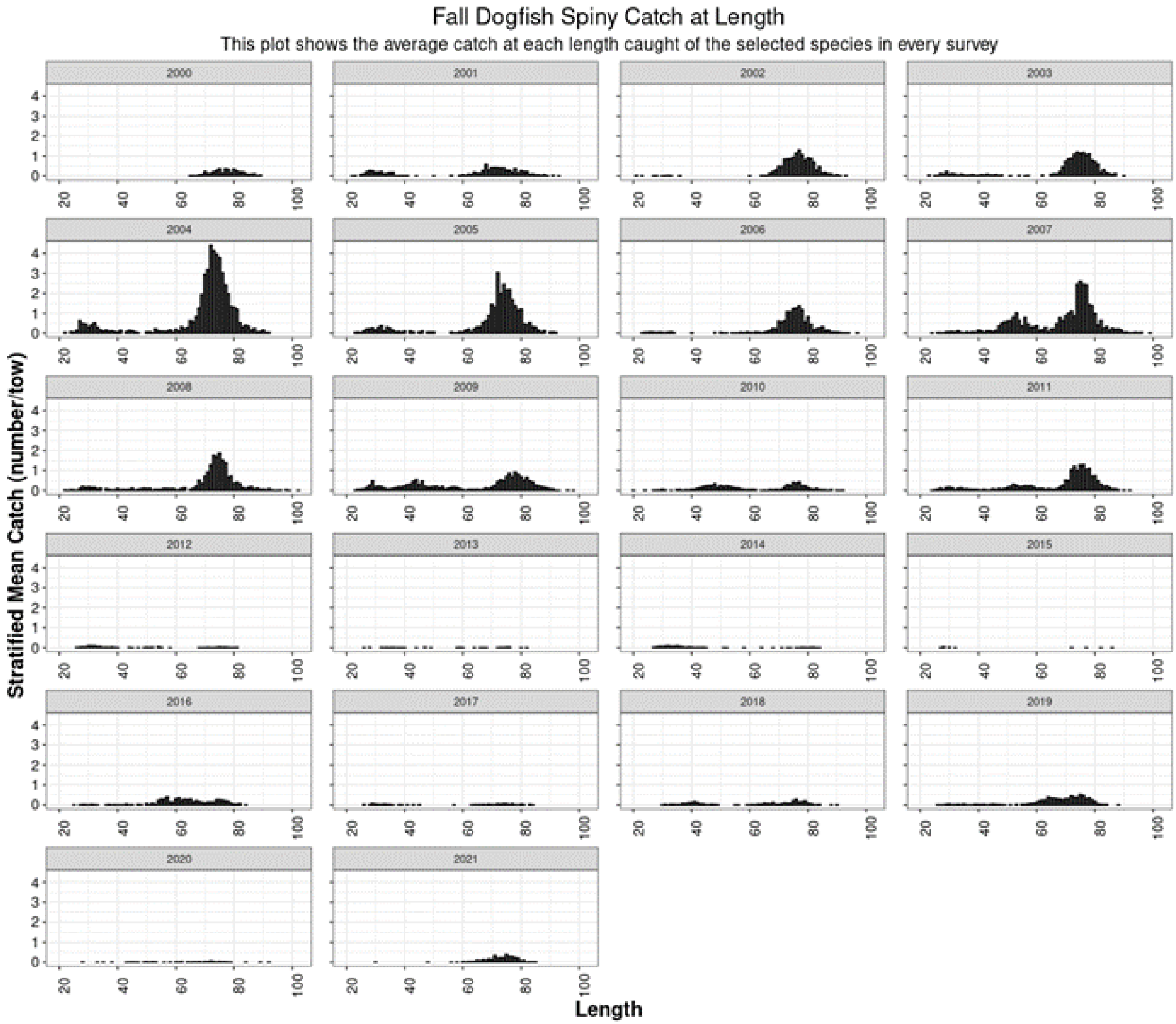


Figure 3.31. ME-NH fall bottom trawl survey length composition by year.

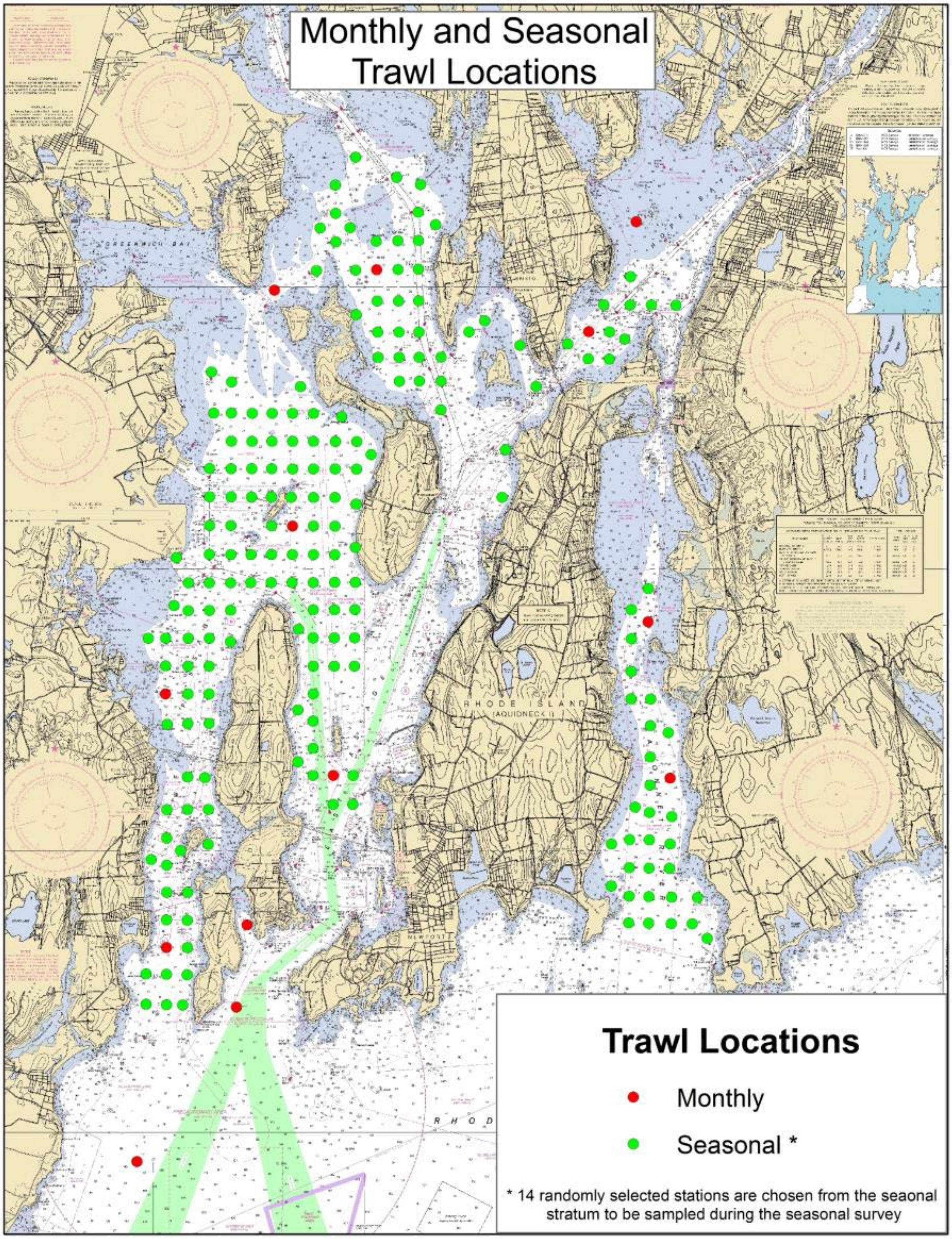


Figure 3.32. RI DEM Monthly and Seasonal Bottom Trawl locations in Narragansett Bay.

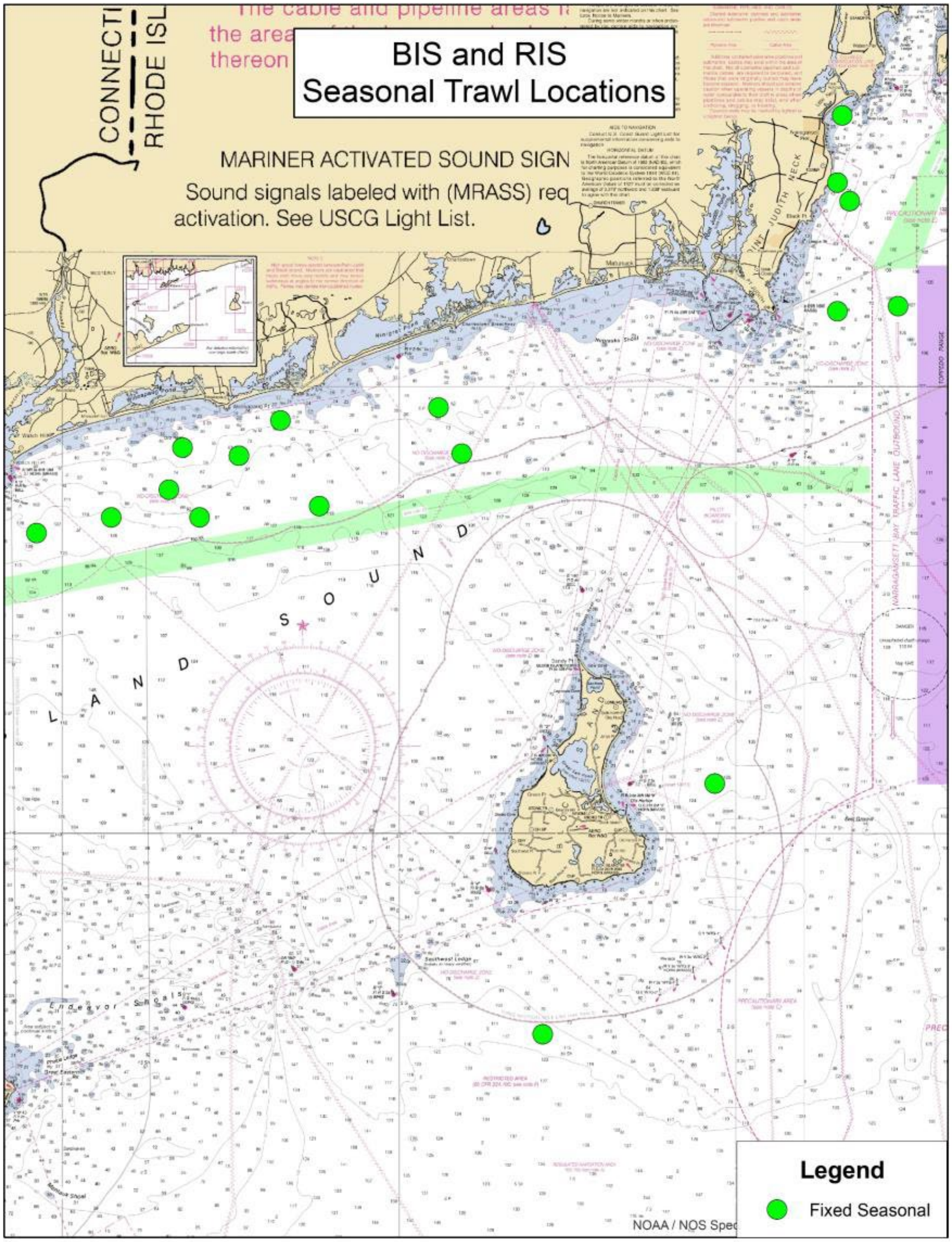


Figure 3.33. RI DEM Seasonal Bottom Trawl locations in Rhode Island and Block Island Sounds.

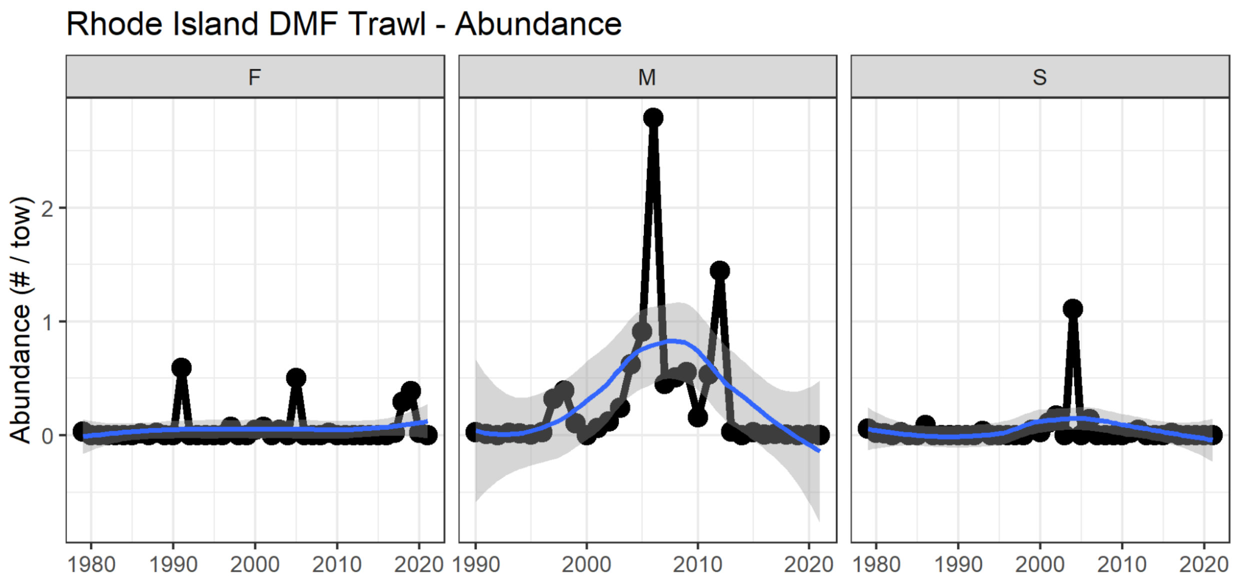


Figure 3.34. Rhode Island Fall (F) and Spring (S) Seasonal and Monthly (M) surveys mean numbers/tow.

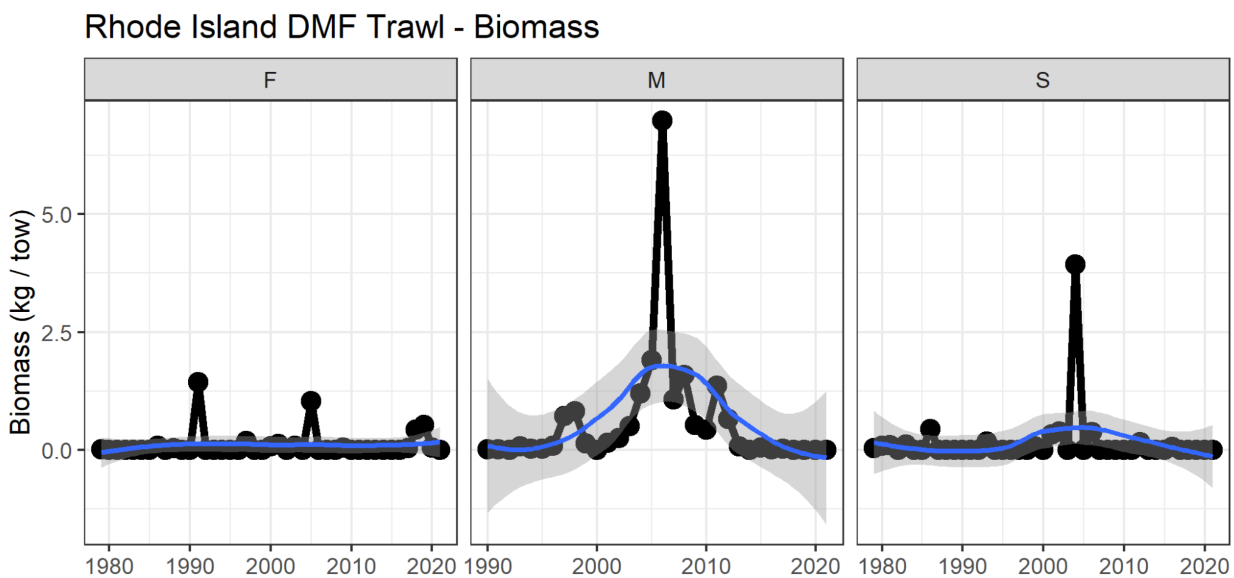


Figure 3.35. Rhode Island Fall (F) and Spring (S) Seasonal and Monthly (M) surveys mean kg/tow.

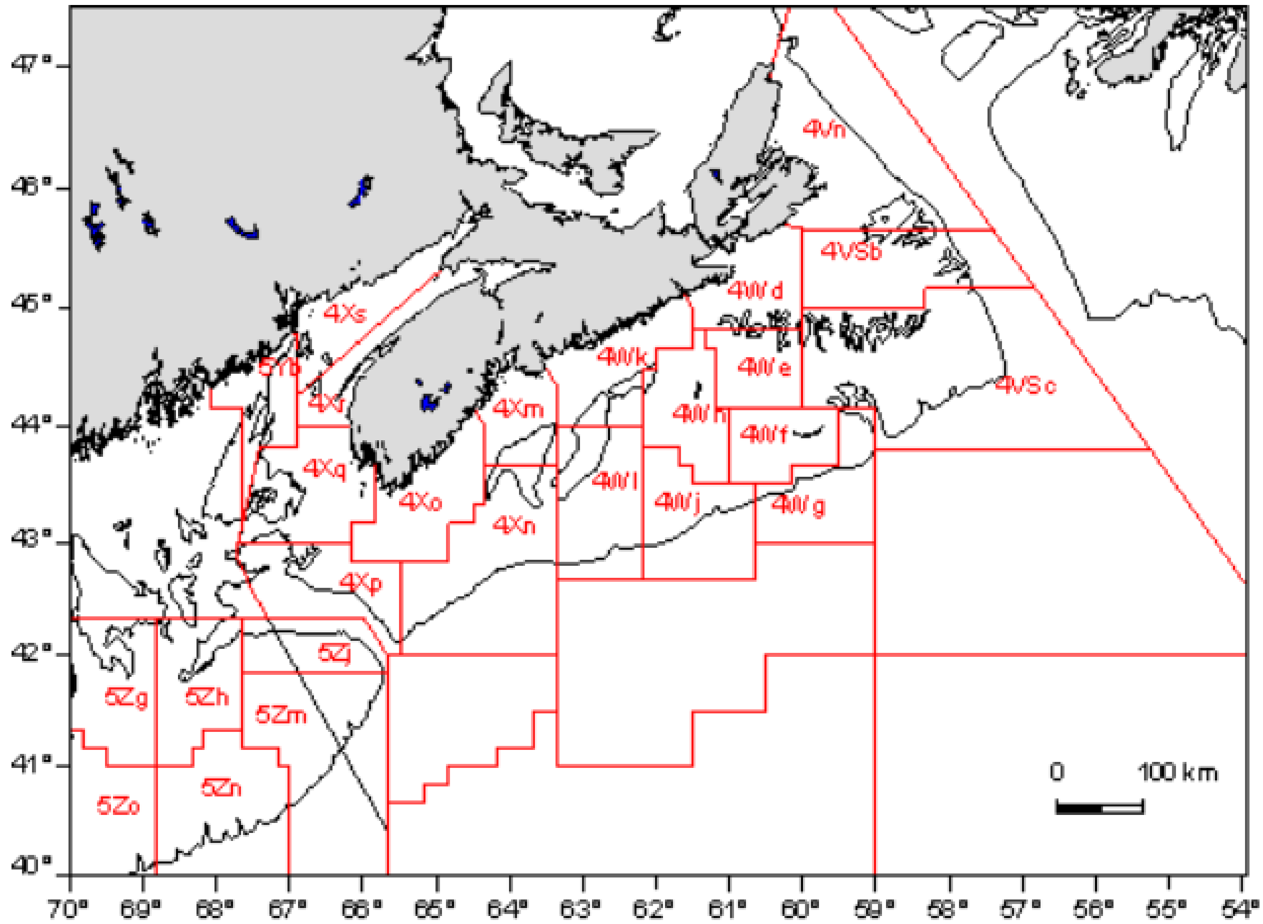


Figure 3.36. Northwest Atlantic Fisheries Organization Divisions

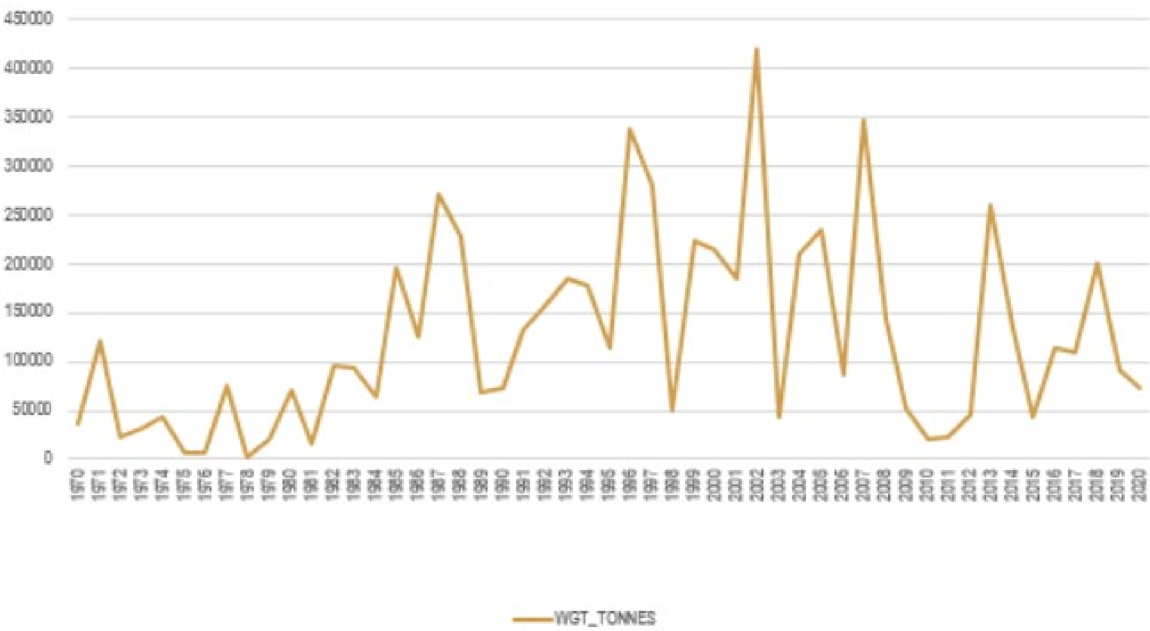


Figure 3.37. Design based biomass index for the Canada DFO Scotian Shelf Summer Survey.

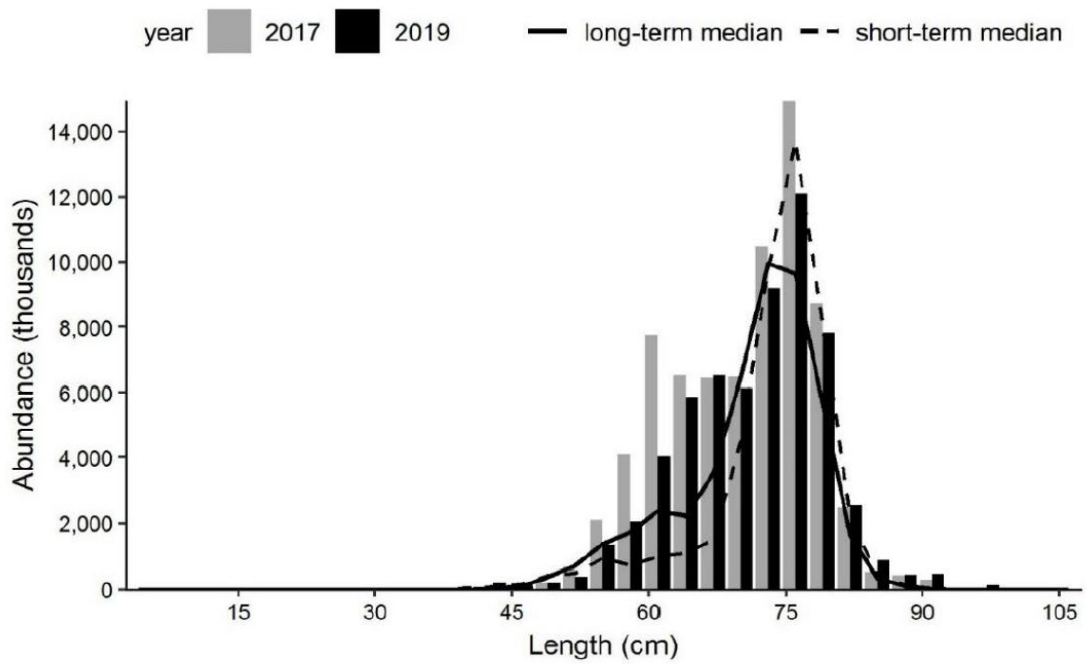


Figure 3.38. Scotian Shelf Survey length frequency. Gray and black bars represent the number in thousands at length for 2017 and 2019, respectively. The solid and dashed black lines represent the median in thousands at length for 1970–2017 and 2008-2017, respectively.

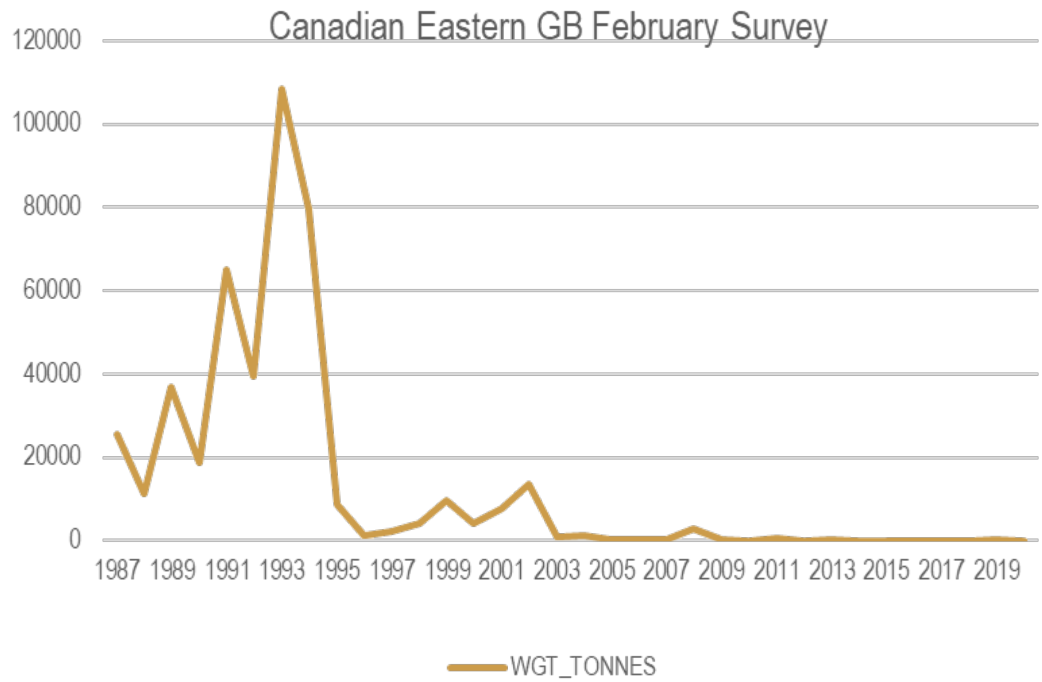


Figure 3.39. Design based biomass index for the Canada DFO Eastern Georges Bank Survey.

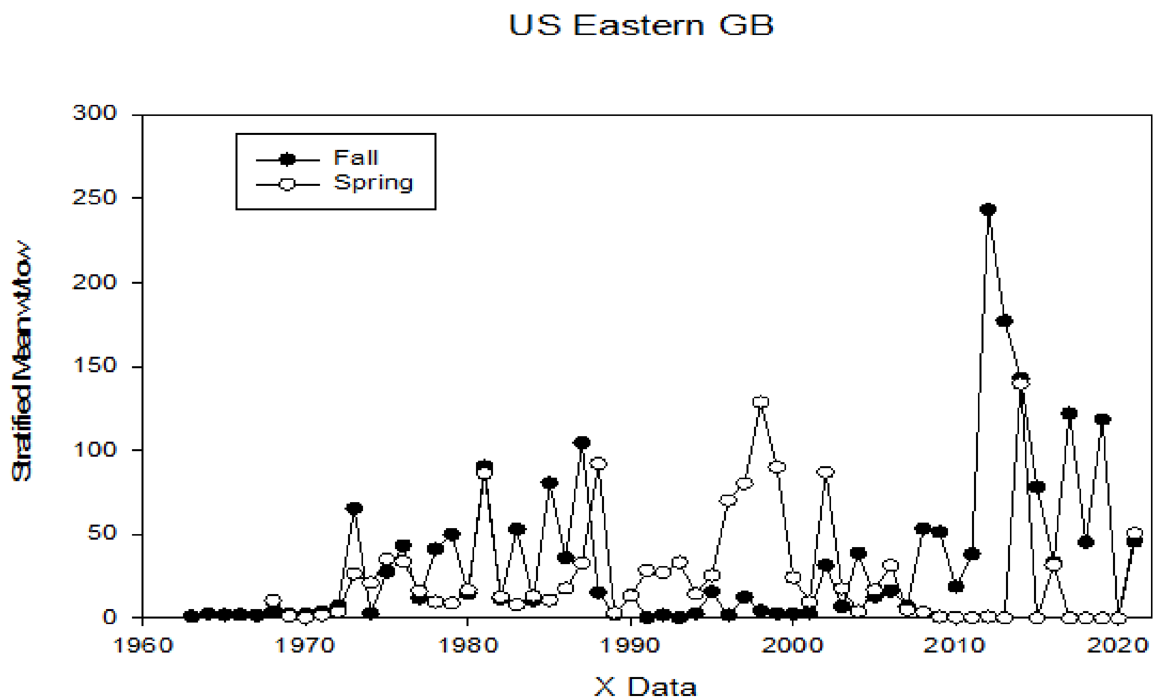


Figure 3.40. NEFSC Fall and Spring Bottom Trawl Survey stratified mean/tow for the US Eastern Georges Bank.

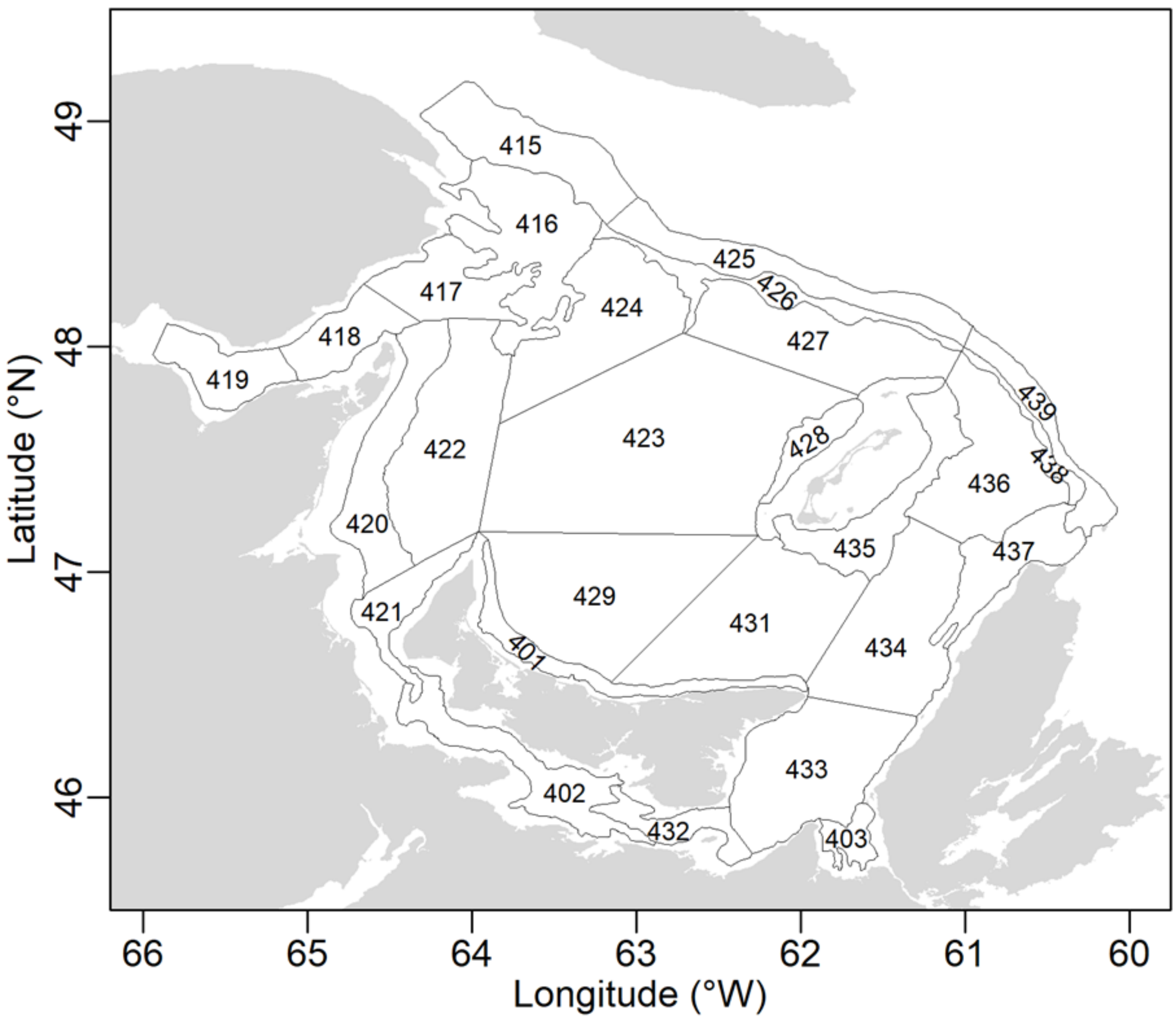


Figure 3.41. Canada DFO Southern Gulf of St. Lawrence survey strata.

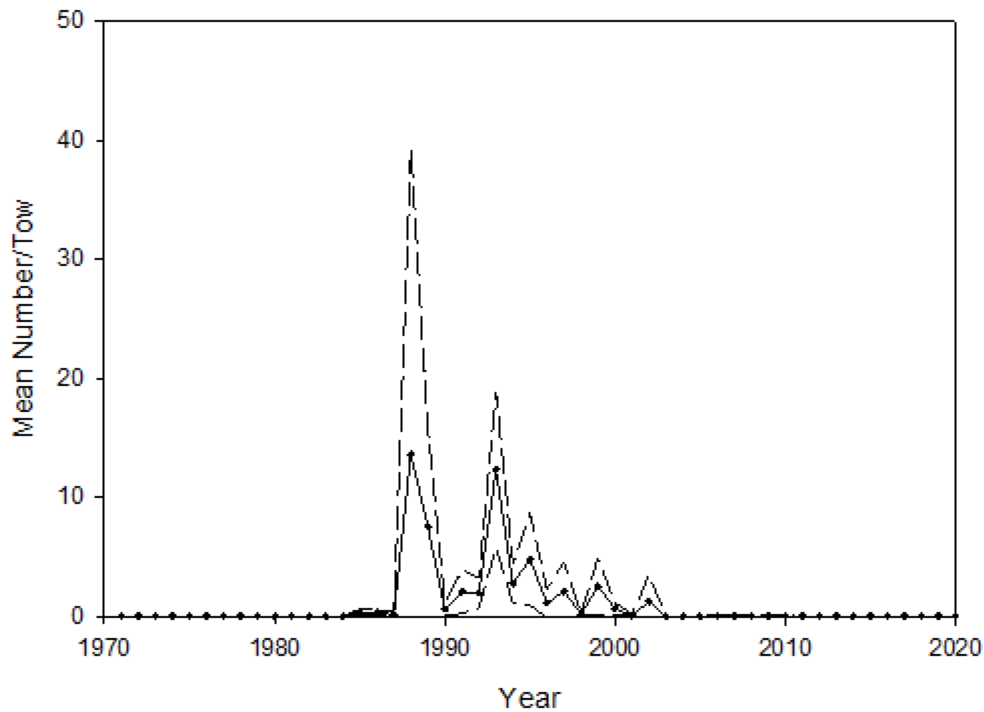


Figure 3.42. Annual mean numbers per tow for the Southern Gulf of St. Lawrence Survey

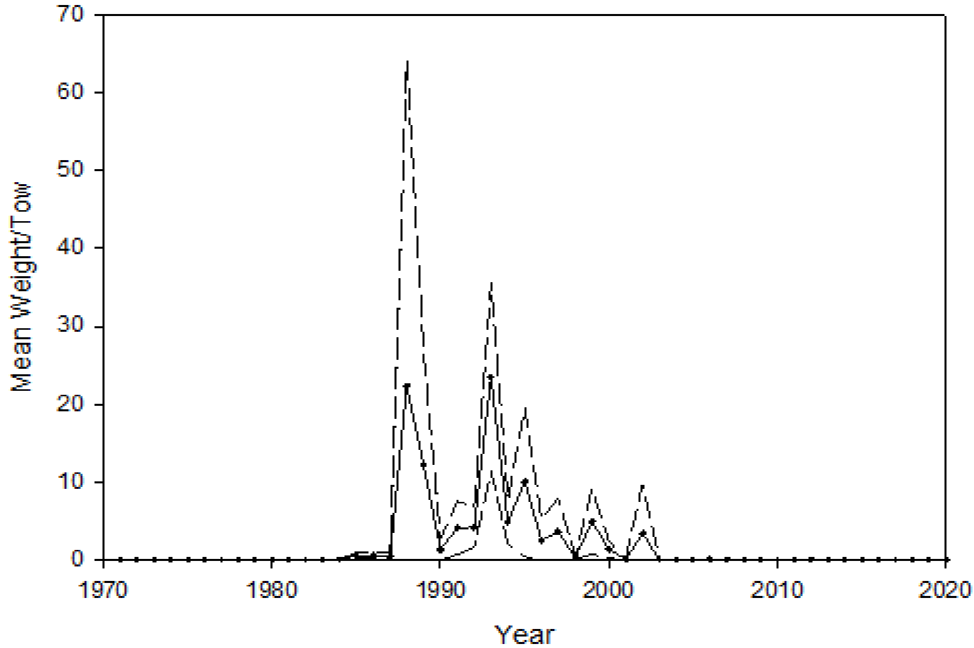


Figure 3.43. Annual mean weight (kg) per tow for the Southern Gulf of St. Lawrence Survey

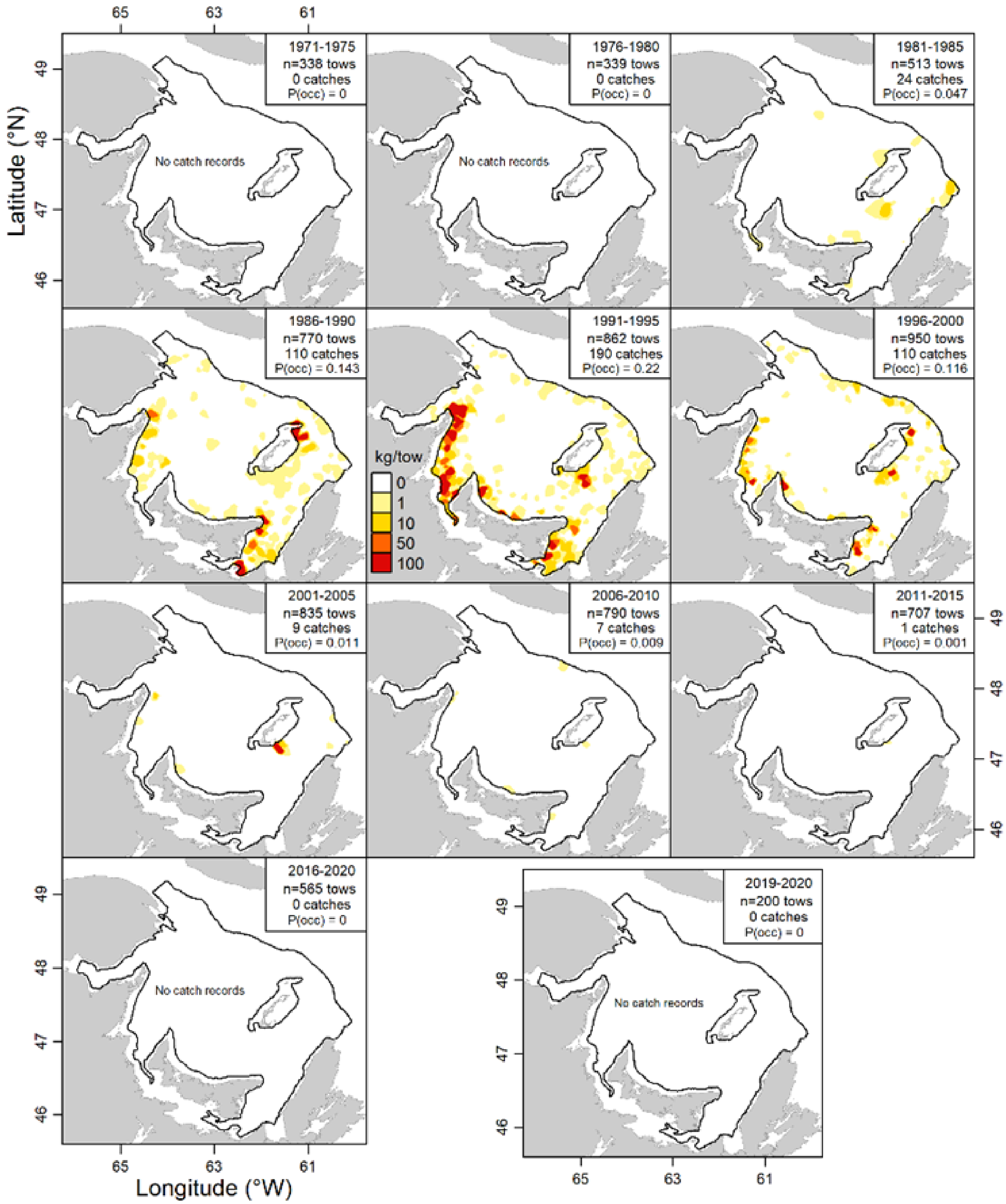


Figure 3.44. Distribution of survey mean weight (kg) per tow within the Southern Gulf of St Lawrence.

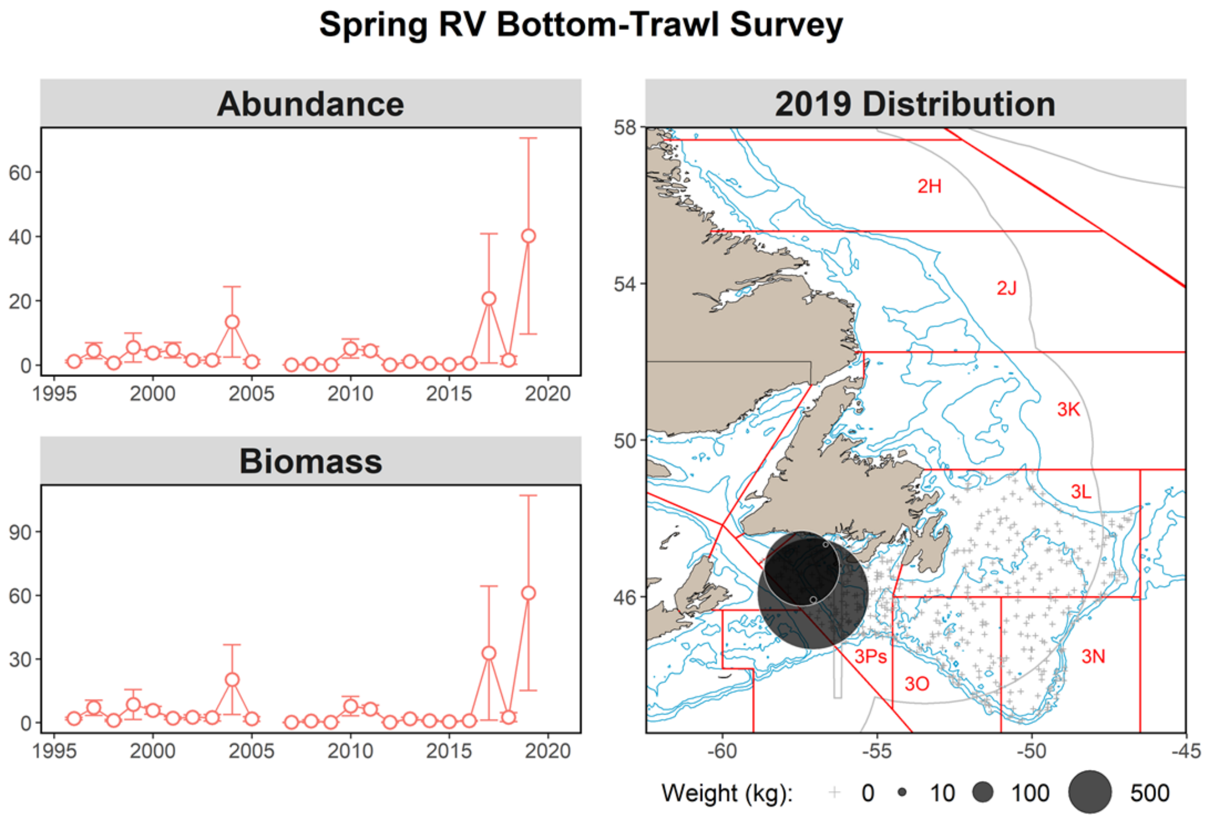


Figure 3.45. Canada DFO Spring Grand Banks abundance and biomass from 1996 - 2019 with 95% confidence intervals and weight (kg) per tow plotted within the survey area for the last year of the survey (2019).

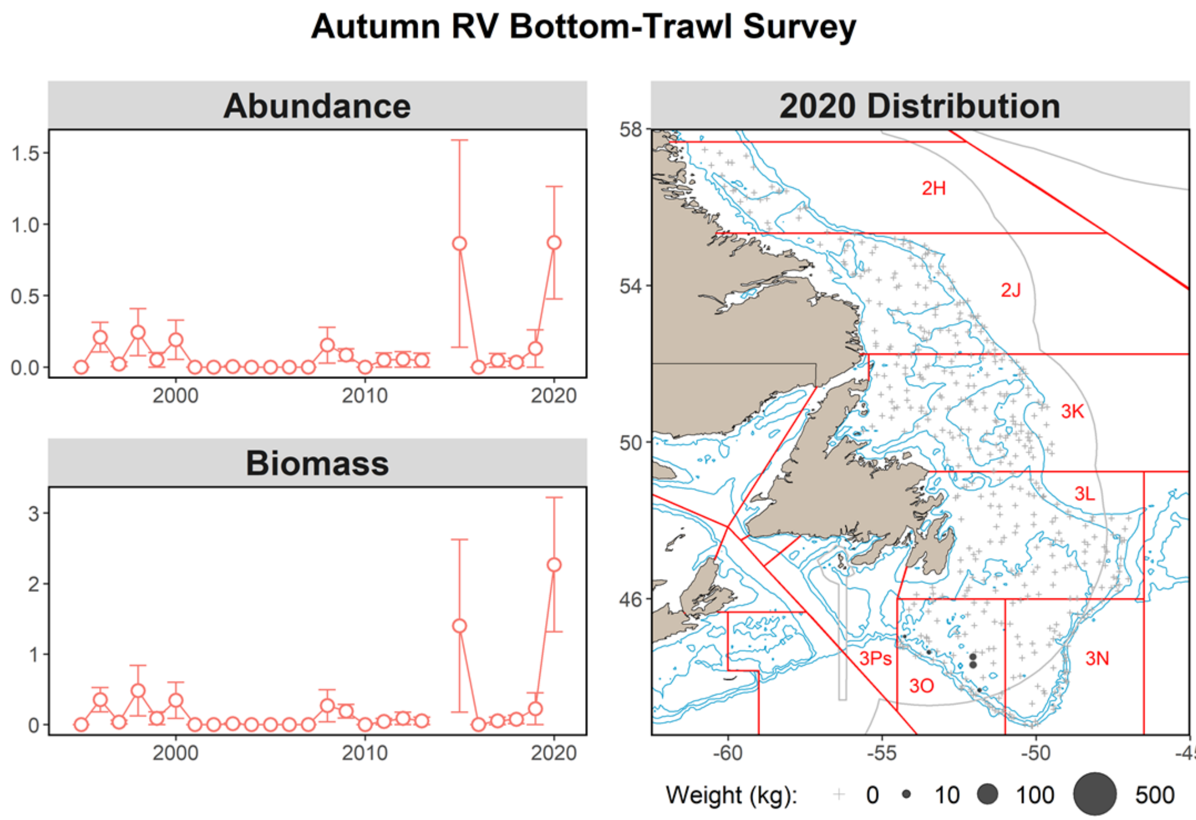


Figure 3.46. Canada DFO Fall Grand Banks and Labrador abundance and biomass from 1996 - 2020 with 95% confidence intervals and weight (kg) per tow plotted within the survey area for the last year of the survey (2020).

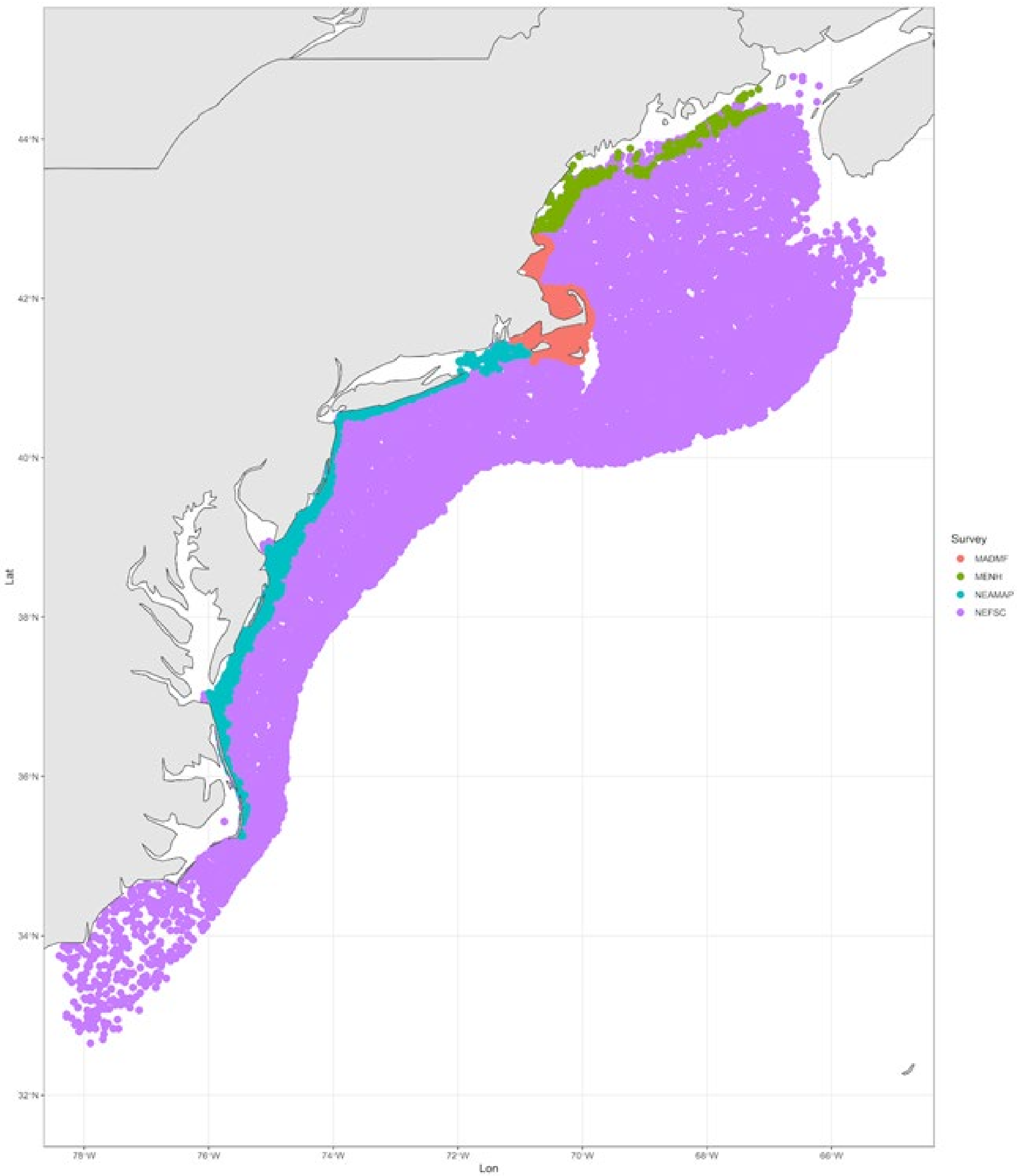


Figure 3.47: Survey data explored in VAST models for spiny dogfish.

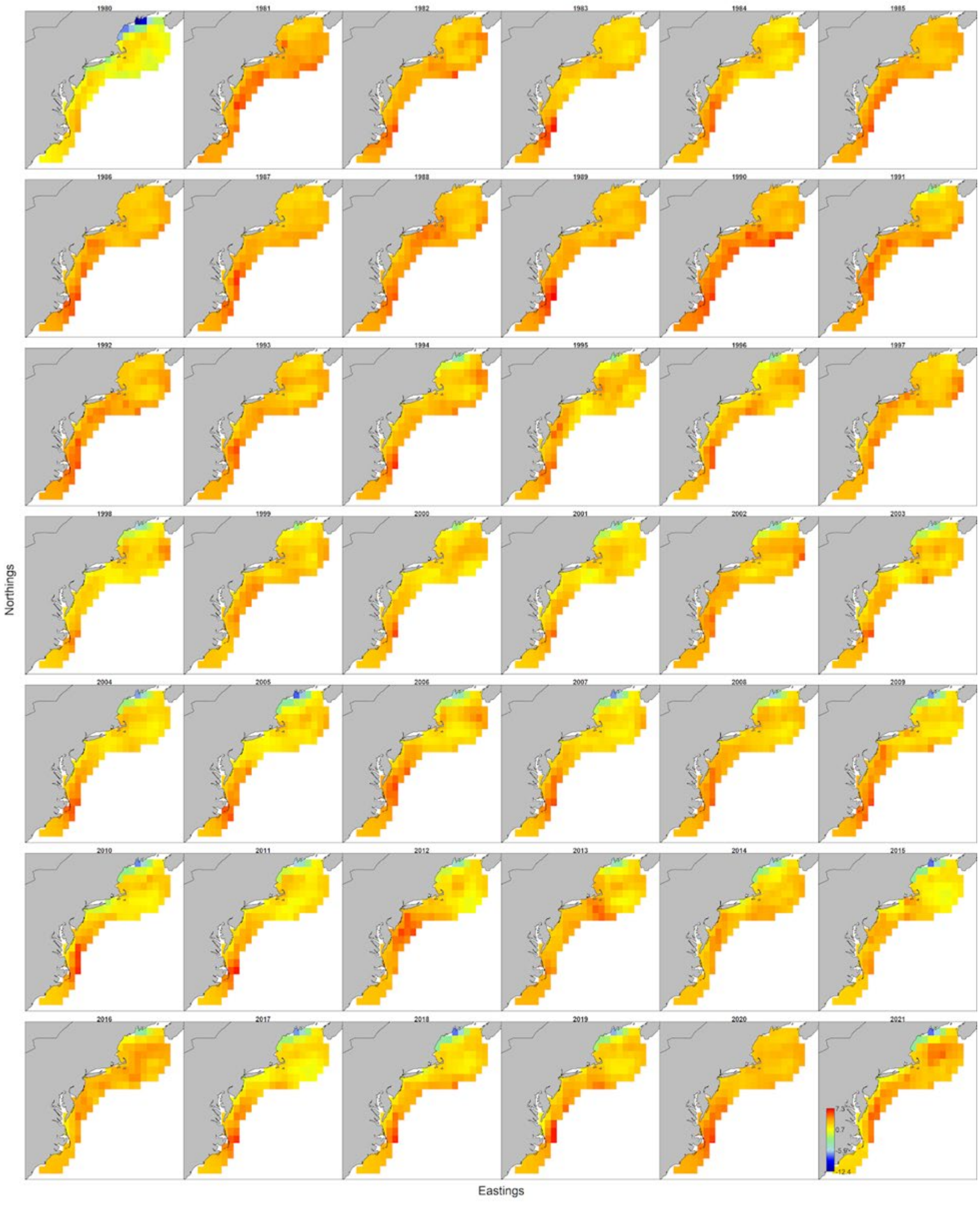


Figure 3.48: VAST estimated abundance for spiny dogfish in the spring.

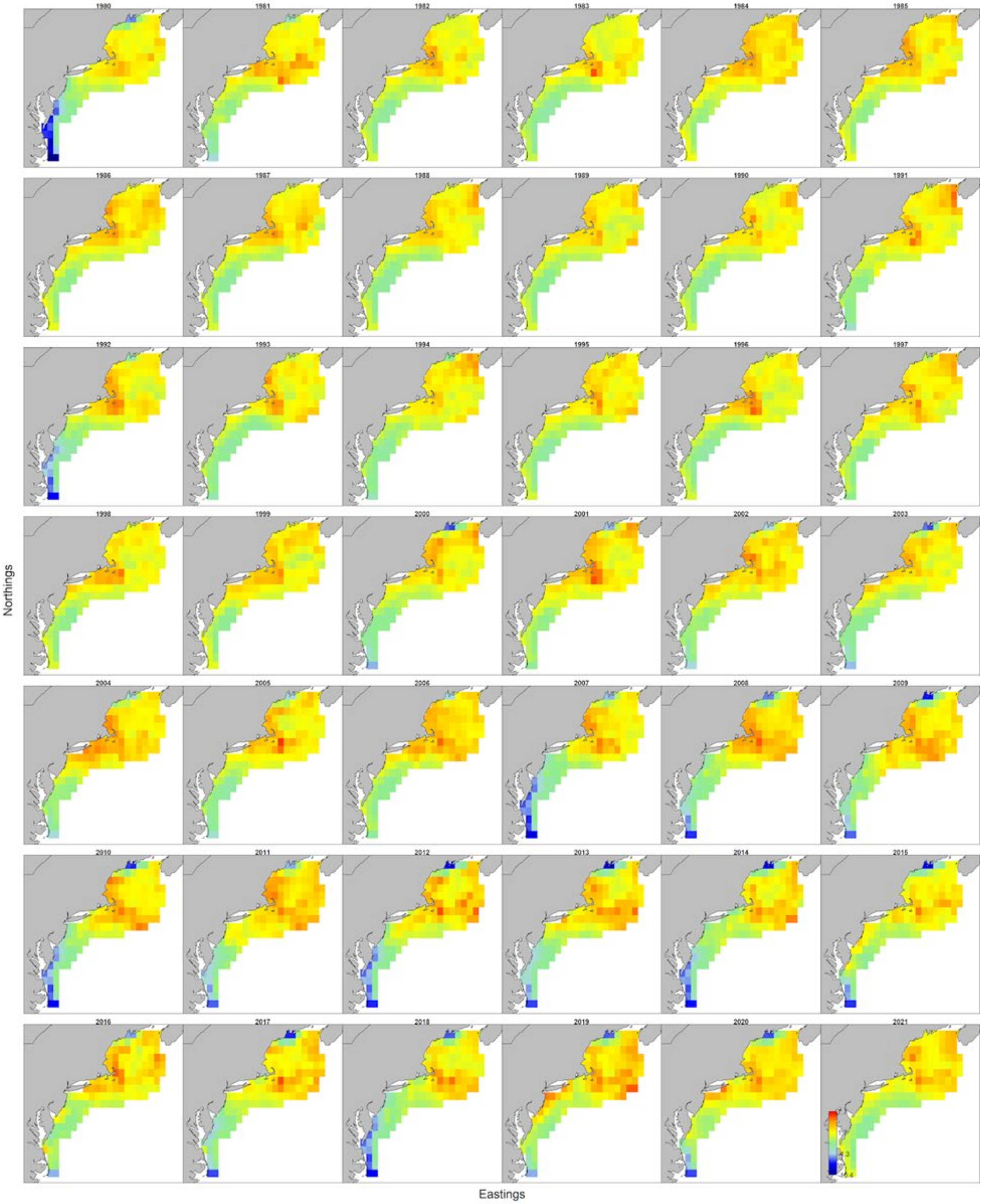


Figure 3.49: VAST estimated abundance for spiny dogfish in the fall.

Figure 3.50: Spring comparison of male relative abundance estimates produced by VAST and design based estimates for inshore and offshore surveys.

Figure 3.51: Spring comparison of female relative abundance estimates produced by VAST and design based estimates for inshore and offshore surveys.

Figure 3.52: Fall comparison of male relative abundance estimates produced by VAST and design based estimates for inshore and offshore surveys.

Figure 3.53: Fall comparison of female relative abundance estimates produced by VAST and design based estimates for inshore and offshore surveys.

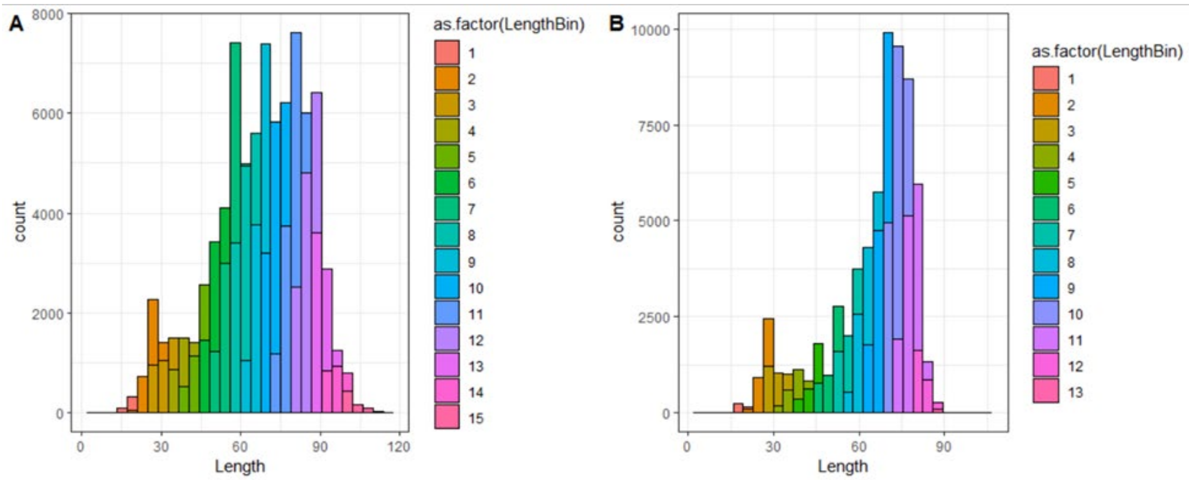


Figure 3.54: Spring length distribution and size bins used in VAST for female (A) and male (B) spiny dogfish

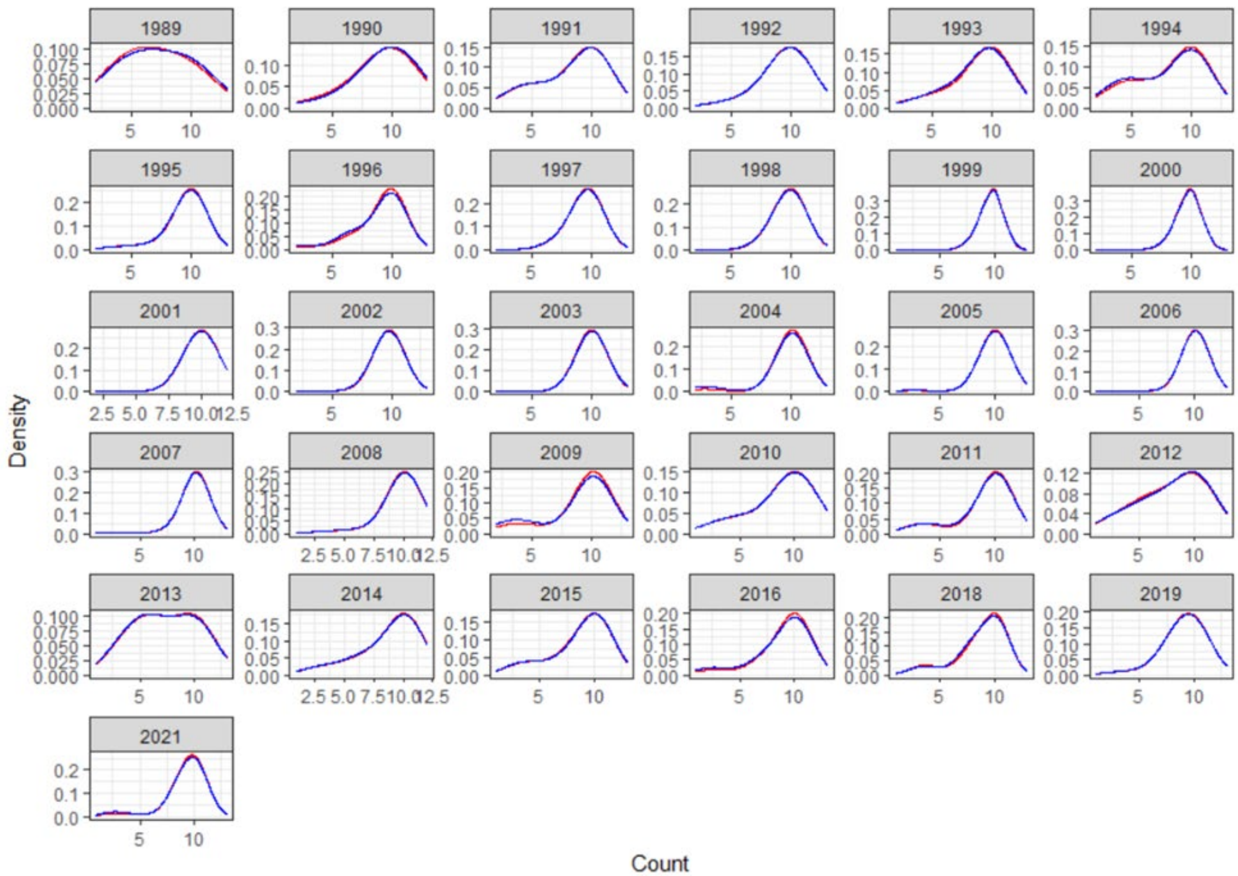


Figure 3.55: Comparison between design and VAST estimates of length composition for male spiny dogfish.

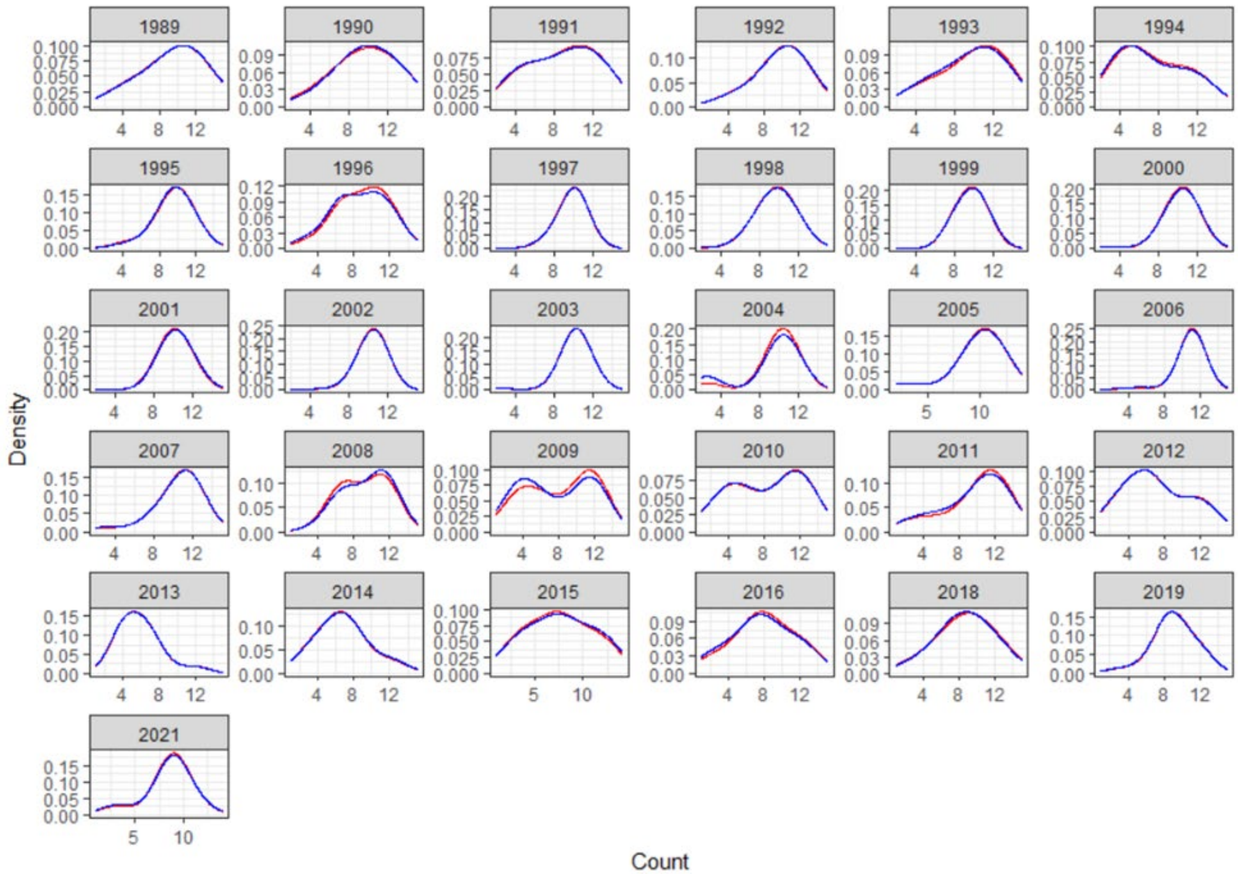


Figure 3.56: Comparison between design and VAST estimates of length composition for female spiny dogfish.

TOR4: ESTIMATE STOCK SIZE AND FISHING MORTALITY

"*Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.*"

Several approaches to stock assessment modeling were evaluated for this research track assessment. Ultimately, the Working Group proposed Stock Synthesis 3 (SS3) as the basis for status determination and fishery management advice (Chang et al. 2022). SS3 provides an analytical advancement over previous spiny dogfish assessments, because it incorporates biological characteristics and rates of the stock, as well as fishery dynamics into estimating stock conditions (e.g., spawning stock biomass, recruitment, fishing mortality).

A number of important life history processes and parameters were investigated and re-estimated in this assessment. These include length-weight relationships, maturity and pups at length for females, and natural mortality (Figure 4.1; see Anstead 2022a, Hart and Sosebee 2022, and Sosebee 2022a for details). Of particular interest is that the mean length at maturity has declined from around 80 cm in 1998 to 73 cm during 2012-2019 (Sosebee 2022a). This decline could be due to earlier maturation or slower growth or both. Natural mortality was chosen to decline with age (Lorenzen 1996), with a 50 year mean averaging 0.102 (Anstead 2022a).

There were also new investigations into growth using mark-recapture (McCandless 2022) and ageing (Passerotti and McCandless 2022) methods. The mark-recapture estimates were not appropriate for use in the assessment model, but provided supporting evidence concerning the decrease in length at maturity (McCandless 2002). Additionally, new ages were produced using the 2nd dorsal spine, but questions regarding the age estimates and uncertainties in the growth estimates prevented the Working Group from using these estimates directly in SS3 at this time (Passerotti and McCandless 2022).

Stock Synthesis

Model Configuration

An Atlantic spiny dogfish stock assessment model was developed in Stock Synthesis version 3.30.18 (SS3; Methot and Wetzel 2013) to provide an alternative to the index-based approach (Stochastic Estimator; NEFSC 2006) that was used in the previous assessments. SS3 is a statistical length-based age-structured population modeling framework. It is one of the most widely used stock assessment packages in the U.S. and globally (Dichmont 2016, 2021) and has many essential features of next-generation stock assessment models (Punt et al. 2020). Unlike most age-structured stock assessment models, SS3 can tune directly to length data, which is necessary when age data are lacking, as in Atlantic spiny dogfish. Additionally, SS3 can model sexes separately, an essential feature for a sexually dimorphic species such as spiny dogfish where the fishery targets only females. SS3 was recently used to assess Pacific spiny dogfish (Gertseva et al. 2021).

A sex-specific SS3 model was constructed for the Atlantic spiny dogfish to account for the life history and fishing differences between sexes. The SS3 runs were conducted solely on length data with assumed/estimated growth parameters within the model. While there was an effort to age Atlantic spiny dogfish and provide up-to-date age information for this assessment, due to several potential issues for the new age data, the Working Group decided not to use it for this assessment (Passerotti and McCandless 2022). Due to the uncertainty associated with growth, extensive sensitivity and profile analyses on various growth assumptions were conducted.

Catch data for the model included: commercial landings (metric tons) for U.S. and distant water commercial fisheries, U.S. recreational landings from 1962 to 2019, and discards from U.S. commercial fisheries and U.S. recreational landings from 1989 to 2019 (see TOR2). Both landings and discards data are available by gear type and summarized in Table 4.1. The discards were converted into dead discards using gear-specific discard mortalities and modeled as “catch” in SS3 (see TOR2). The commercial data by gear were aggregated into five modeled fleets (two fleets for landings and three fleets for discards) based on examining the similarities of their length compositions (Table 4.1 and Figures 4.2-4.3).

Spring NEFSC bottom trawl survey data were used as the primary abundance index for the SS3 modeling because that survey best covers the range of the stock (see TOR3). The survey has operated in the spring and fall since 1968. Fall data were not used because a greater portion of dogfish is outside of the bottom trawl survey domain in the fall due to seasonal migrations. The 2014 spring bottom trawl survey data were excluded from SS3 modeling because of missing data from critical survey strata in the Mid-Atlantic region. The annual stratified mean number per tow index was expanded using a factor of 5,260,450, the ratio of the total area surveyed divided by the swept area of a tow (wings only), the same expansion factor used in the Stochastic Estimator. This expansion allows the survey catchability (q) estimated in SS3 to be interpretable as gear efficiency combined with availability.

Additional abundance/biomass indices considered in SS3 modeling were the NEFSC bottom longline survey data (2014-2021; Nieland and McElroy 2022) and a vector auto-regressive spatio-temporal model-based index (VAST) that combined four trawl surveys from NEFSC (1980-2021), Massachusetts Division of Marine Fisheries (1980-2021); Maine/New Hampshire (2005-2021); and Northeast Area Monitoring and Assessment Program (NEAMAP; 2007-2021; Hansell and McManus 2022; see TOR3). These abundance/biomass indices, along with the NEFSC fall bottom trawl survey index, were included in SS3 as sensitivity runs.

The abundance/biomass indices are assumed to have a lognormal error structure, and the standard error of $\sqrt{ln(1+CV^2)}$ where CV is the coefficient of variation. A constant parameter added to the inputted standard error of the survey indices was estimated in SS3 for each survey.

Sex-specific length composition data from catch and survey for all fleets and years, except for the 2014 NEFSC spring bottom trawl survey, were available for this assessment. Total length data were partitioned into 31 length bins, from 20 to 110+ cm with a 3 cm increment. SS3 estimated population numbers at length (population length bins), structured the same as the length composition data. Length composition data were excluded and not used in the modeling when the effective sample size was one, or the number of length bins covered was less than five, as they are less credible (Figure 4.3). Comparing preliminary model runs using the complete data versus the reduced data showed no difference in population estimates, suggesting that the excluded data were not informative.

SS3 model runs started in 1989, the first year quantitative discard information was available from observer data. Discards before 1989 are a significant source of mortality for spiny dogfish (NEFSC 1994); thus, the Working Group was reluctant to start the model before 1989. Since fishing for dogfish occurred before 1989, an initial equilibrium catch was assumed, and initial fishing mortality was estimated for each fleet in SS3. The initial equilibrium catch by fleet was estimated using an average of the 1962-1988 catch data. Total landings from 1962 to 1988 were obtained from Sosebee (2019). Total discards from 1962 to 1988 were hindcasted using the observed ratio of discarded dogfish to landings of all species in 1989 from otter trawl and gill net fisheries (NEFSC 2006). Hindcasted total discards are likely underestimated because they only rely on two types of gears. Total landings and hindcasted total discards were assigned to each fleet using the averaged by-fleet proportion from the 1989-1993 catch data. An SS3 run starting the model from 1962 and assuming fishing morality to be negligible prior to 1962 was conducted in the sensitivity analysis.

Life history characteristics, including the sex-specific length-weight relationship, female maturity, and fecundity relationship, were updated using NEFSC bottom trawl survey data during this assessment and fixed at the updated values in SS3 (Hart and Sosebee 2022; Sosebee 2022a). During the preliminary model explorations, the Working Group found evidence of changing life history characteristics, including growth, maturity, and fecundity for Atlantic spiny dogfish in recent years. In particular, the estimated length at 50% maturity for females declined from 80 cm in 1998- 2011 to 73 cm during 2012-2019 (Figure 4.1; see Sosebee 2022a, Figure 1). Therefore, time blocks of 1989-2011 and 2012-2019 (referred to as biology blocks) were implemented in SS3 to allow growth, maturity, and fecundity to vary through time. Different growth, maturity, and fecundity parameter values were assumed/estimated for each block in SS3. Several sensitivity runs were conducted to examine the biology block assumption.

In the past assessments, the sex-specific growth for Atlantic spiny dogfish was assumed to follow a von Bertalanffy (VB) relationship estimated by Nammack et al. (1985; Table 4.2). A new growth study was conducted during this assessment to provide up-to-date growth information (Passerotti and McCandless 2022). During the preliminary model explorations, the new age data was compiled as conditional distributions of age-at-length, and VB growth parameters were estimated for each sex in SS3 (Figure 4.4). However, due to the high variability in length by age classes, especially for older females (Figure 4.4), the estimated standard deviations around the estimated growth curve were unrealistically large. As a result, the estimated selectivities for landings and surveys became dome-shaped, which the Working Group found unreasonable. SS3 runs that fixed the growth parameters at the values estimated by Passerotti and McCandless (2022) using the new growth data were also conducted. However, the results were similarly unrealistic. Given the uncertainties of the new growth data identified in Passerotti and McCandless (2022) and the unrealistic SS3 model results, the new growth data were not used in this assessment. Performances of the model using Nammack et al. (1985) growth and models with time-varying growth where the VB parameters were estimated for the biology block 2012-2019 were examined during the preliminary model explorations. The results showed a significant improvement in Akaike information criterion (AIC), resulting from the reduced VB asymptotic length (L_∞), especially for the females (Table 4.2-4.3). The reduction of L_∞ reflects the absence of large females in both catch and survey data for recent years (Figure 4.5). The Working Group decided to estimate L_∞ for both sexes in SS3 for the 2012-2019 period but fix the VB length at age-0 (L_{Amin}) and growth coefficient (k) at the values of Nammack et al. (1985) for the base case model. Sensitivity and profile analyses with various growth assumptions were conducted. The maximum age in SS3 was fixed at 50 years based on the approximate maximum age observed (Passerotti and McCandless 2022).

Sex-specific length-weight relationships in SS3 were estimated using NEFSC bottom trawl survey data from 1993 to 2019 from generalized linear mixed-effects models (Hart and Sosebee 2022; Figure 4.6):

$$W = 1.899348e-06L^{3.188} \text{ for females,} \quad (1)$$

$$W = 3.656515e-06L^{3.006} \text{ for males,} \quad (2)$$

where W is total weight (kg) and L is total length (cm).

Female maturity relationships were estimated for 1998-2011 and 2012-2019, respectively, using NEFSC bottom trawl survey data and used in SS3 (Sosebee 2022a; Figure 4.7):

$$Mat = \frac{1}{1+exp(0.4098361(79.9-L))} \text{ for biology block: 1989-2011,} \quad (3)$$

$$Mat = \frac{1}{1+exp(0.2832861(73.1-L))} \text{ for biology block: 2012-2019,} \quad (4)$$

where *Mat* is proportion mature and *L* is total length (cm).

Fecundity relationships were estimated for 1998-2011 and 2012-2019, respectively, using the pups/embryo data found in a subsample of female dogfish in the NEFSC bottom trawl survey and used in SS3 (Hart and Sosebee 2022; Figure 4.8):

$$P = 5.525074e-06L^{3.046335} \text{ for biology block: 1989-2011,} \quad (5)$$

$$P = 7.893089e-06L^{2.950182} \text{ for biology block: 2012-2019,} \quad (6)$$

where *P* is number of pups (age-0) and *L* is total length (cm).

The past Atlantic spiny dogfish assessments assumed a natural mortality (*M*) of 0.092 (Hoenig 1983; Rago et al. 1998). Several age-constant and age-varying *M* estimator approaches were evaluated for this assessment. Each approach required different life history parameters as inputs (see Anstead 2022a, Table 1). Many approaches were age-constant or time-invariant, providing one *M* estimate for all ages or lengths of spiny dogfish. Several age-constant approaches were revised and updated by Then et al. (2015), which were considered in this assessment. Two age-varying approaches were also used to consider different values of *M* by either age or length for spiny dogfish. All approaches were done by sex.

Life history parameters used in the *M* estimator approaches were tabulated by sex for spiny dogfish using various sources (see Anstead 2022a, Table 2). While the Working Group recommended the values in Table 2 in Anstead (2022a), other values were considered, including those for maximum age (Nammack et al. 1985), VB growth parameters (Campana et al. 2009;

Bubley et al. 2012), and length-weight relationship parameters (Wigley et al. 2003). As part of the 2022 assessment, the growth and length-weight relationship were re-estimated using updated data (Hart and Sosebee 2022; Passerotti and McCandless 2022). Several issues were identified in the growth analysis, so the values from Nammack et al. (1985) were used for the *M* estimators that use growth parameters, although the revised length-weight parameters were used.

The Working Group decided that approaches that rely heavily on the VB growth rate coefficient, *k*, should not be used for spiny dogfish (e.g., Alverson and Carney 1975, Jensen 1996). The Working Group supported the length-varying Lorenzen (1996) estimates by sex that were scaled to the average Then et al. (2015) estimate (*M* = 0.102) being used for the base case model (Figure 4.9). Sensitivity runs were conducted to examine various *M* assumptions.

Spawner-recruitment (SR) relationship in SS3 models the relationships between age-0 fish and spawning output, i.e., the number of pups the mature females produced (1,000s) at the beginning of each year (Methot et al. 2021). Ricker, Beverton-Holt, and survival SR relationships were explored during this assessment. The survival SR relationship developed by Taylor et al. (2013) is an SR model that explicitly models the survival between embryos and age-0 recruits, which is particularly useful for low fecundity species that produce fewer offspring per litter and exhibit a more direct relationship between spawning output and recruitment (Taylor et al. 2013; Methot et al. 2021). The survival SR relationship was assumed for the Pacific spiny dogfish assessment (Gertseva et al. 2021) and is parameterized as (Taylor et al. 2013):

$$R_y = SSB_y e^{ln(S_0)(1 - Z_{frac}(1 - \frac{SSB_y}{SSB_0}^{\beta}))} \qquad (7)$$

where R_y is recruitment in year *y*, SSB_y is spawning output in year *y*, $S_0 = \frac{R_0}{SSB_0}$ is survival of per-recruit individuals at unfished equilibrium, R_0 is unexploited equilibrium recruitment, SSB_0 is the corresponding equilibrium spawning output, β is a shape parameter controlling the shape of the density-dependent relationship between $\frac{SSB_y}{SSB_0}$ and S_0 (with limit $\beta > 1$), and Z_{frac} is a fraction of pre-recruit instantaneous mortality rate at equilibrium ($-\,ln\,(S_0)$) and range $0 < Z_{frac} < 1$.

During the preliminary model explorations, the parameters for all three SR models were estimated within SS3, and model results were compared. The SS3 model with the Beverton- Holt SR relationship failed to converge, and the models that assumed Ricker and survivorship SR relationships showed very differently estimated stock trajectories. Thus, the Working Group decided to estimate the SR relationship outside of SS3, fix the SR parameters in SS3 at these values, and then compare their model performances.

The Ricker and Beverton-Holt SR relationships parameterized by a and b were estimated using the NEFSC bottom trawl survey data (McManus et al. 2022). The survivorship SR relationship was explored using the same data set (with S_0 and SSB_0 estimated by averages of various SS3 preliminary runs) but failed to converge because the two parameters Z_{frac} and β are highly correlated. Therefore, the survivorship SR parameters estimated in a preliminary model run (Z_{frac} = 0.93 and β = 1.6) were assumed for exploratory SS3 runs.

In SS3, the Ricker and Beverton-Holt SR models were parameterized using $ln\ (R_0)$, the steepness parameter (*h*; Methot and Wetzel 2013). To estimate the Ricker and Beverton-Holt steepness from the *a* and *b* form models, S_0 is required (Miller and Brooks 2021):

$$h = \frac{a\phi_0}{4+a\phi_0} \text{ for Beverton-Holt SR model,} \qquad (8)$$

$$h = \frac{(a\phi_0)^{\frac{4}{5}}}{5} \text{ for Ricker SR model,} \qquad (9)$$

where $\phi_0 = \frac{1}{S_0}$ can be interpreted as unexploited spawning per recruit. The survivorship SR relationship is not parameterized in the form of steepness in SS3, but steepness was calculated for comparison purposes. S_0 is also required to estimate steepness for the survivorship SR parameters (Taylor et al. 2013):

$$h = 0.2e^{S_0 Z_{frac}(1-0.2^{\beta})} \qquad (10)$$

To get an estimate of S_0 , various preliminary SS3 runs were examined. The estimated S_0 in SS3 is invariant with different model settings, e.g., maturity, fecundity, SR relationships, but varies with natural mortality. Therefore, three S_0 values derived using three *M* assumptions, static *M* = 0.092 (Hoenig 1983), static *M* = 0.102 (Then et al. 2015), and Lorenzen (1996) *M* scaled to an average of 0.102 were assumed, steepness were estimated from these values for the Ricker and Beverton-Holt SR models, and SS3 runs were conducted with the fixed steepness values. For the survivorship SR relationship, parameters were fixed at Z_{frac} = 0.93 and β = 1.6, and model runs were conducted with three different M assumptions.

The estimated steepness was around 0.4 for *M* = 0.092, around 0.3 for *M* = 0.102, and around 0.2 for scaled Lorenzen (1996) *M* for both Ricker and Beverton-Holt SR models. However, the steepness is around 1 for *M* = 0.092, around 0.8 for *M* = 0.102, and around 0.6 for scaled Lorenzen *M* for the survivorship SR models. AIC values from these runs suggested that survivorship SR outperformed Ricker and Beverton-Holt models regardless of *M* assumptions; the survivorship SR model coupled with *M* = 0.102 performed the best, followed by the scaled Lorenzen (1996) *M*. These conclusions were the same with or without estimating recruitment deviations in the model.

Because assuming *M* = 0.102 resulted in an unrealistically high steepness/productivity for spiny dogfish, a long-lived and low fecundity stock, the Working Group decided to assume a survivorship SR relationship, coupled with the Lorenzen (1996) *M* scaled to an average of 0.102 as the base case model configuration. The survivorship SR parameters were updated based on a profile analysis and fixed at Z_{frac} = 0.9, β = 1.5, and σ_R = 0.3 (standard deviation of log recruitment deviations) for the base case model. Recruitment deviations were estimated for the entire time series and bias-adjusted so that the estimated recruitments are mean unbiased (Methot and Taylor 2011; Methot et al. 2021). Uncertainty of the SR relationship assumptions were further explored in the sensitivity and profile analysis.

A double normal selectivity function was assumed for all six fleets in SS3 to fit the length composition data for its ability to estimate either an asymptotic or a domed-shaped selectivity pattern from data (Methot and Wetzel 2013; Methot et al. 2021). The double normal selectivity function has six parameters: p1 - peak value, p2 - top logistic, p3 - ascending width, p4 - descending width, p5 - selectivity at first length bin, and p6 - selectivity at last length bin. The sex-specific selectivity was estimated using a parameter offset approach with a maximal selectivity greater than or equal to one for the dominant sex and an additional parameter to determine the relative apical selectivity value for the offset sex. The selectivity parameters allowed to be offset in SS3 are p1, p3, p4, and p6. For the catch fleets 1-5, male selectivity was estimated as an offset from the female parameters, so the maximum selectivity for both sexes is one; thus, the resulting apical fishing mortality is comparable among fleets. The shape of the selectivities was freely estimated in SS3 for all fleets. Parameters p5 and p6 were skipped for all fleets, except for p5 for the discard fleet 5 and survey because they caught small dogfish. The offset of descending parameter p4 for landings fleets and the survey was turned off because it was estimated at zero during the preliminary model explorations. Selectivity time blocks were implemented for the NEFSC spring bottom trawl survey to estimate different selectivities for the two different research vessels conducting the survey: RV *Albatross IV* (1989-2008) and FRV *Henry B. Bigelow* (2009-2019). A sensitivity run was conducted to examine the selectivity time block assumption.

Three data weighting approaches were explored to rescale the effective sample size to reduce conflicts between data sources during the preliminary model exploration: McAllister-Ianelli, Francis, and Dirichlet-Multinomial (McAllister and Ianelli 1997; Francis and Hilborn 2011; Thorson et al. 2017). The scalers estimated using McAllister-Ianelli and Francis data weighting approach significantly down-weighted the survey length composition data relative to the catch length composition data. Thus, the Working Group decided to use the Dirichlet-Multinomial data weighting approach, which involves estimating a parameter (θ) to scale each fleet's inputted effective sample size. For comparison purposes, the θ parameter was fixed at the base case value for the jitter and profile analysis but re-estimated for the retrospective analysis. Sensitivity analysis was conducted without weighting the length composition data.

In summary, the parameters fixed in SS3 include length-weight, maturity, fecundity, SR relationships, growth for the first biology block, and the fixed p4-6 parameters mentioned in the selectivity paragraph above. Within the estimated parameters, the peak, ascending, and apical selectivity parameters were time-varying for fleet 6, and L_{∞} for both sexes were estimated for biology block 2012-2019. Non-informative priors were used for all the parameters except for the θ parameter for the Dirichlet-Multinomial error distribution used to weight the length data. A

Normal N(0, 1.813) prior was assumed for $ln\ (\theta)$ to counteract the log transformation effect between θ and data weighting (Methot et al. 2021).

The model convergence was evaluated based on whether the final gradient is < 0.0001 and whether the Hessian matrix for the parameter estimates is positive definite. Parameters estimated at a bound were examined, and correlations between estimated parameters were produced to see if highly correlated parameter pairs or non-informative parameters exist for possible unstable model or model misspecification. The residual analysis proposed by Carvalho et al. (2021) was performed on indices and length composition data to check for model fits. Profile of R_0, jitter, and retrospective analyses were also conducted to check for data consistency and model stability (Carvalho et al. 2021).

Model Results

The base case model converged (gradient 2.3 x 10^{-5}) and the Hessian matrix was positive definite. All parameters were estimated within their bounds, correlations between parameters were low (< 0.95), and all parameters were informative (correlation > 0.01). The 100 iterations of jittering the starting values by 10% resulted in 60% of the runs converging at the total likelihood value of the base case (-23409.9) and above the base case total likelihood value for the rest of the runs with a maximum change of 36.6 in likelihood. This result indicated that the base case model is slightly sensitive to starting values but stable and is likely to converge at a global rather than a local minimum.

The overall model fit of the abundance index data and length composition data was evaluated using joint-index residual plots from the fit to the index data and the mean length of the length composition data (Carvalho et al. 2021). The residual plot for the NEFSC spring bottom trawl survey index showed a residual pattern where the residuals are positive during the 1990s, negative during the 2000s, and positive in recent years, with RMSE = 39.6% (Figure 4.10). The residual plot for mean length of the length composition data showed a good fit with RMSE = 6.3%. The loess-smoother of this plot indicated a positive residual pattern at the beginning of the time series but no apparent residual pattern for recent years (Figure 4.11). The above analysis indicated a reasonably good overall fit to the data for the base case model.

The time-varying growth curve and the assumed/estimated VB growth parameters by sex are shown in Table 4.2 and Figure 4.12. The estimated L_∞ for the biology block 2012-2019 were smaller than those estimated by Nammack et al. (1985) for both sexes. The reduction is more significant for females (11.26 cm) than males (3.35 cm) and is likely reflecting the absence of large females in both catch and survey data (Figure 4.5).

The observed and model-predicted NEFSC spring bottom trawl abundance index is shown in Figure 4.13. The predicted index is within the 95% uncertainty level, except for 2004. The estimated catchability q was 0.83 for this survey.

The estimated selectivities by sex and fleet are shown in Figures 4.14-4.19. The estimated selectivities were asymptotic (logistic) for all landings fleets and NEFSC spring bottom trawl survey (fleets 1, 2, and 6) and dome-shaped for all discard fleets (3-5; Table 4.1). Estimated apical male selectivity was smaller than females for landings and discard fleets (1-5; Table 4.1), which is reasonable for a female-targeted fishery. Time-varying selectivity for the NEFSC spring bottom trawl survey showed an increased selectivity for small dogfish and reduced selectivity for the median-sized females during the *Bigelow* period (2009-2019), which is consistent with the survey data. Figure of length compositions from 2005 to 2012 showed systematic changes between the *Albatross* to *Bigelow* period for both sexes (Figure 4.20).

The observed and model-predicted length compositions aggregated across time by fleet and sex are shown in Figure 4.21. The fits to the aggregated length compositions appear to be fairly accurate, suggesting that the estimated fisheries and survey selectivities are reasonable. The observed and model-predicted annual length composition data and the residuals from the fits by fleet and sex are shown in Figures 4.22-4.33. Fit to the annual length composition data showed some systematic poor fit for the large females for the landings fleets (1 and 2) and the survey, as well as the median size males for the survey. There were large residuals for small (around 30 cm) dogfish for fleets 1, 3, and 4 and large dogfish for fleets 3 and 4. The fixed survivorship SR relationship, along with the estimated recruitment from both the SR relationship and recruitment deviations, are shown in Figure 4.34. The estimated recruitment decreased from 1989 to the early 2000s, when the lowest recruitments of the entire time series were estimated, followed by a large increase through 2010, and then dropped to half of the peak value and stayed stable since (Table 4.4 and Figure 4.35). The estimated time series of total biomass by sex and spawning output are provided in Table 4.4 and Figure 4.36. The estimated spawning output declined during the beginning of the time series, increased starting in the early 2000s, peaked in 2012, and then decreased since. The estimated annual fishing mortality, which is defined as the number-based exploitation rate for age 12+ dogfish (roughly age at 50% fishery selectivity), peaked around 1989 to 1999, decreased to the lowest point in 2005, and stayed around 0.03 in recent years (Table 4.4 and Figure 4.36).

Sensitivity Analysis

For the base case model, L_∞ was the only growth parameter estimated for the biology block 2012-2019. The sensitivity of this assumption was examined with three additional runs:

- estimating L_∞ and k but fixing L_{Amin} at the Nammack et al. (1985) values,
- estimating all three growth parameters L_∞, k, and L_{Amin}, and
- fixing L_∞, k, and L_{Amin} at the Nammack et al. (1985) values

for both sexes for the biology block 2012-2019. The estimated spawning output from the two growth scenarios with estimating two or all three VB parameters are similar to the estimates from the base case model, with slightly higher terminal spawning outputs (Figure 4.37). However, the run assuming Nammack et al. (1985) growth produced a very different spawning output trajectory than the base case model (Figure 4.37). The estimated L_∞ is similar with or without estimating k and L_{Amin} (Table 4.2). The estimated k is slightly higher than that estimated by the Nammack et al. (1985) study. Although runs estimating two or all three VB parameters performed better than the base-case model, the differences in AIC were small (Table 4.3). When the VB growth parameters were fixed at the Nammack et al. (1985) values, the AIC was much worse. These results support the Working Group's decision on estimating the L_∞ for the biology block 2012-2019 for the base case model.

Sensitivity runs were performed assuming:

- M = 0.092 (Hoenig 1983) for all ages and sexes, as used in the previous assessments,

- *M* = 0.102 for all ages and sexes derived using Then et al. (2015) method, and
- the sex- and age-specific Lorenzen (1996) *M* scaled to asymptote at 0.102.

These were compared to the base case model where the sex- and age-specific Lorenzen (1996) *M* was scaled to an average of 0.102. A summary of performance statistics and several critical parameter estimates for these runs can be found in Table 4.3. The two static natural mortality runs performed better than the base case in AIC, likely contributed by the higher *M* for older dogfish (Figure 4.38). However, the estimated NEFSC spring bottom trawl survey *q* and steepness *h* were both over 1 for the static natural mortality runs, indicating possible model misspecifications. This supports the Working Group's decision not to use static natural mortality for the base case model. The run with Lorenzen (1996) *M* scaled to asymptote at 0.102, which assumed the highest natural mortality at age of all the runs, performed worse than the base case. The estimated spawning output for this run is much higher than the two static *M* runs and the base case model (Figure 4.39).

The performance of the base case model with a fixed survivorship SR relationship and estimated recruitment deviations was compared to two additional sensitivity runs:

- fixed Ricker SR parameters with recruitment deviations and
- fixed Beverton-Holt SR parameters with recruitment deviations.

The Ricker and Beverton-Holt SR relationship parameters were derived from the NEFSC bottom trawl survey and translated into steepness using the ϕ_0 estimated from the base case model. The estimated steepness was 0.28 for both Ricker and Beverton-Holt SR and 0.68 for the survivorship SR from the base case model. Different SR assumptions resulted in different trajectories of spawning output and likely different management advice (Figure 4.40). These two SR sensitivity runs performed worse than the base case model in terms of AIC (Table 4.3). The recruitment likelihood increased when assuming a Ricker (recruitment likelihood = 126.99) or a Beverton-Holt (recruitment likelihood = 107.97) SR relationship, reflecting a poorer fit to the recruitment data compared to the base case model (recruitment likelihood = 0.24). The recruitment time series estimated from the Ricker and Beverton-Holt models were far from what was observed in the NEFSC spring bottom trawl survey (Figure 4.41; see McManus et al. 2022,

Figure 1). In both cases, the estimated NEFSC spring bottom trawl survey *q* was over 1, which indicated possible model misspecifications (Table 4.3).

Sensitivity runs were conducted with different time block assumptions:

- biology block 2011-2019,
- biology block 2013-2019,
- no biology block, and
- no survey block.

These were compared to the base case model where the biology block 2012-2019 and survey block 2009-2019 was assumed. For the runs with plus and minus one year of the base case biology block (2012-2019), the maturity and fecundity relationships remain the same as the base case model, and L_∞ was estimated for both sexes within the model. The run with no biology block, maturity, fecundity, and growth was assumed to be the same as the settings for the biology block 1989-2011 in the base case model. The model run with no biology block could not track the large population increases observed in surveys around 2010, and performed worse in terms of AIC (Table 4.3 and Figure 4.42; see TOR3). Assuming different lengths of the biology block only affected the earlier years' spawning output and did not change the terminal estimates (Figure 4.42). Therefore, even though the 2011-2019 biology block slightly outperformed the base case model, given that the terminal year estimates are insensitive to this assumption, the Working Group decided to proceed with the base case model configuration. The fit for length composition data was worse with no survey blocks in the model (Table 4.3).

A sensitivity run was conducted that examined a longer time series 1962-2019. The population is assumed to be unfished prior to 1962. Landings and discards from 1962 to 1988 were estimated using the same method used to derive the initial equilibrium catch for each fleet in the base case model. NEFSC spring bottom trawl survey time series data were available from 1979 for this run. The estimated spawning output is smaller for the 1962-2019 model; however, the trend is similar to the base case model (Figure 4.43).

Sensitivity runs were conducted using different survey data:

- NEFSC fall bottom trawl survey (as an additional abundance index),
- NEFSC spring longline survey (as an additional abundance index),
- NEFSC fall longline survey (as an additional abundance index), and
- VAST spring index (as the sole biomass index).

These were compared to the base case model that used only the NEFSC spring bottom trawl survey index. The estimated spawning output trend is similar to the base case model in all cases (Figure 4.44). The NEFSC fall bottom trawl survey was split into *Albatross* and *Bigelow* time series and entered as separate fleets in the model because their length composition is distinctly different (see TOR3). The estimated survey q for the NEFSC fall bottom trawl is much smaller than the spring survey (Table 4.3), reflecting the seasonal migration of dogfish out of the survey domain in the fall. The estimated selectivity for the NEFSC fall bottom trawl survey is logistic for the *Albatross* years but flat domed-shaped for the *Bigelow* period. Further investigations regarding the fall survey data and the model are required to examine whether this result is reasonable. Adding the NEFSC longline survey to the model did not change the spawning output (Figure 4.44). The model constructed using the model-based VAST index performed worse than the base case model in AIC (Table 4.3). The VAST length composition was estimated at a 6 cm length bin and was interpolated to a 3 cm length bin using a moving average method. It is not clear whether this mismatch is the cause of its low performance. The Working Group suggested continuing to develop the VAST index, and this index should be reevaluated in future assessments.

Profile and Retrospective Analysis

For the R_0 profile analysis, the $ln\ (R_0)$ parameter was fixed at values above and below the value estimated by the base case model (9 to 15 with an increment of 0.5, base case $ln\ (R_0)$ = 12) and the models were refitted. The results indicated that the length composition data was the most informative and the survey index was the least informative for estimating R_0 (Figure 4.45). Among the length composition data, the catch data support the base case R_0; however, the survey data slightly favored a smaller R_0 value (Figure 4.46). This result indicated a slight conflict between catch and survey length composition data and that the maximum likelihood estimate of R_0 landed at the spot where conflicts between different sources of data were balanced (Figure 4.46).

Likelihood profiling was conducted over a wide range of values for the female VB growth parameters L_∞ and k while the rest of the VB parameters were fixed at the Nammack et al. (1985) values. The model had a tendency to favor smaller L_∞ and slightly larger k values compared to Nammack et al. (1985; Figure 4.47). The run with the smallest total likelihood was L_∞ = 88 and k = 0.12, which is close to the maximum likelihood estimates (Tables 4.2-3 and Figure 4.47), suggesting that the estimated growth parameters in the base case model or sensitivity analysis are likely global instead of local minimums.

The survivorship SR parameters, Z_{frac}, β, and σ_R were profiled over a wide range of values, and the resulting total likelihoods are in Figure 4.48. Among the combination of parameters tested, the parameter values fixed in the base case model (Z_{frac} = 0.9, β = 1.5, and σ_R = 0.3) produced the smallest total likelihood. The β parameter is the least influential to the model, which is likely why this parameter is hard to estimate in SS3. The model performance is the most sensitive to Z_{frac}, where larger Z_{frac} values were favored.

A 7-year peel retrospective analysis was conducted for the base case model. The results indicated that the model has a minor retrospective pattern with Mohn's ρ= 0.06 for the spawning output and -0.05 for the fully recruited fishing mortality (Figures 4.49-4.50).

Stochastic Estimator

In addition to SS3, the Working Group used the Stochastic Estimator model to estimate the spiny dogfish population size and fishing mortality rates. The Stochastic Estimator uses swept area calculations based on the NEFSC spring bottom trawl survey and catch (landings and dead discard) data to estimate biomass and fishing mortality, under the assumption that survey efficiency (between the wingtips) is 1. It uses bootstrapping to better quantify the uncertainties of these quantities. It was the primary method used in recent previous assessments; a full description can be found in NEFSC (2006), pages 35-42.

Only minor changes to the Stochastic Estimator were done for this assessment. These include updating length-weight, maturity, and fecundity relationships, and changing the assumed logistic (landed) fishery selectivity curve to better match that from SS3 (in particular, the L_{50} for the selectivity curve was reduced from 80 to 73 cm). Spawning stock biomass (females greater than 80 cm) were replaced by spawning output (pups); these quantities are strongly correlated. Additionally, a call to a proprietary subroutine that calculates normal quantiles was replaced by public code, so the Fortran source code can be compiled using the 'gfortran' open source compiler.

The Stochastic Estimator was run for the 1989-2019 time series using spring trawl survey, landings, and discard data. Results show high fishing mortality on females, a decline in total and exploitable female biomass and spawning output during 1989-2000, a recovery after fishing mortality was reduced during 2000-2010, and a more modest decline in biomass and spawning output in the last years of the time series as fishing mortality increased somewhat (Figure 4.51). Fishing mortality for males has remained low.

The results from the Stochastic Estimator can be compared to those from the SS3 base run (Figure 4.52). These estimates are strongly correlated during 2000-2019, with SS3 estimating somewhat higher biomass and lower fishing mortality due to its lower survey efficiency estimate (*q*). In the early portion of the time series (1989-2009), the Stochastic Estimator shows greater declines in spawning output and much higher fishing mortalities than SS3. It is likely that this is due to some misspecification in life history parameters (e.g., growth, natural mortality) or in catch data (e.g., discards, discard mortality) in SS3 during that period.

Table 4.1. Summary of Atlantic spiny dogfish data by gear and fleet used in SS3.

Type	Gear	Fleet	Label
Landings	Sink Gill Net + Others Recreational	1	Landings_SGN_Rec_Others
Landings	Longline Otter Trawl + Foreign	2	Landings_LL_OT_Foreign
Discard	Sink Gill Net Scallop Dredge	3	Discard_SGN_SD
Discard	Longline Large Mesh Otter Trawl Recreational	4	Discard_LMOT_LL_Rec
Discard	Small Mesh Otter Trawl	5	Discard_SMOT
Survey	NEFSC Spring Bottom Trawl	6	NEFSC_Spring_BTS

Table 4.2. Summary of von Bertalanffy (VB) growth parameters assumed/estimated in SS3 for Atlantic spiny dogfish. Values with an asterick indicate an estimated value.

Sex	VB Parameters	Base Case 1989-2011 Nammack et al. (1985)	Base Case 2012-2019 Est. $L\infty$	Sensitivity 2012-2019 Est. $L\infty$, k	Sensitivity 2012-2019 Est. $L\infty$, k, L_{Amin}
F	$L\infty$	100.50	89.24*	88.64*	88.67*
F	k	0.1057	0.1057	0.1258*	0.1259*
F	L_{Amin}	26.53	26.53	26.53	27.33*
M	$L\infty$	82.49	79.14*	78.02*	78.02*
M	k	0.1481	0.1481	0.1657*	0.1666*
M	L_{Amin}	26.94	26.94	26.53	27.46*

Table 4.3. Summary of Atlantic spiny dogfish SS3 model runs.

Version	Sensitivity Category	Scenario	AIC	Delta AIC	Catchability (*q*)	Steepness (*h*)
3.6.2 1.5	Base Case	Dirichlet-Multinomial Data Weighting	-46624	-	0.83	0.68
3.6.2 1	Base Case	No Data Weighting	5504	0	0.88	0.68
3.6.2_2	Growth	Nammack et al. (1985)/Est L_∞, and *k*	5488	-17	0.85	0.68
3.6.2_3	Growth	Nammack et al. (1985)/Est L_∞, *k*, and L_{Amin}	5485	-19	0.85	0.68
3.6.2_4	Growth	Nammack et al. (1985)	5931	427	1.03	0.68
3.6.2 8.1	Natural Mortality	*M*=0.092 (Hoenig 1983)	5108	-396	1.11	1.23
3.6.2 8	Natural Mortality	*M*=0.102 (Then et al. 2015)	5059	-446	1.13	1.01
3.6.2 8.2	Natural Mortality	Lorenzen (1996) scaled asymptote 0.102	5938	433	0.47	0.36
3.6.2 6	SR Relationship	Ricker SR with recruitment deviation	5833	328	1.21	0.28
3.6.2 5	SR Relationship	Beverton-Holt SR with recruitment deviation	5804	300	1.18	0.28
3.6.2 10	Time Block	Biology Block 2011-2019	5387	-117	0.86	0.68
3.6.2 11	Time Block	Biology Block 2013-2019	5601	69	0.89	0.68
3.6.2 1.2	Time Block	No Biology Block	5938	434	1.02	0.68
3.6.2 9	Time Block	No Survey Block	5648	143	0.95	0.68
3.6.2 13.1	Model Starting Year	1962-2019 Model	6974	-	0.87	0.68
3.6.2 14	Survey Data	Additional NEFSC fall bottom trawl survey	7202	-	0.94/0.33/0.48	0.68
3.6.2 15	Survey Data	Additional NEFSC spring longline survey	5606	-	0.89/0.0004	0.68
3.6.2 16	Survey Data	Additional NEFSC fall longline survey	5590	-	0.89/0.0002	0.68
3.6.2 18	Survey Data	VAST spring index	5778	274	0.03	0.68

Table 4.4. Summary of total biomass by sex, spawning output, recruitment (in 1,000, age 0+) and fishing mortality (age 12+) by year estimated by SS3 for Atlantic spiny dogfish.

Year	Female Total Biomass	Male Total Biomass	Spawning Output	Recruitment (1,000s)	F
1989	379,672	432,328	228,469	218,249	0.076
1990	386,663	437,351	232,245	223,706	0.118
1991	382,068	440,461	221,779	213,925	0.087
1992	384,717	447807	217,034	209,429	0.17
1993	373,117	447,218	199,000	192,048	0.107
1994	371,731	453,841	187,884	181,317	0.084
1995	376,160	461,839	183,010	176,608	0.109
1996	375,467	466,877	174,570	168,454	0.101
1997	373,842	472,231	165,600	159,660	0.068
1998	380,404	478,322	167,817	156,426	0.079
1999	381,356	480,471	169,694	102,990	0.067
2000	384,201	480,566	178,975	99,774	0.044
2001	389,329	478,825	196,331	73,343	0.031
2002	395,526	474,807	219,984	76,663	0.029
2003	398,997	468,448	244,437	74,109	0.017
2004	403,791	461,401	271,988	87,065	0.02
2005	405,289	452,780	296,758	85,641	0.016
2006	406,741	444,746	319,904	115,680	0.02
2007	406,047	436,859	338,467	122,918	0.024
2008	404,749	431,073	351,125	176,522	0.019

2009	406,500	429,058	360,845	196,595	0.023
2010	410,016	430,333	364,526	234,935	0.017
2011	418,240	435,756	365,877	235,805	0.026
2012	425,115	444,996	388,326	288,488	0.029
2013	409,991	443,991	353,179	120,648	0.027
2014	401,195	445,024	325,491	167,354	0.041
2015	389,002	444,033	296,337	123,237	0.028
2016	383,112	444,474	276,850	137,889	0.039
2017	375,398	444,646	256,708	159,111	0.032
2018	371,603	444,323	245,197	136,947	0.026
2019	371,635	445,385	239,877	176,963	0.032

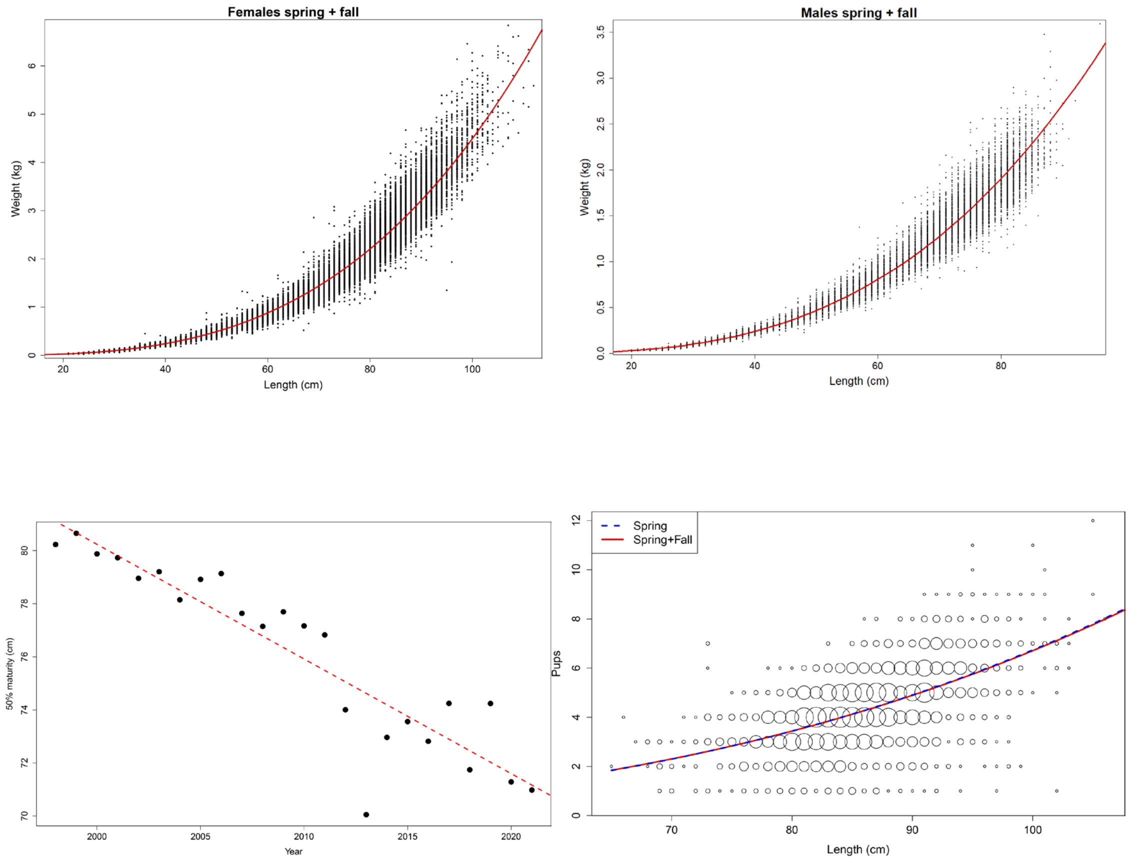


Figure 4.1. Plots of female and male length-weight relationships (top left and right), 50% (L_{50}) maturity at length over time (bottom left), and pups at length (bottom right).

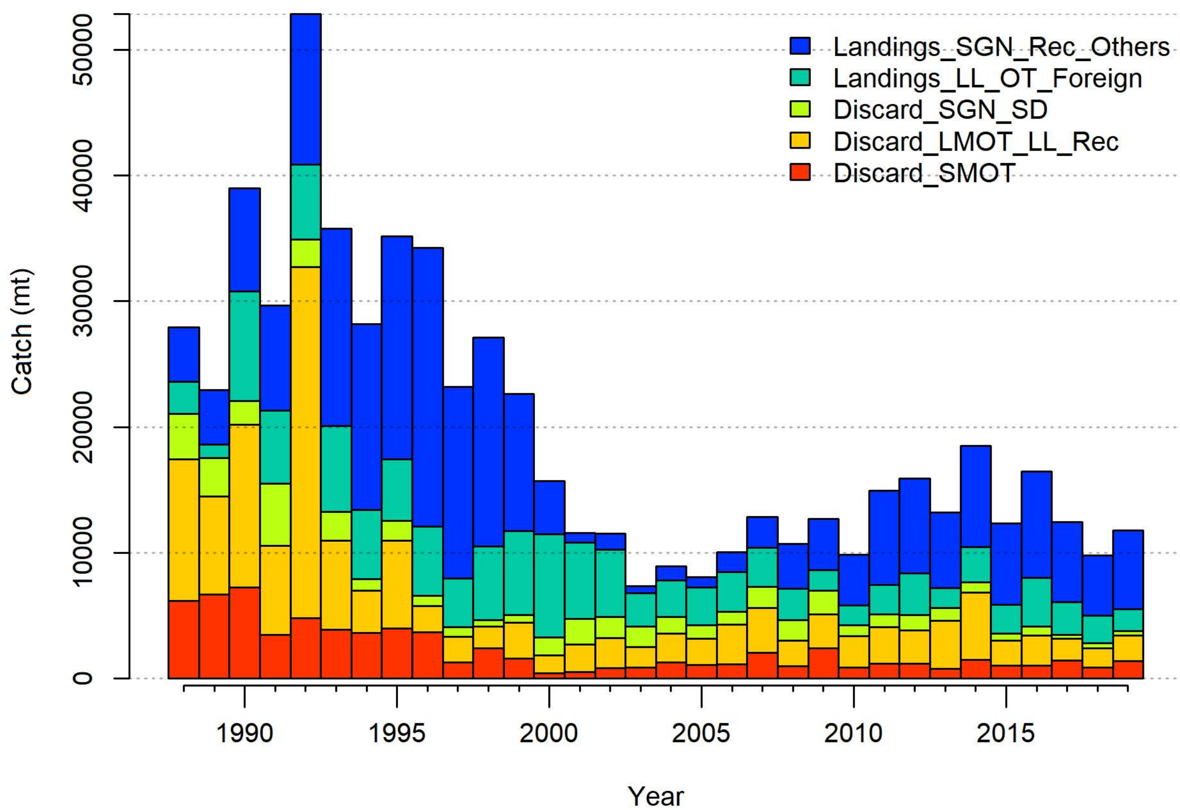


Figure 4.2. Time series of Atlantic spiny dogfish catch by fleet.

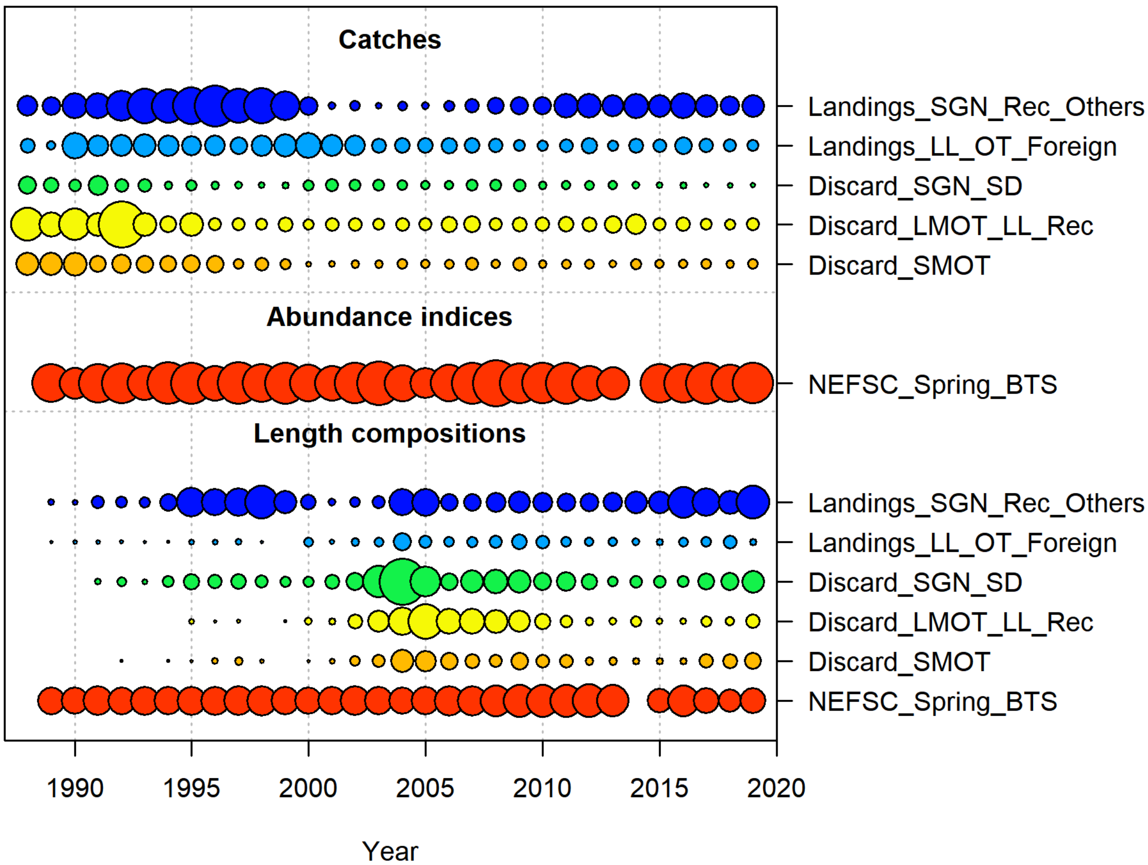


Figure 4.3. Catch and survey data by year for each fleet used in SS3. Circle area is relative within a data type. Circles are proportional to total catch for catches, to precision for indices, and to total sample size for length compositions. Note that since the circles are scaled relative to the maximum within each type, the scaling within separate plots should not be compared.

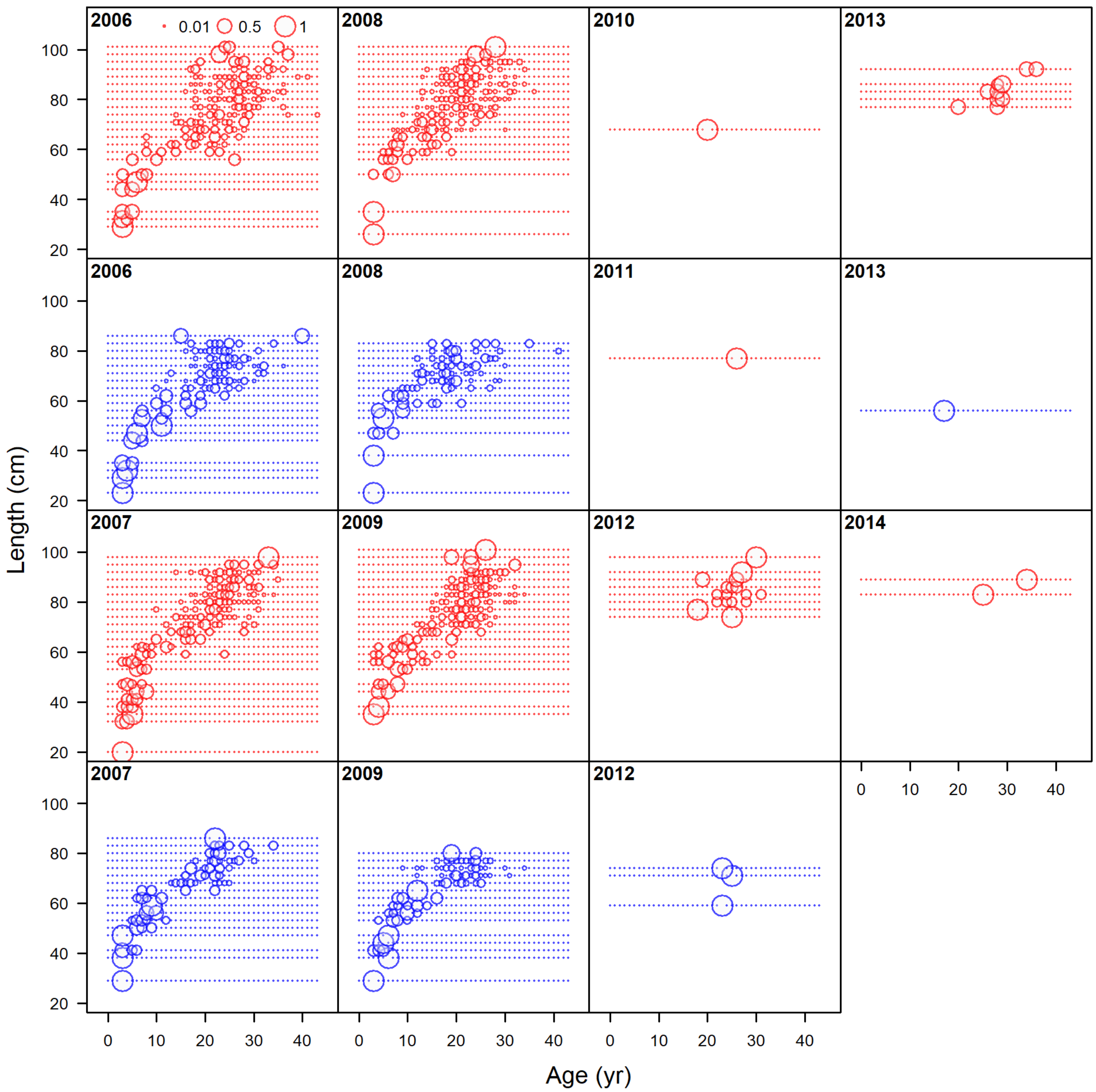


Figure 4.4. Conditional age-at-length data from NEFSC spring bottom trawl survey.

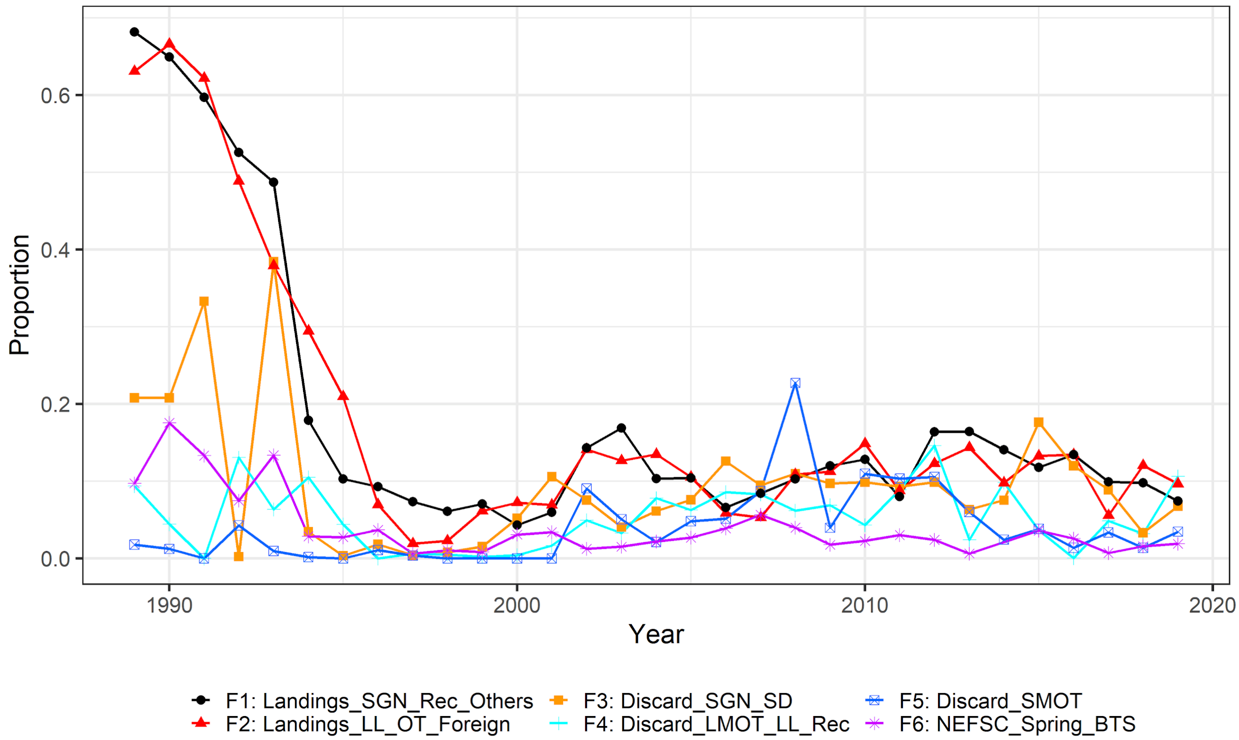


Figure 4.5. Proportion of 90+ cm females by fleet and year.

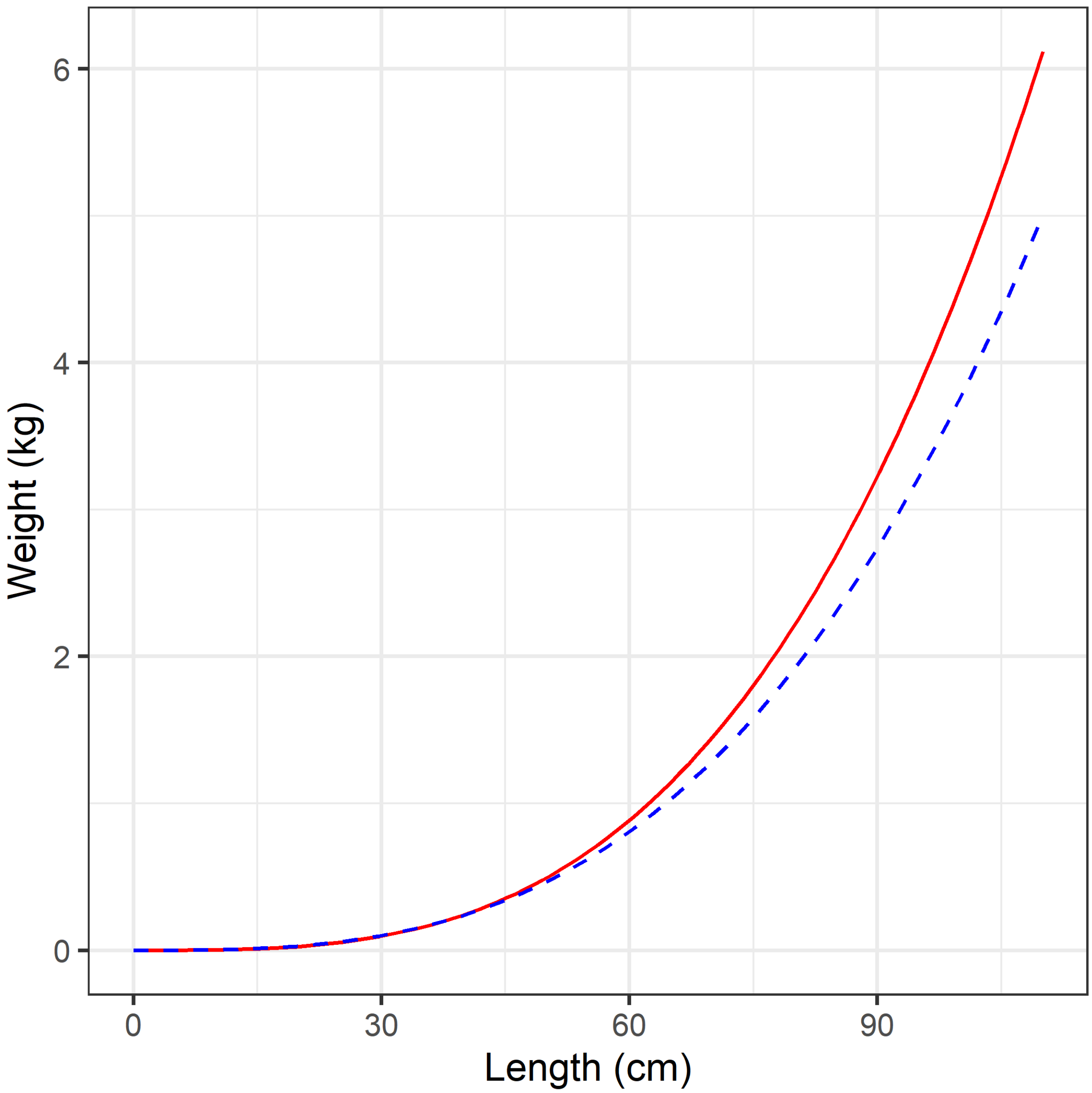


Figure 4.6. Length-weight relationships for females (red solid line) and males (blue dash line).

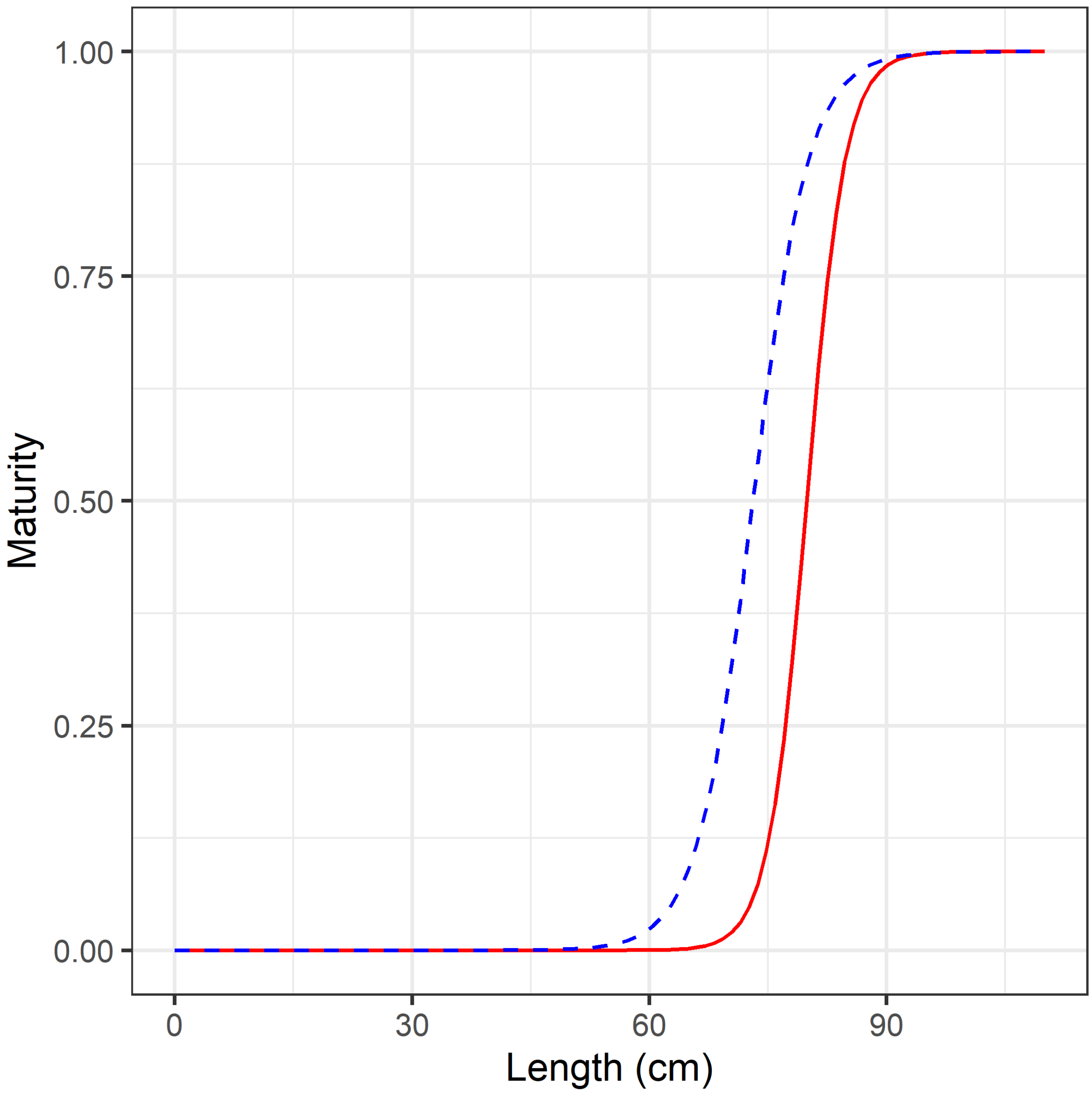


Figure 4.7. Maturity at length for biology blocks 1989-2011 (red solid line) and 2012-2019 (blue dash line).

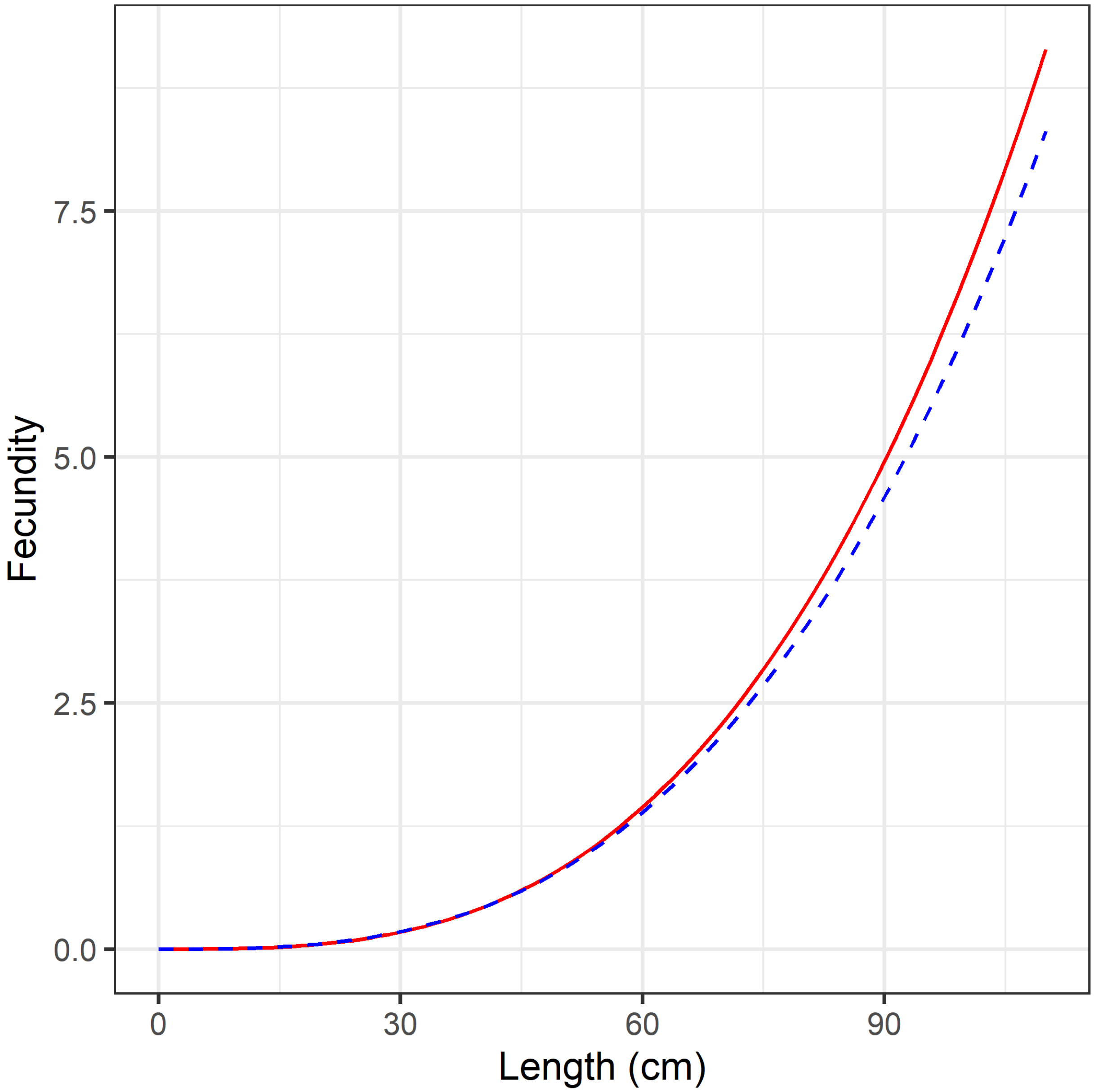


Figure 4.8. Fecundity at length for biology blocks 1989-2011 (red solid line) and 2012-2019 (blue dash line).

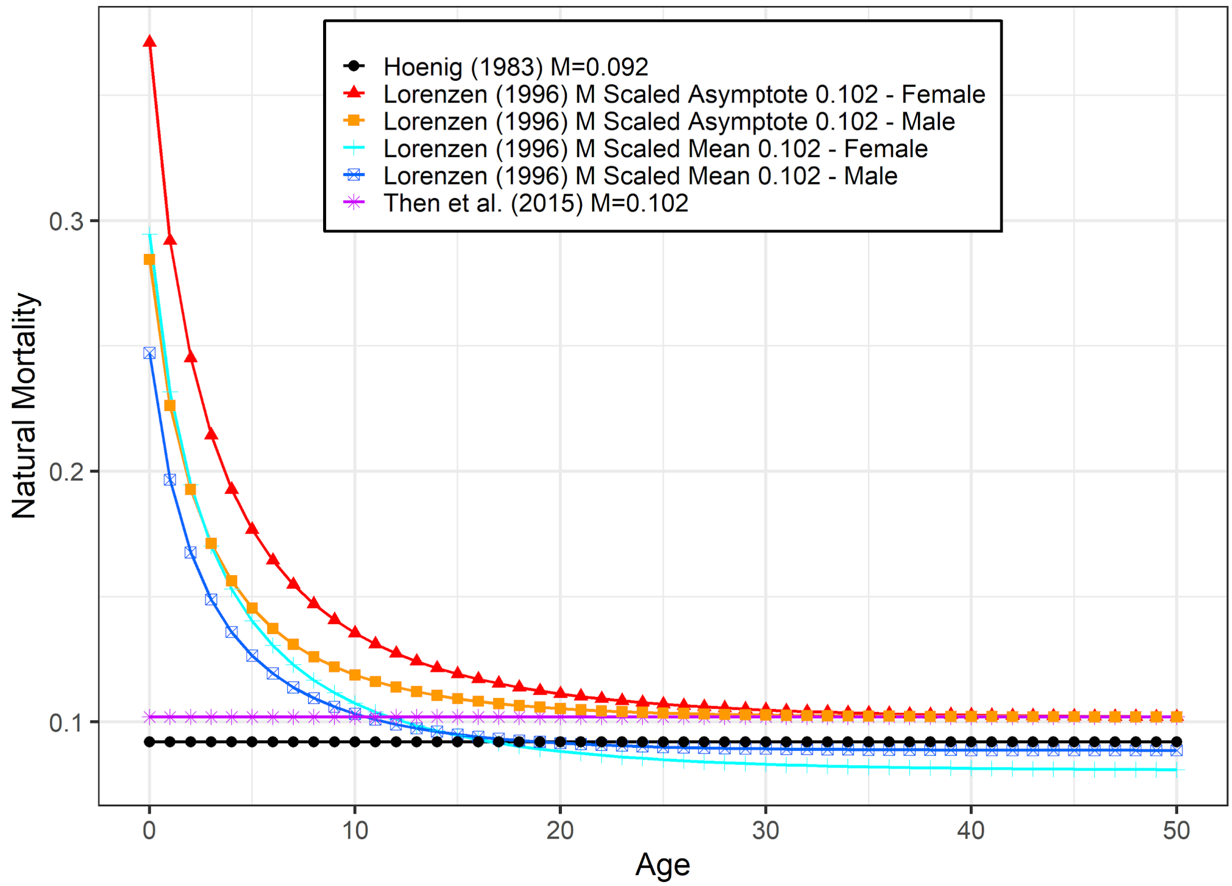


Figure 4.9. Natural mortality estimates explored in SS3 for Atlantic spiny dogfish.

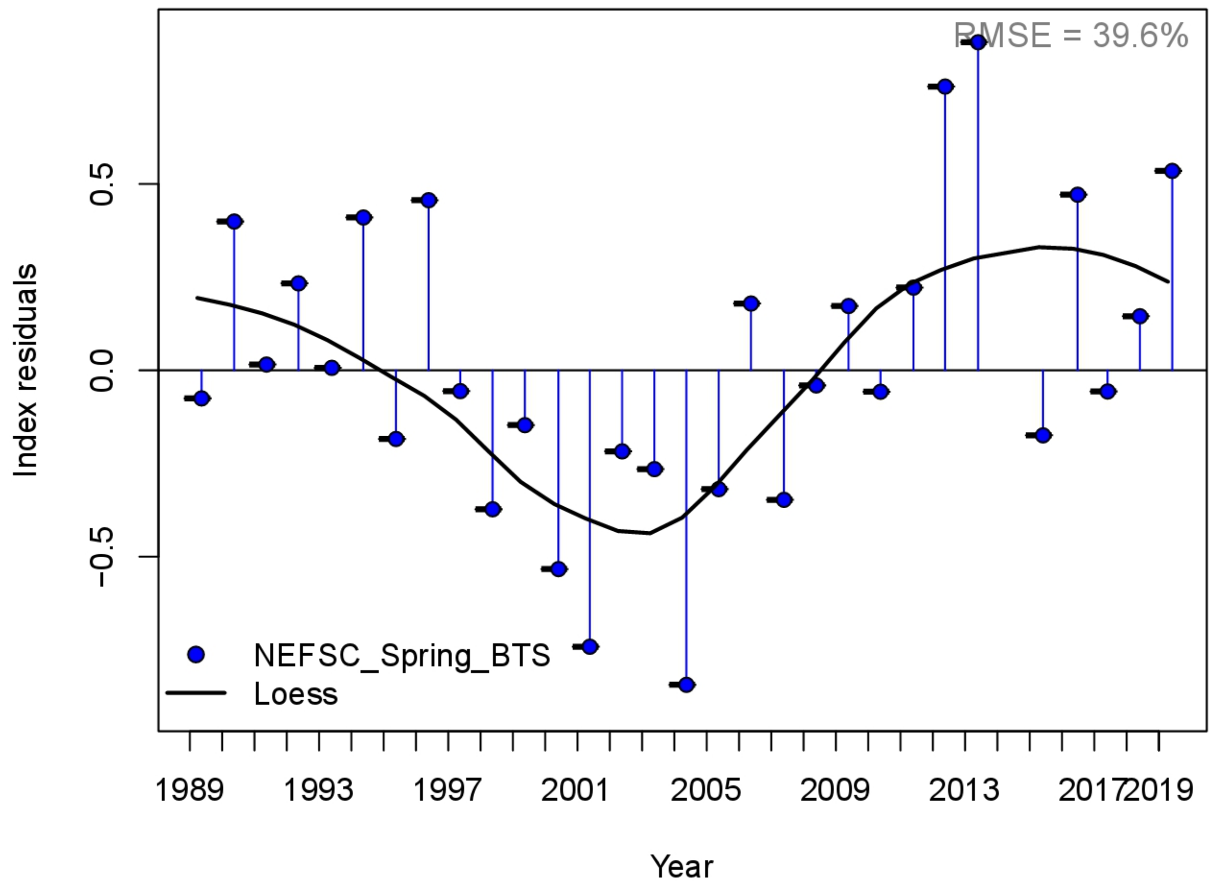


Figure 4.10. Joint residual plot from fit to annual index data.

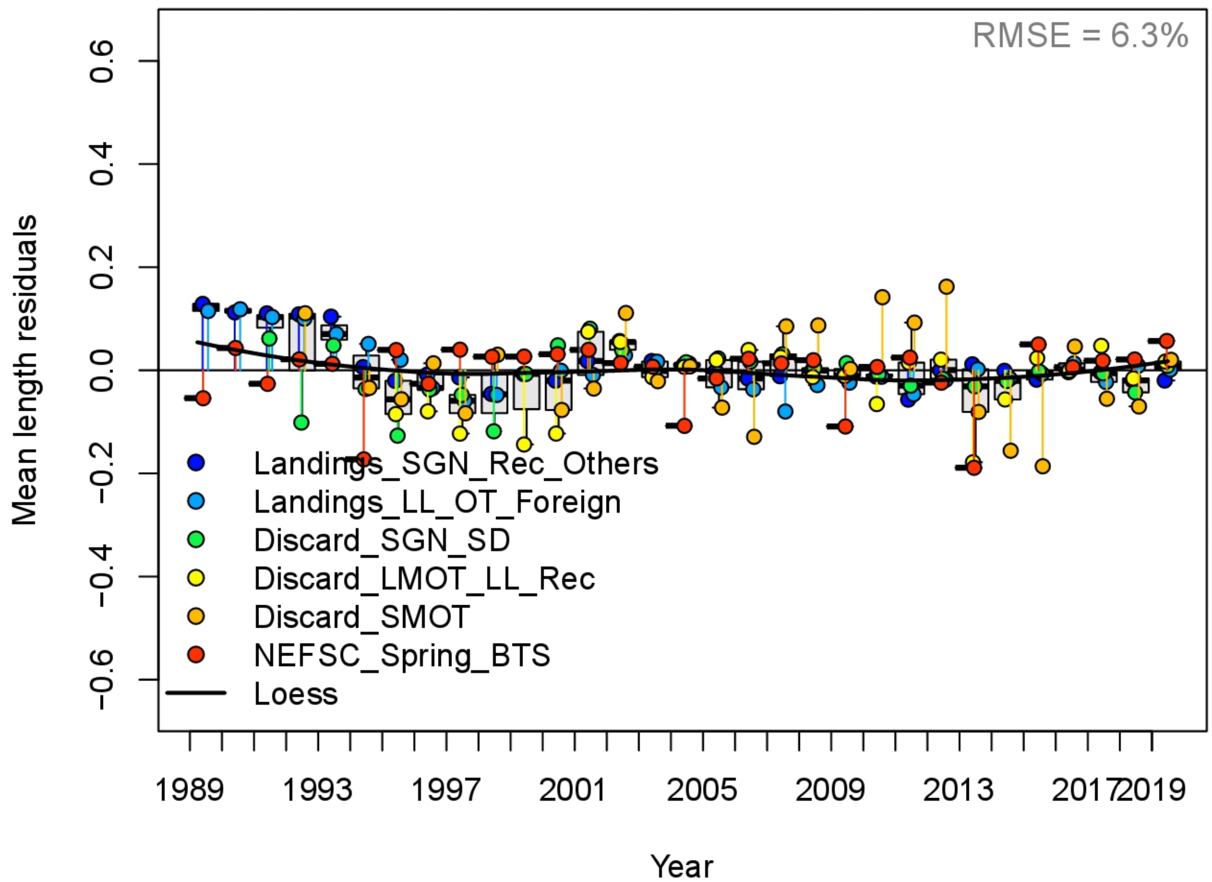
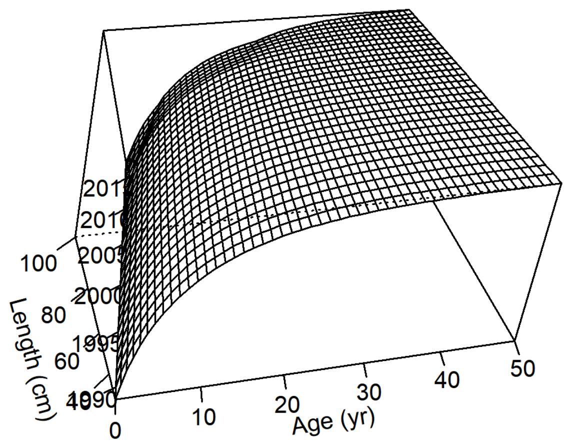


Figure 4.11. Joint residual plot from fit to annual mean length from length composition data.



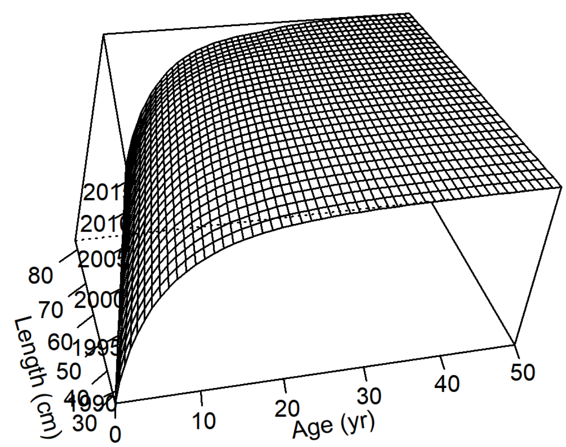


Figure 4.12. Surface plot of time-varying growth for females (top) and males (bottom) from 1989 to 2019.

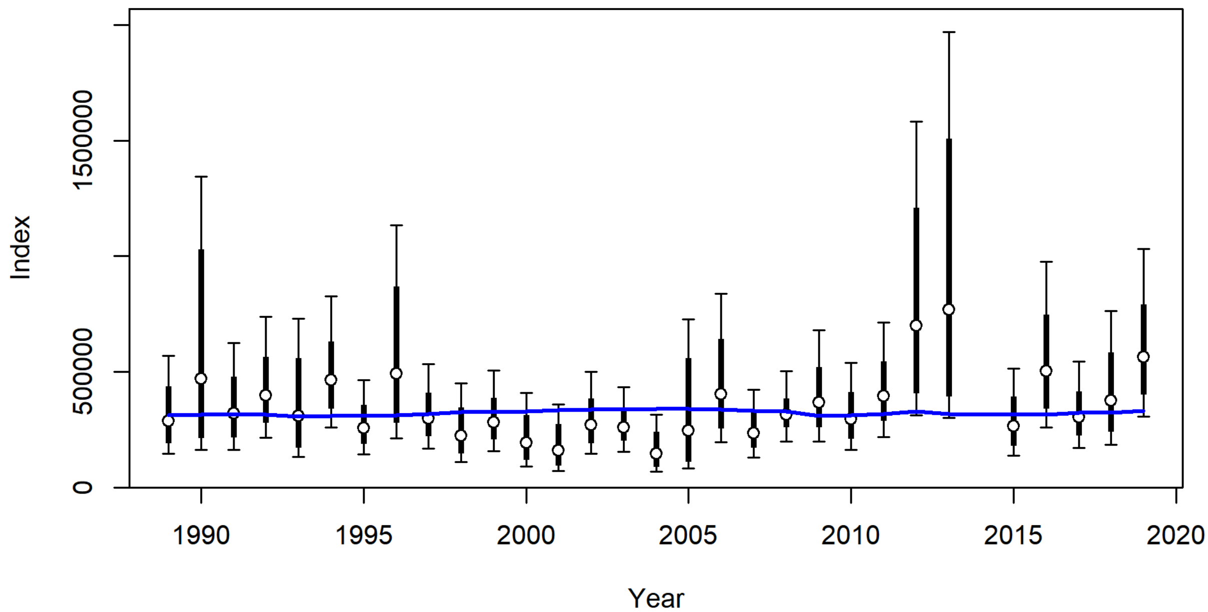


Figure 4.13. Observed and model-predicted abundance index (1,000s) for the NEFSC spring bottom trawl survey. Lines indicate 95% uncertainty interval around index values based on the model assumption of lognormal error. Thicker lines indicate input uncertainty before addition of estimated additional uncertainty parameter.

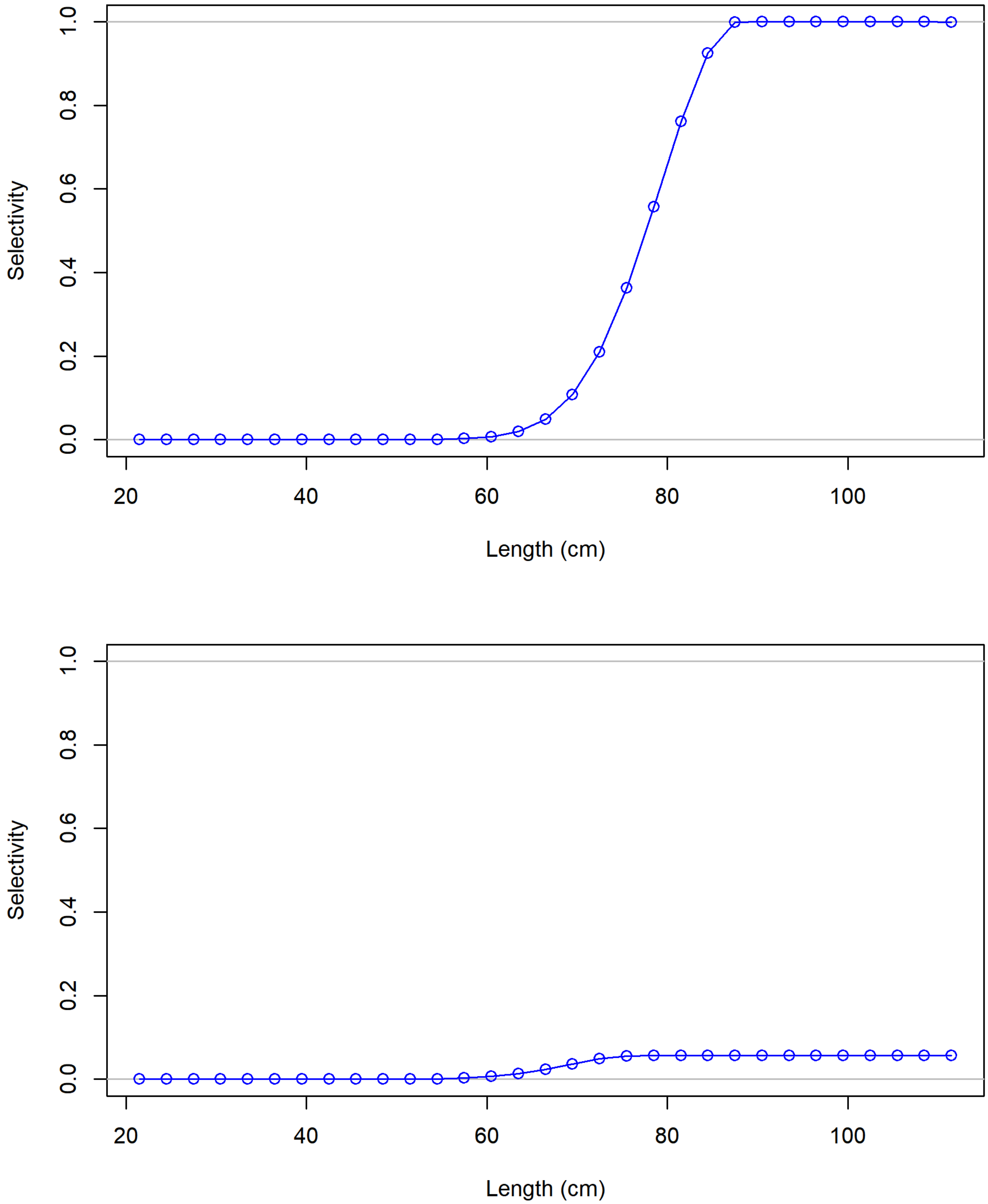


Figure 4.14. Estimated selectivity for females (top) and males (bottom) for fleet 1: Landings_SGN_Rec_Others.

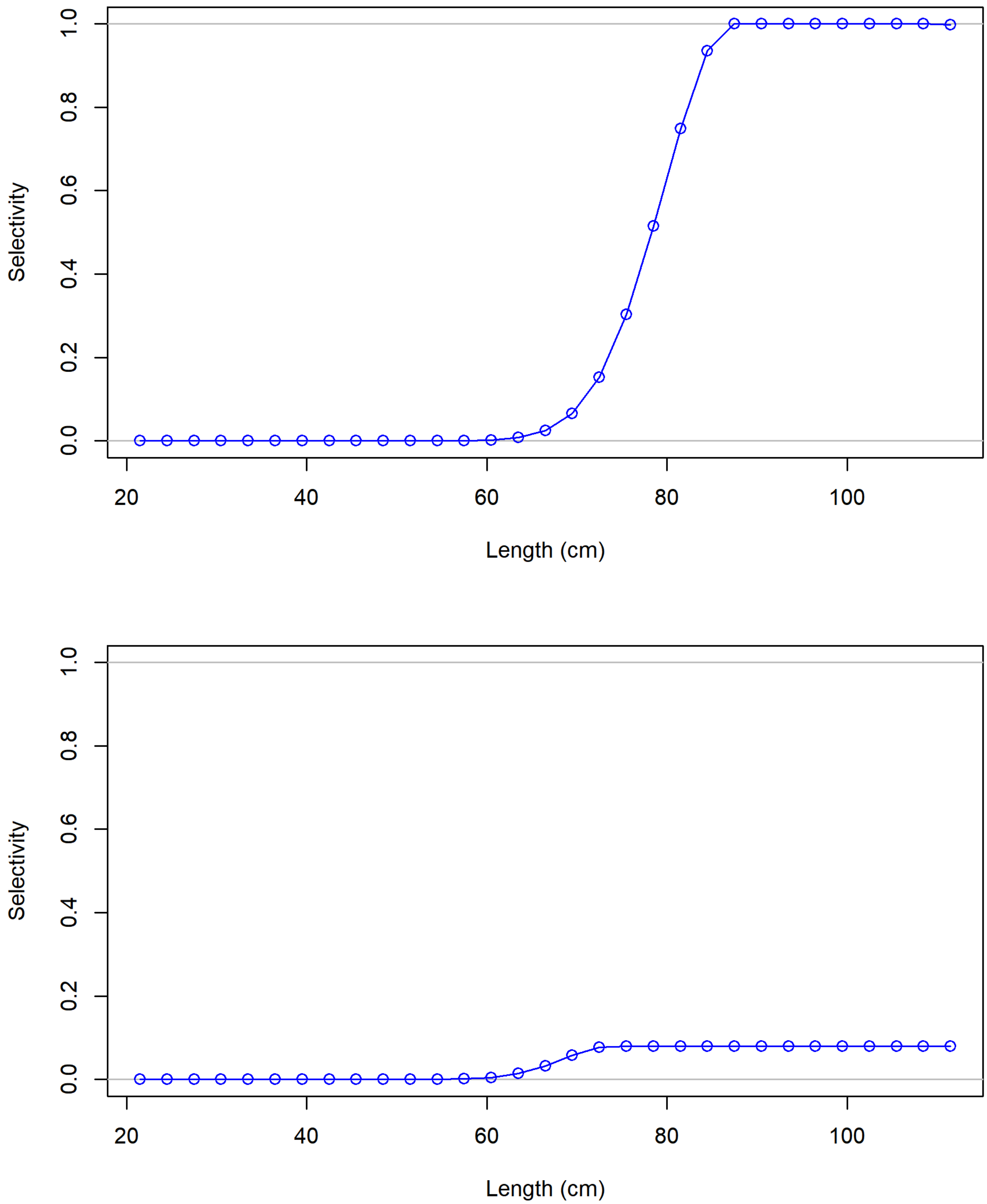


Figure 4.15. Estimated selectivity for females (top) and males (bottom) for fleet 2: Landings_LL_OT_Foreign.

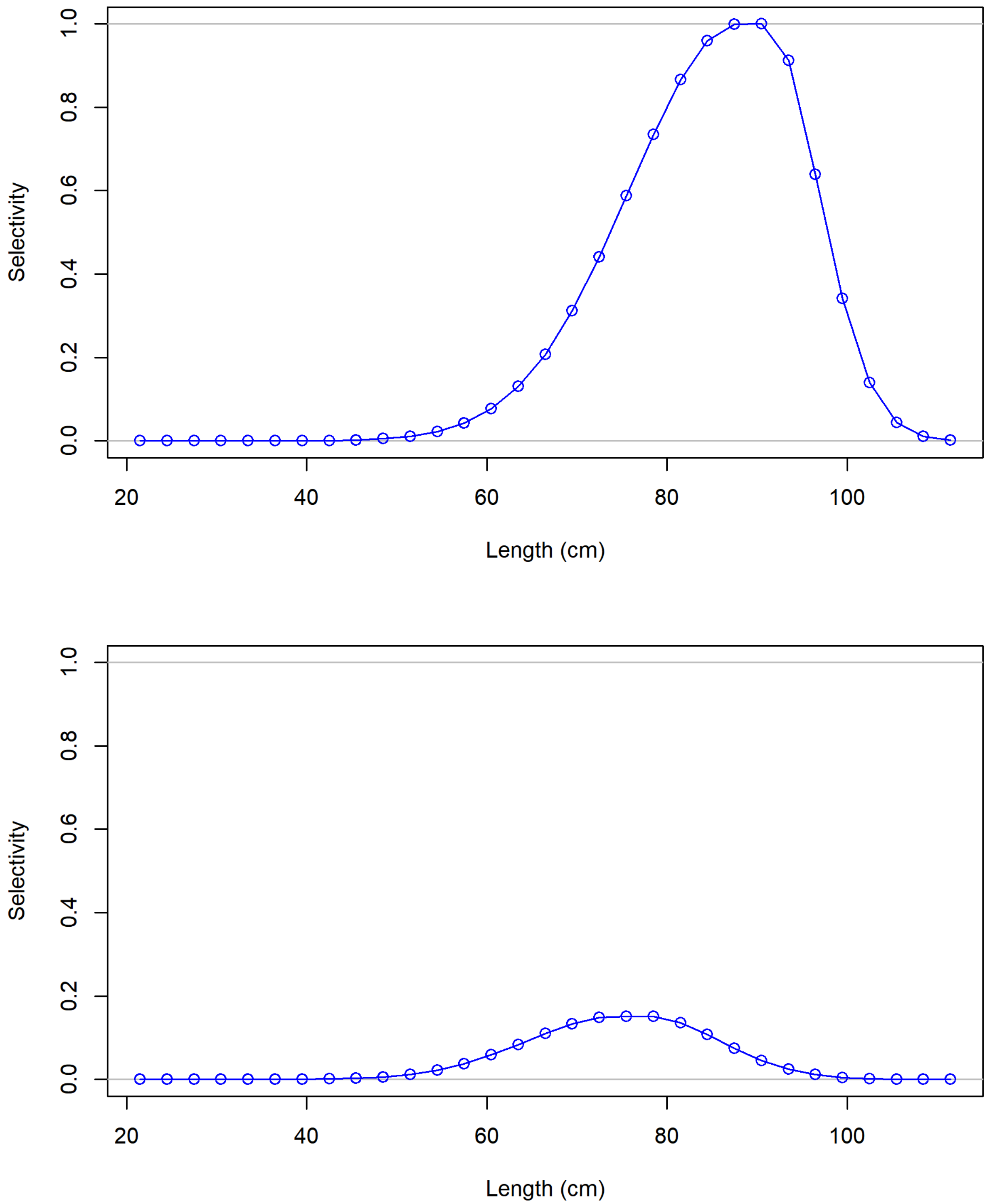


Figure 4.16. Estimated selectivity for females (top) and males (bottom) for fleet 3: Discard_SGN_SD.

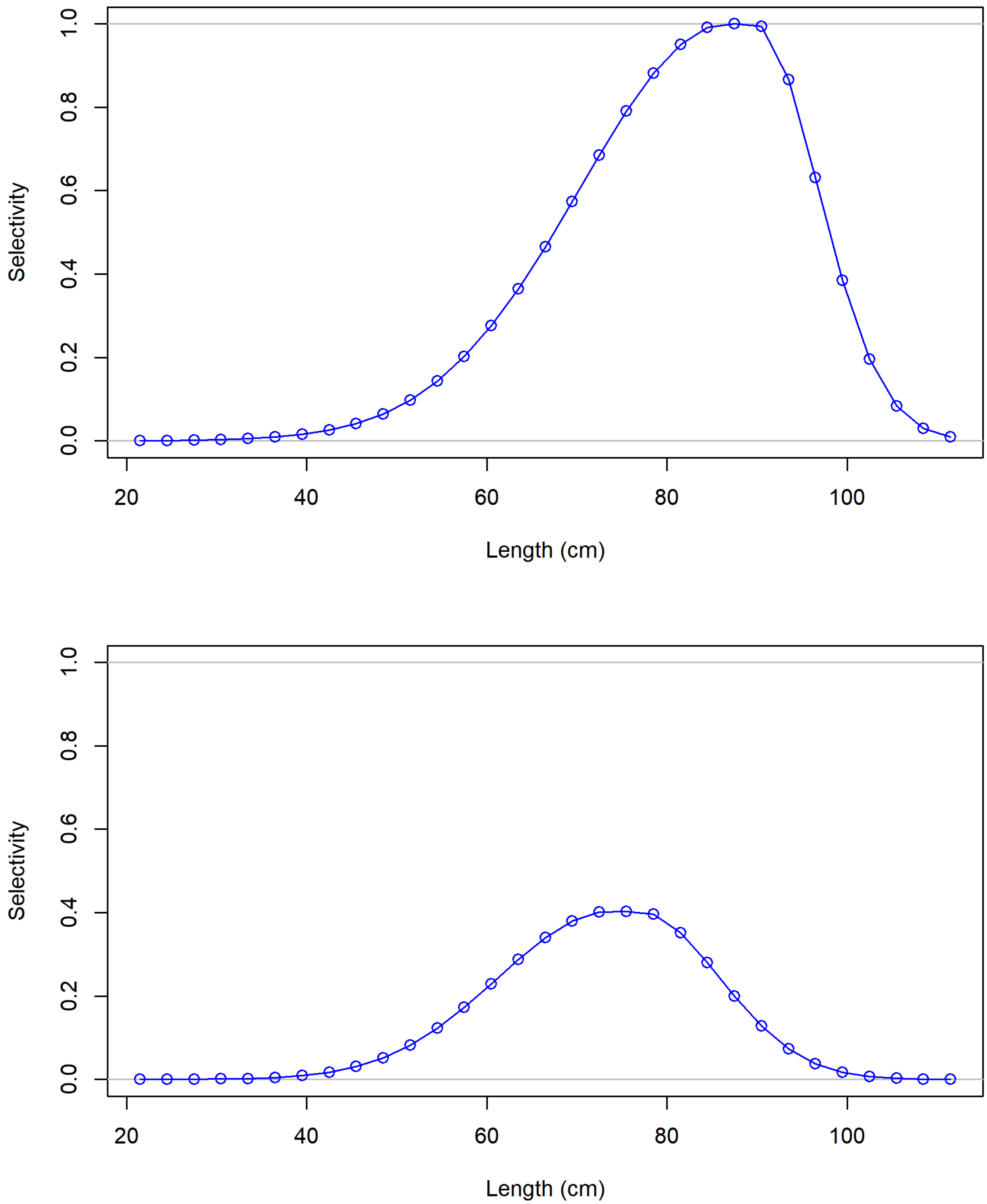


Figure 4.17. Estimated selectivity for females (top) and males (bottom) for fleet 4: Discard_LMOT_LL_Rec.

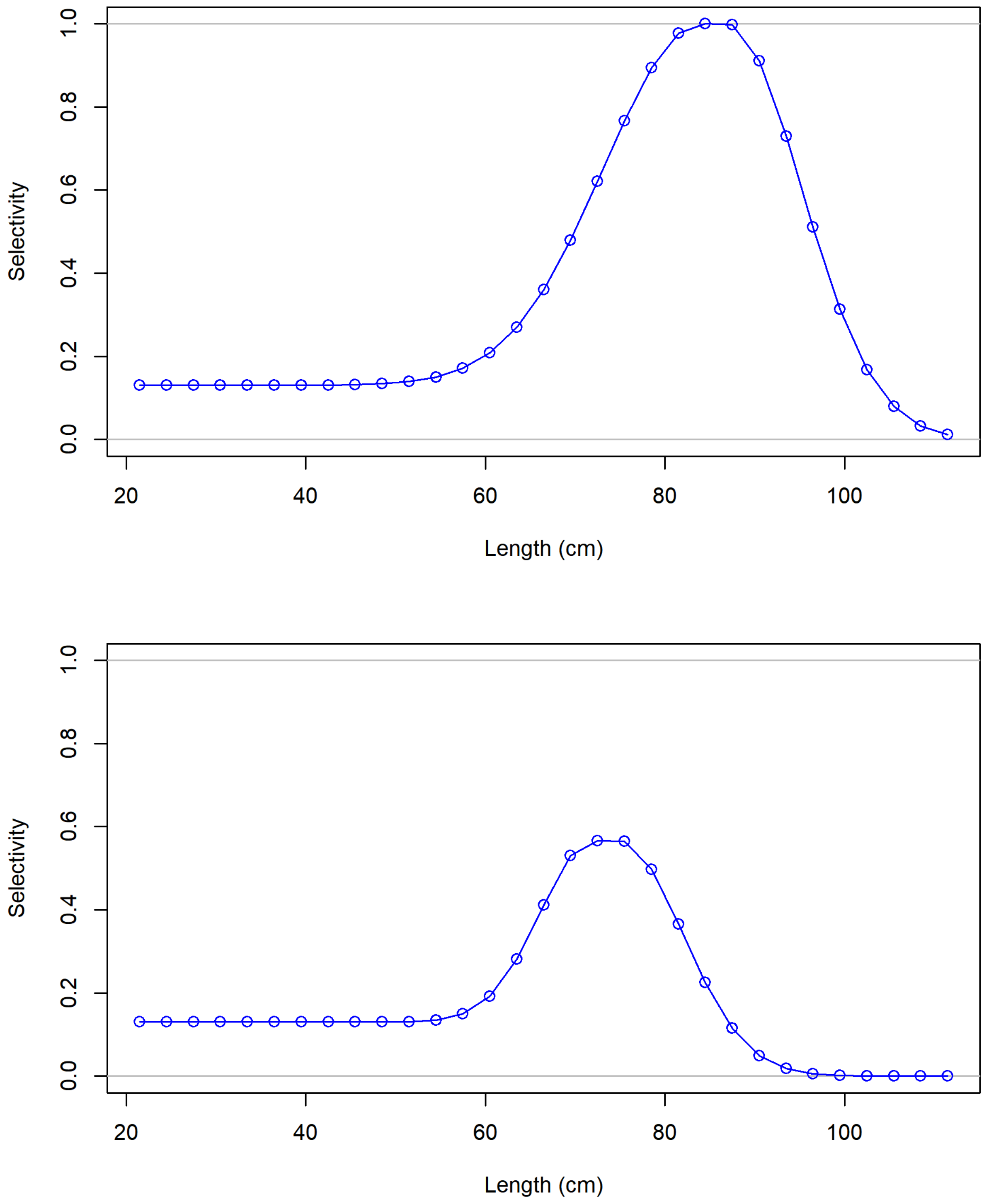
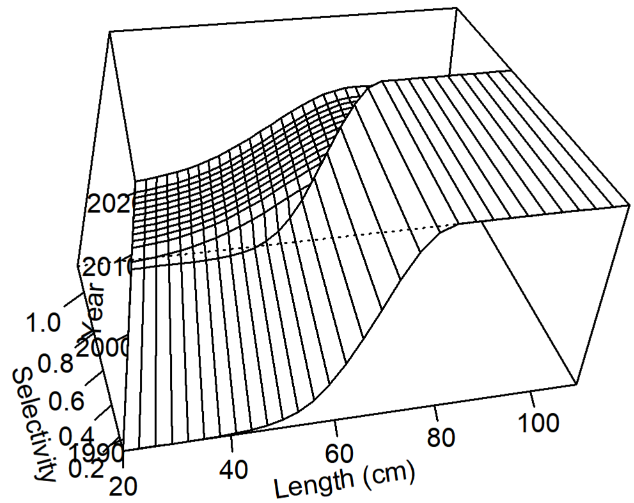


Figure 4.18. Estimated selectivity for females (top) and males (bottom) for fleet 5: Discard_SMOT.



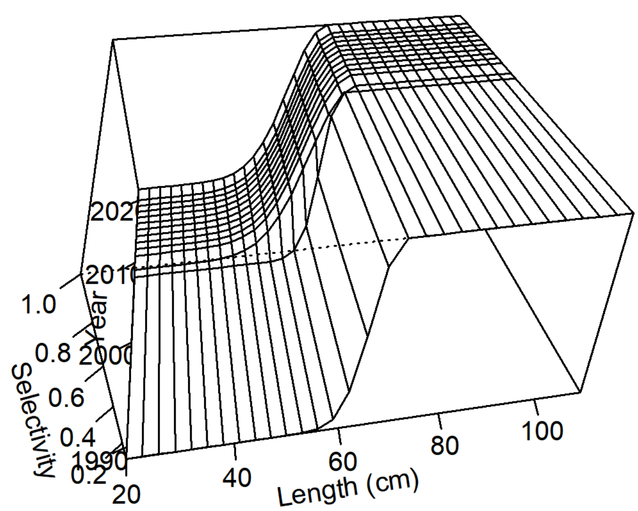


Figure 4.19. Surface plot of time-varying selectivity for females (top) and males (bottom) from 1989 to 2019 for NEFSC spring bottom trawl survey.

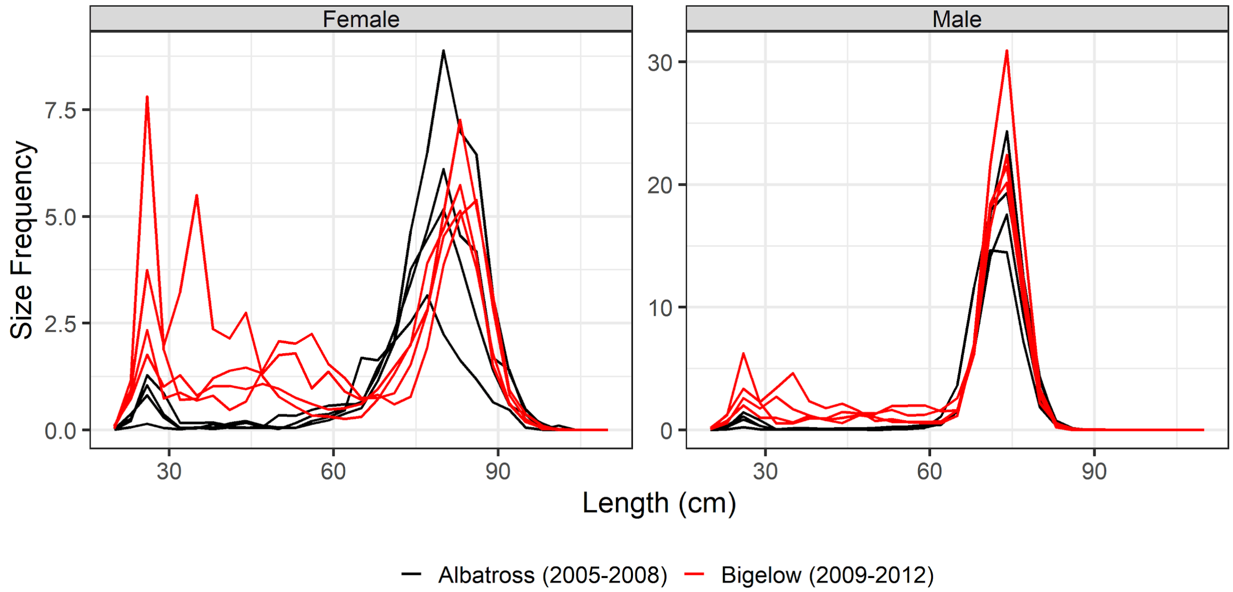


Figure 4.20. Observed length composition data from 2005 to 2012 for NEFSC spring bottom trawl survey by *Albatross* and *Bigelow* period.

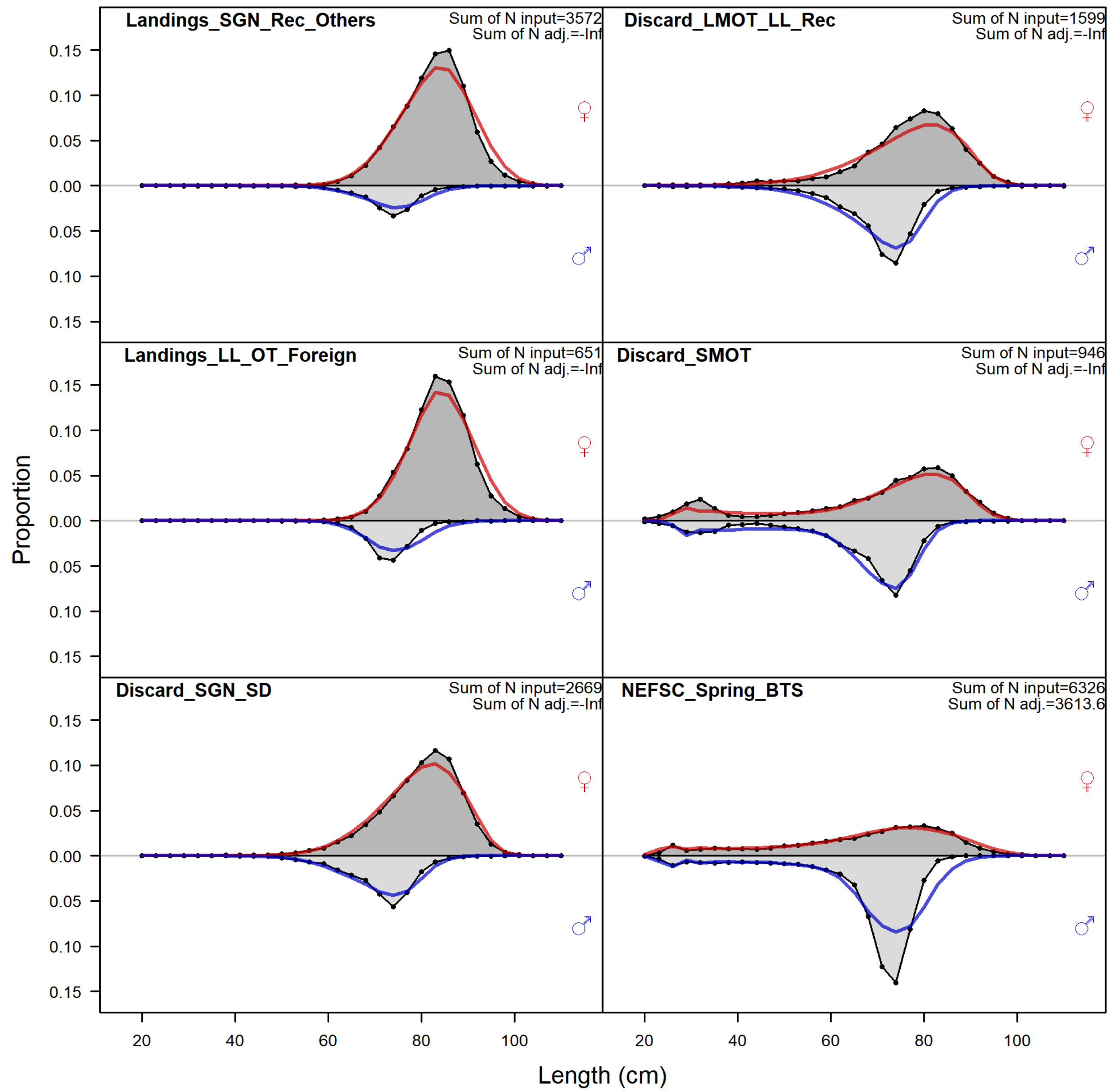
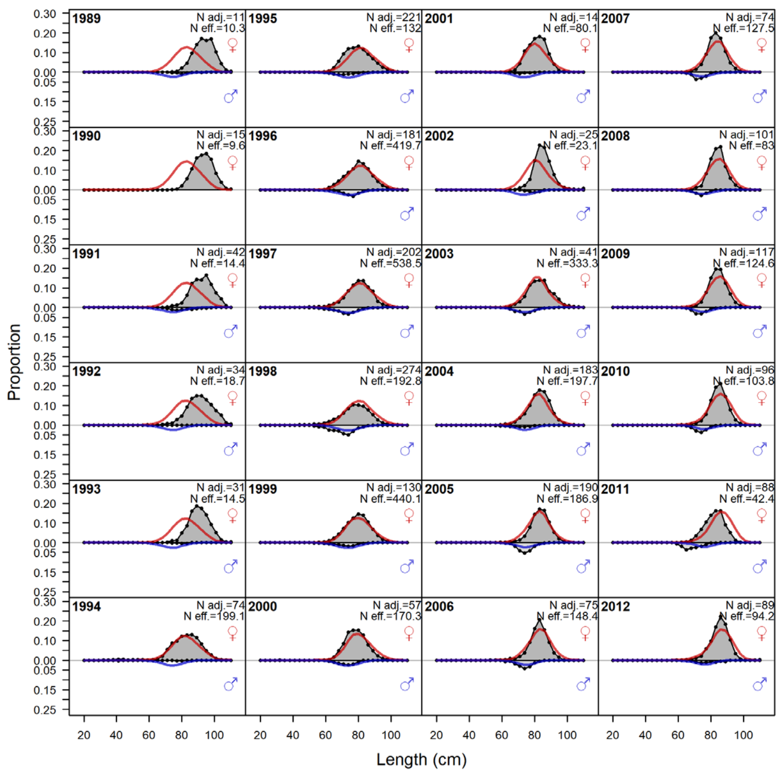


Figure 4.21. Observed (shaded) and model-predicted (line) length compositions, aggregated across time by fleet and sex.



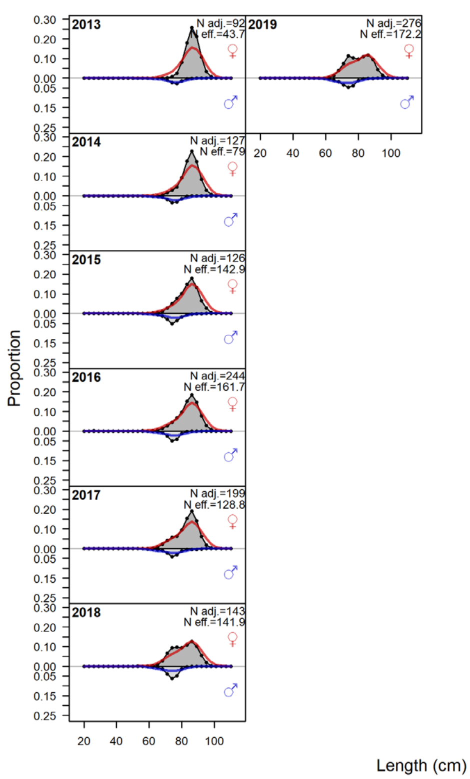


Figure 4.22. Fit to length compositions by year and sex for fleet 1: Landings_SGN_Rec_Others.

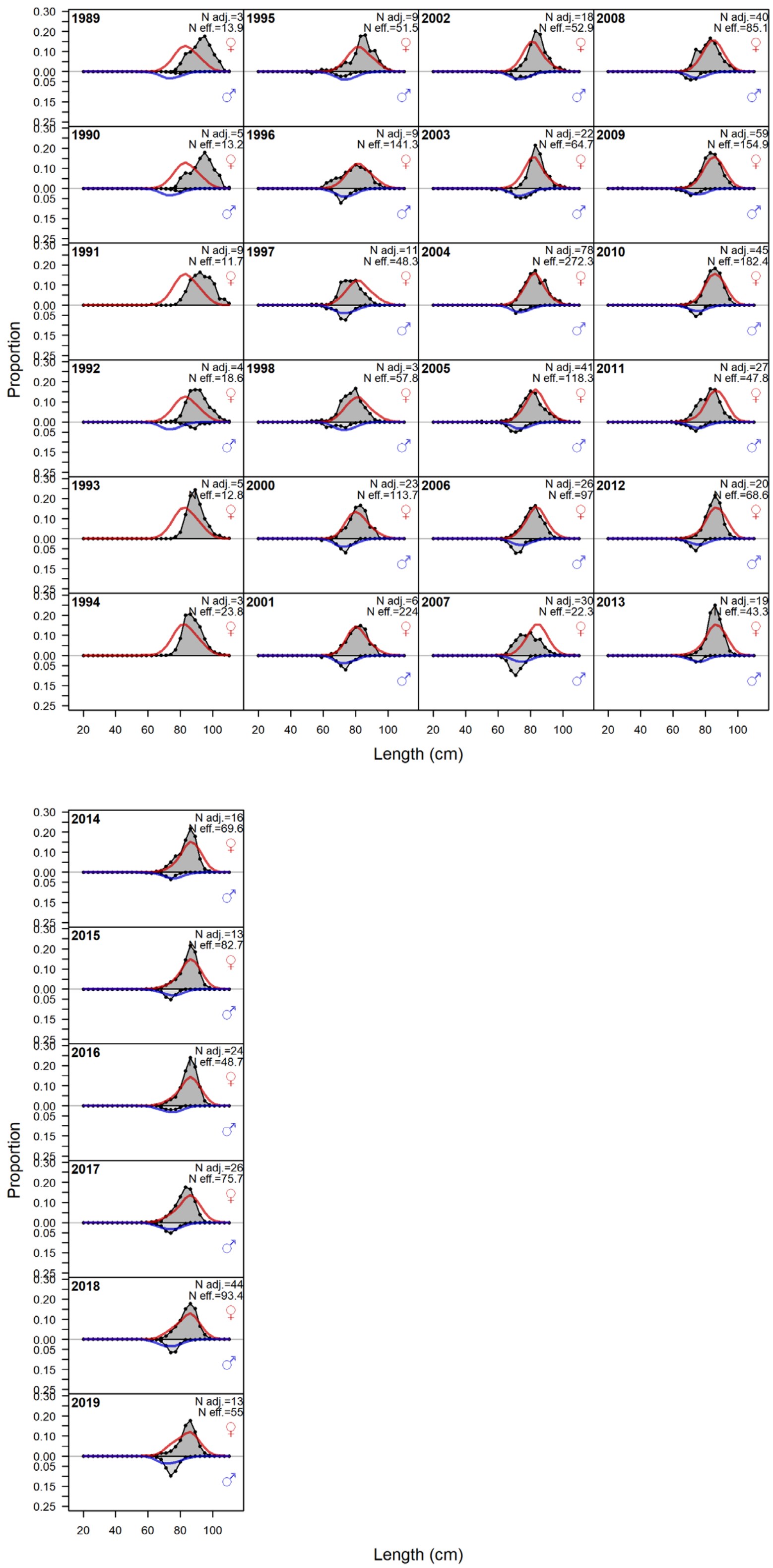


Figure 4.23. Fit to length compositions by year and sex for fleet 2: Landings_LL_OT_Foreign.

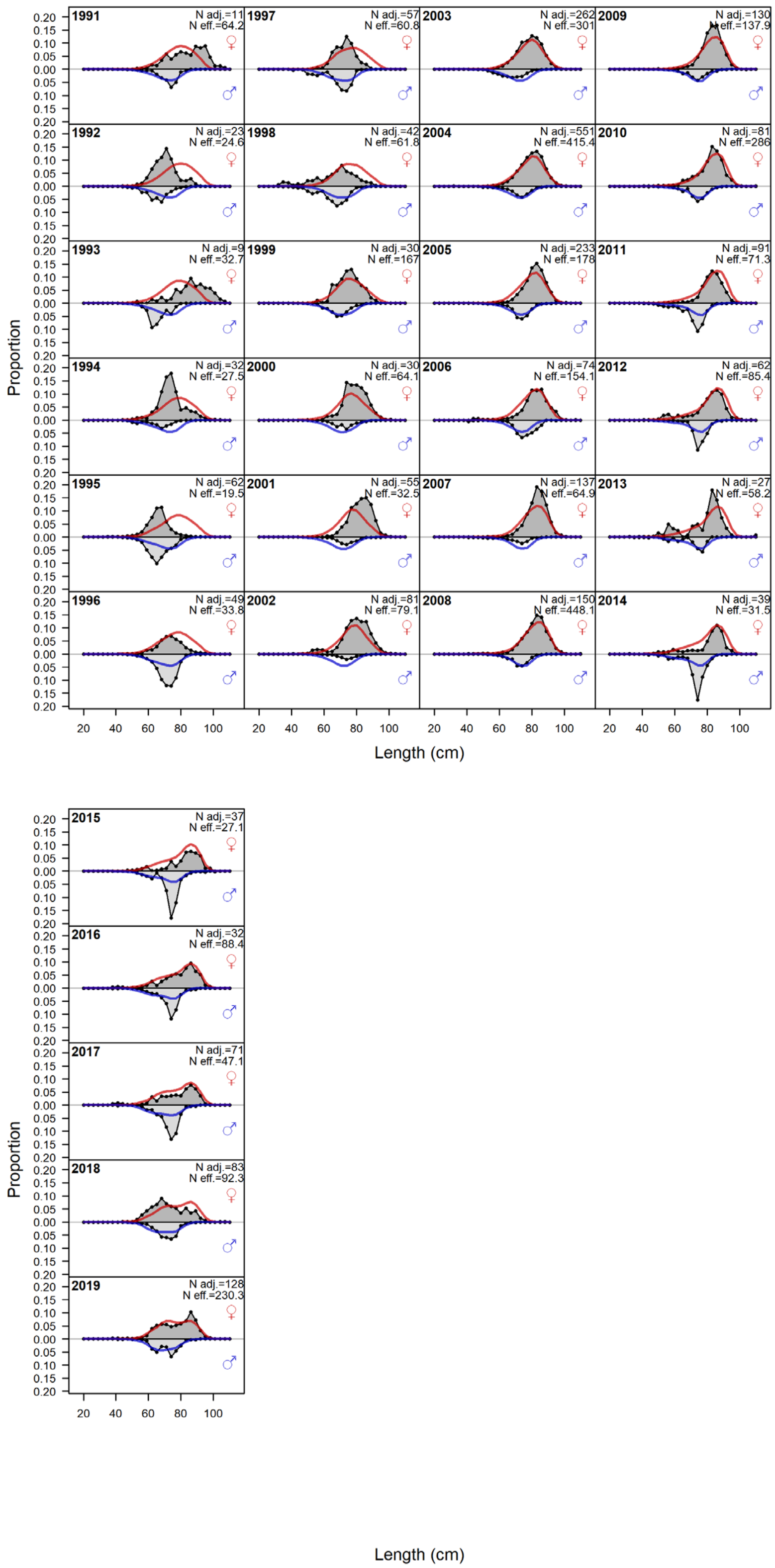


Figure 4.24. Fit to length compositions by year and sex for fleet 3: Discard_SGN_SD.

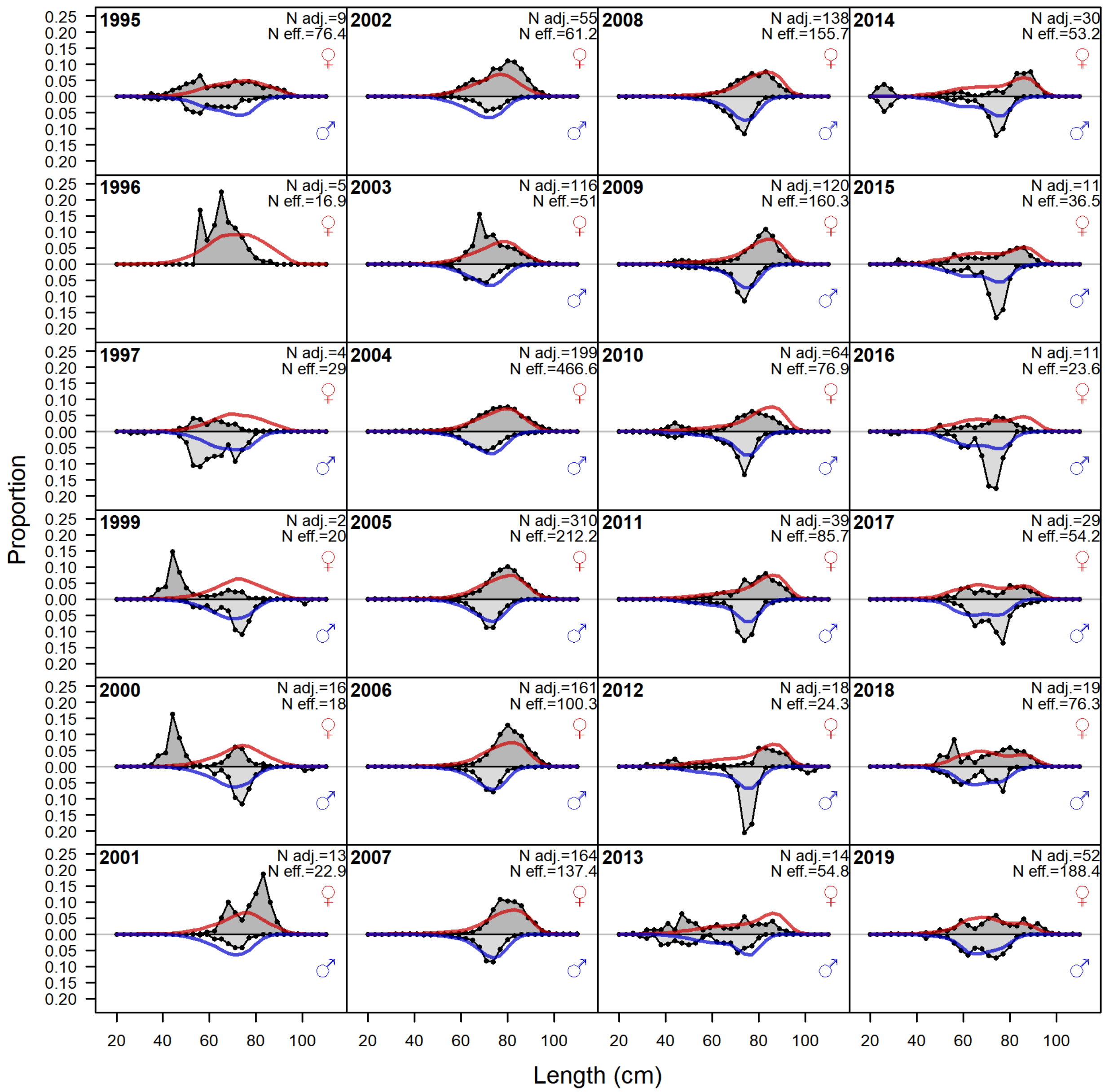


Figure 4.25. Fit to length compositions by year and sex for fleet 4: Discard_LMOT_LL_Rec.

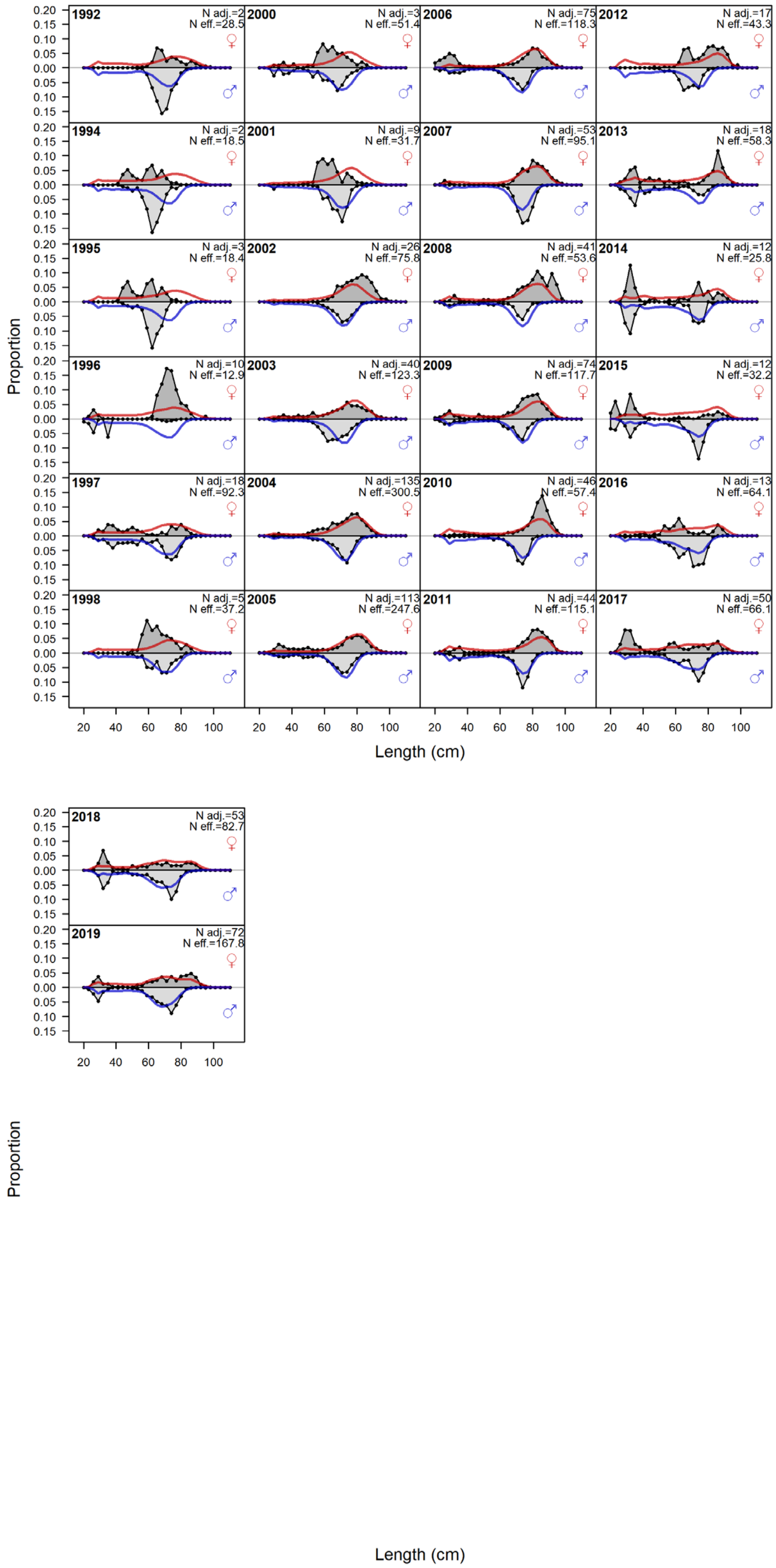
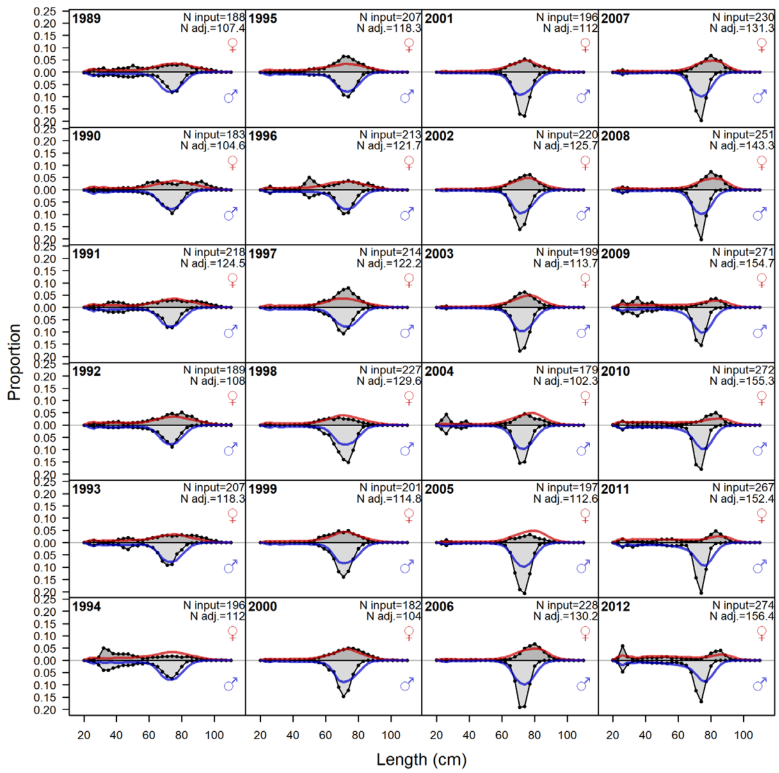


Figure 4.26. Fit to length compositions by year and sex for fleet 5: Discard_SMOT.



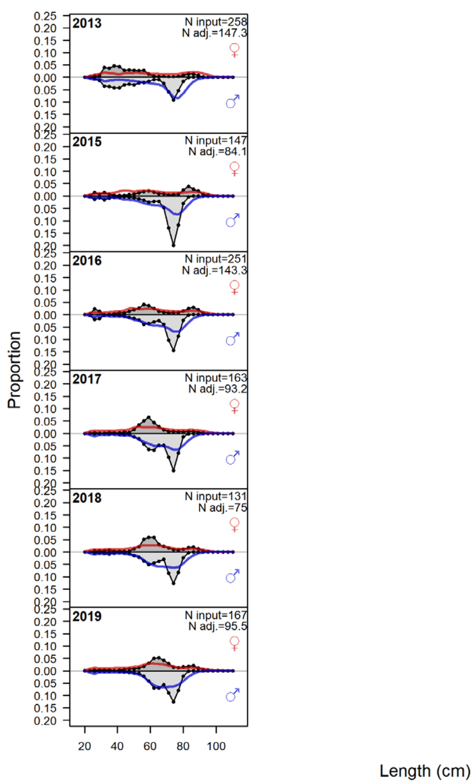


Figure 4.27. Fit to length compositions by year and sex for NEFSC spring bottom trawl survey.

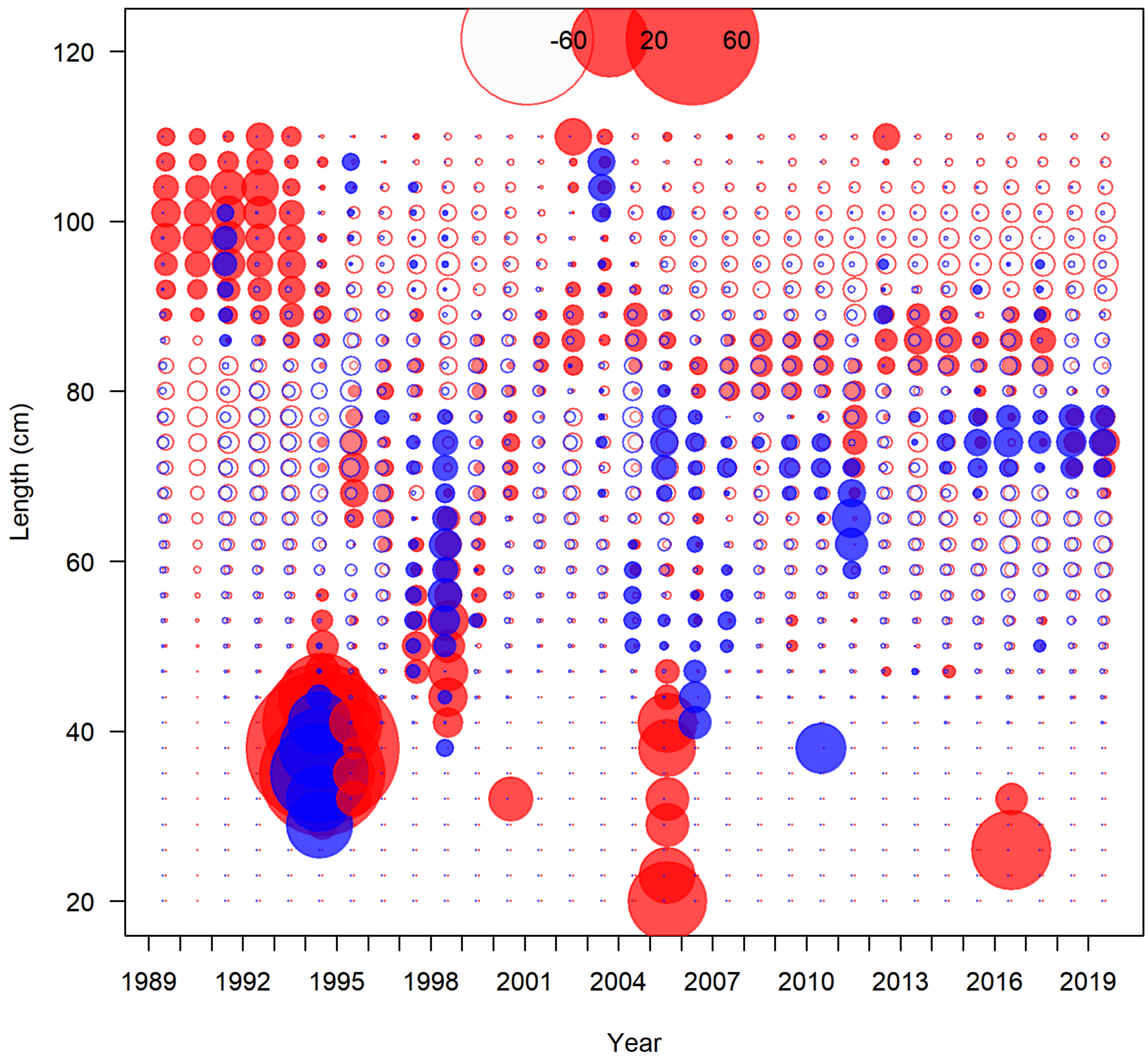


Figure 4.28. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for fleet 1: Landings_SGN_Rec_Others. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

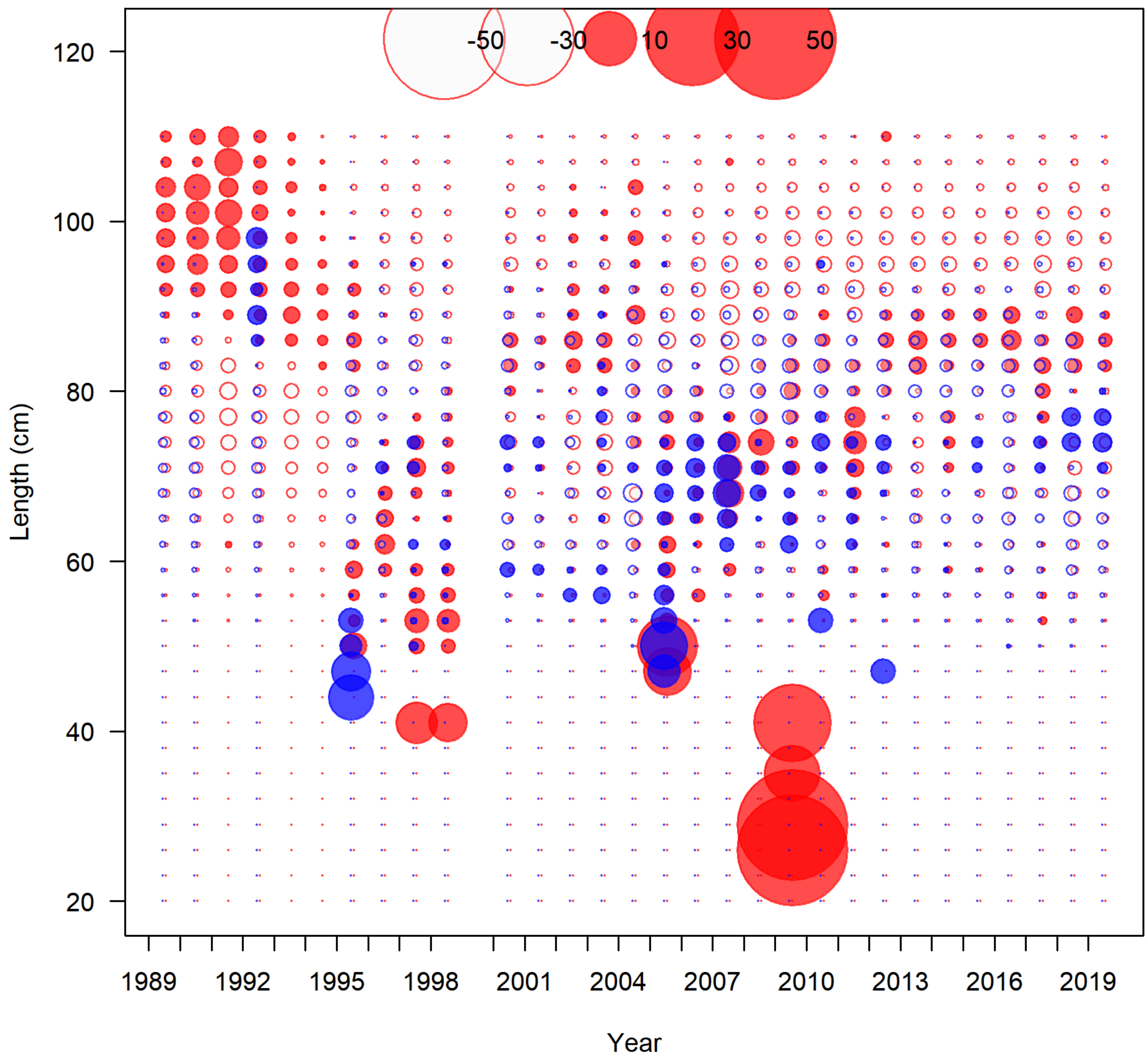


Figure 4.29. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for fleet 2: Landings_LL_OT_Foreign. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

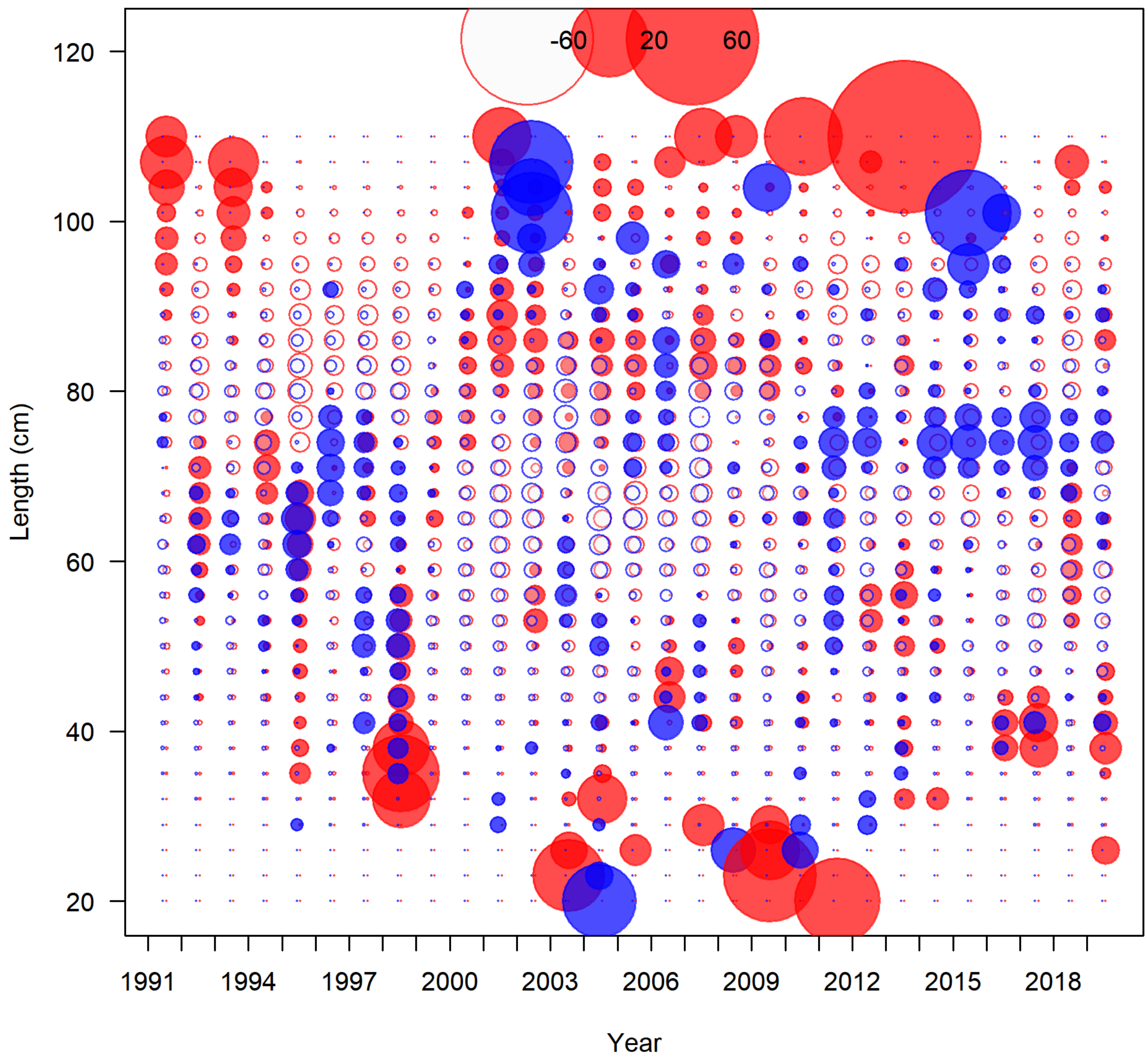


Figure 4.30. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for fleet 3: Discard_SGN_SD. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

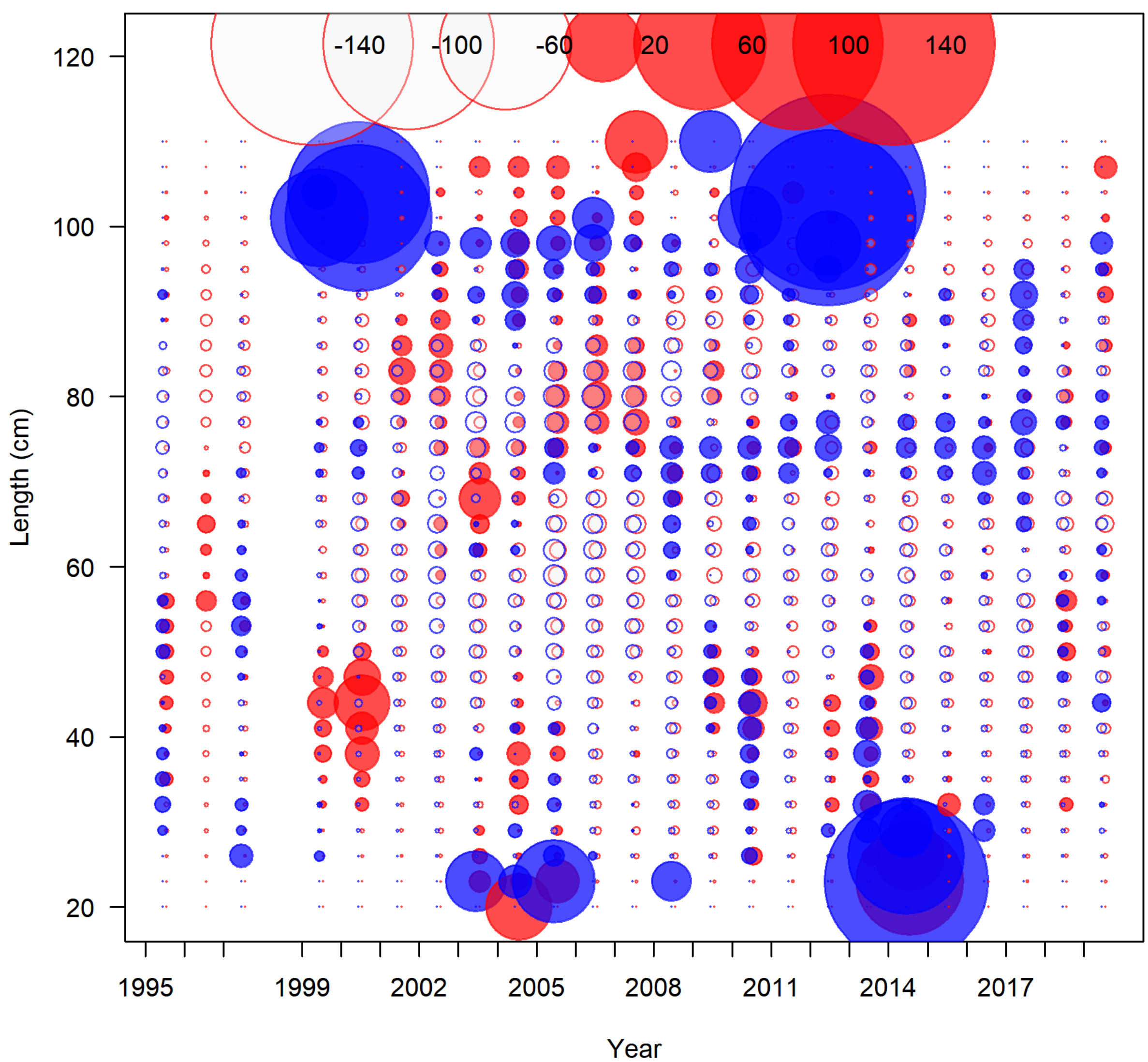


Figure 4.31. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for fleet 4: Discard_LMOT_LL_Rec. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

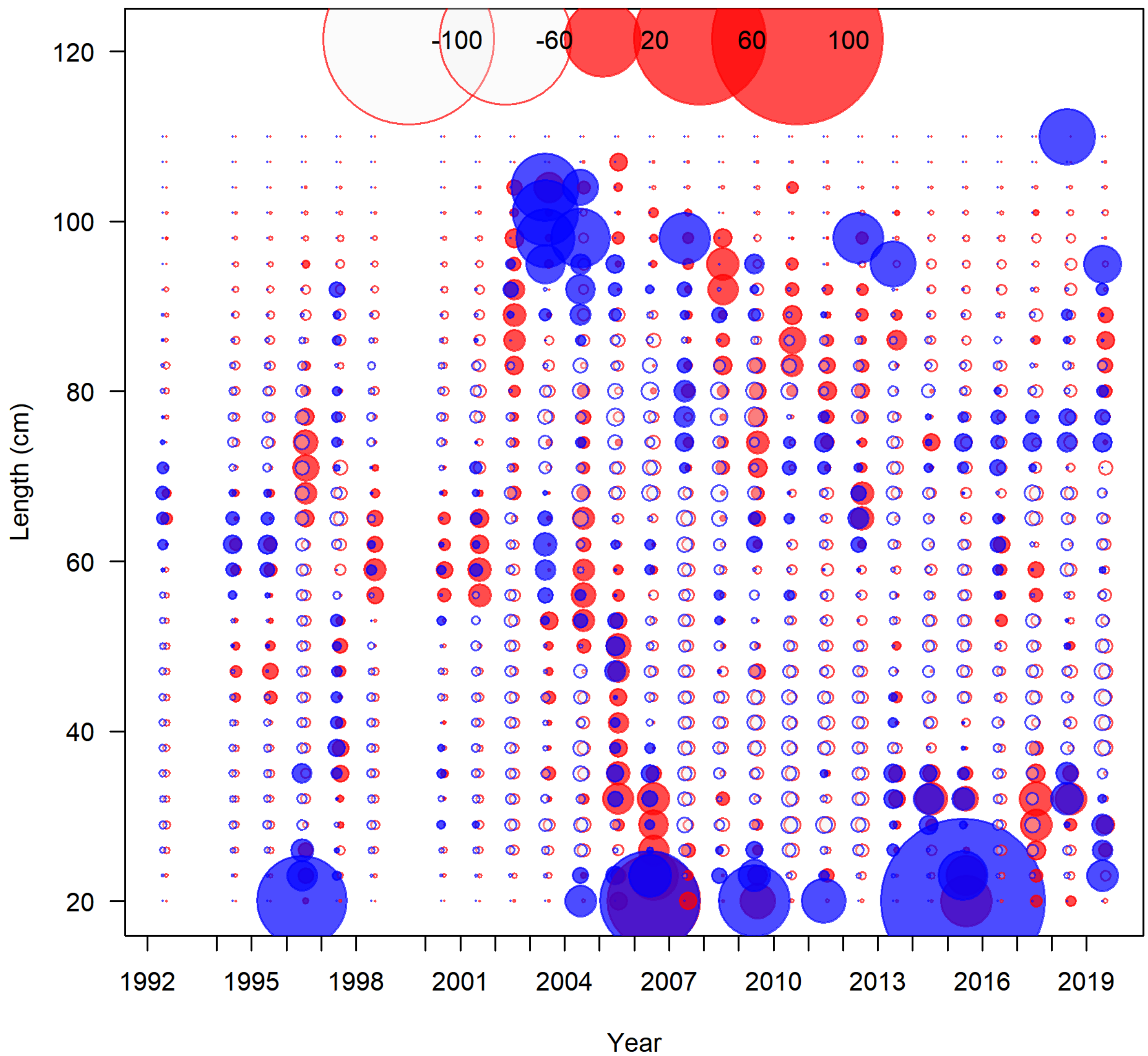


Figure 4.32. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for fleet 5: Discard_SMOT. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

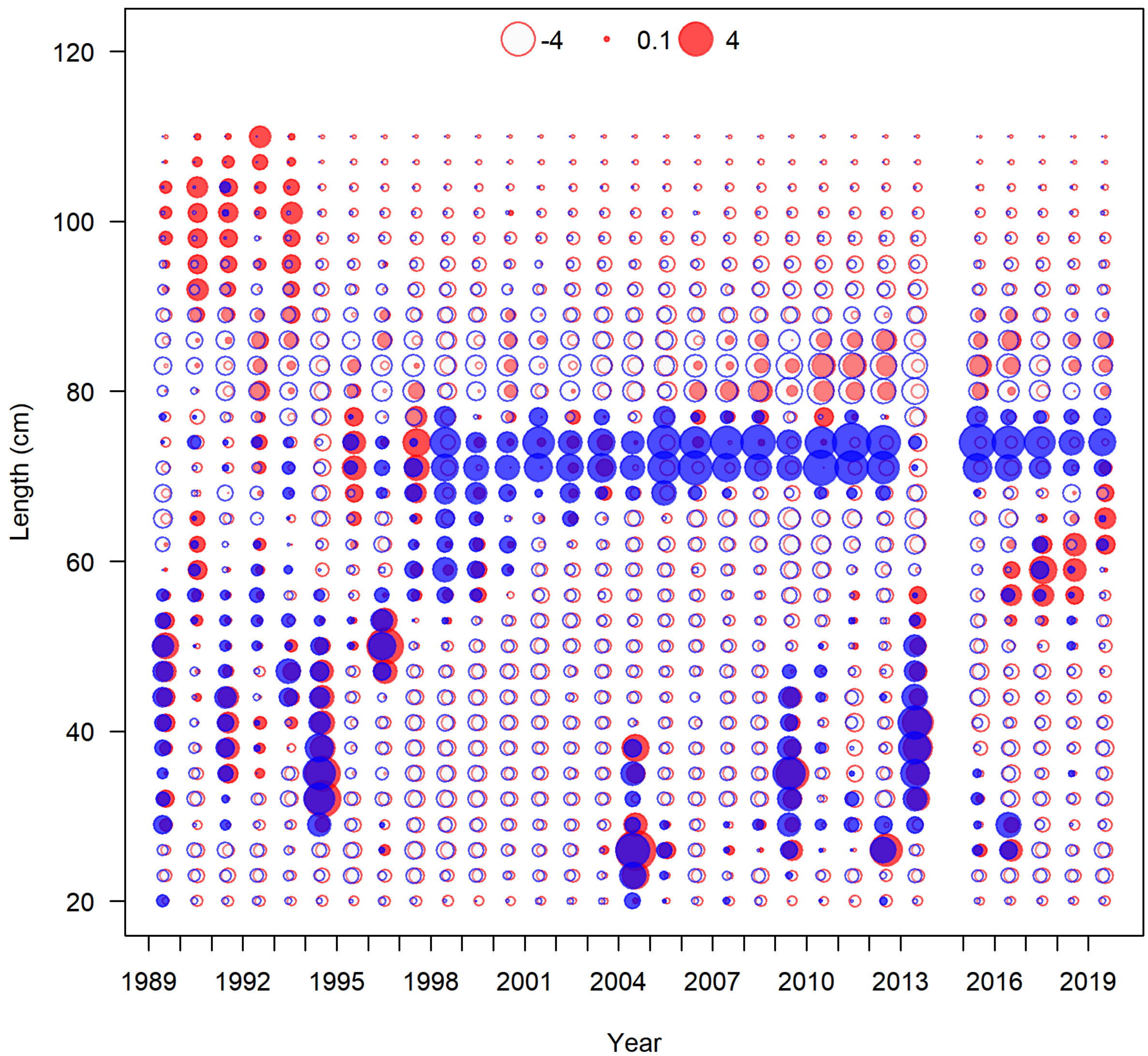


Figure 4.33. Pearson residuals for the fit to length compositions by year and sex (red = female, blue = male) for NEFSC spring bottom trawl survey. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

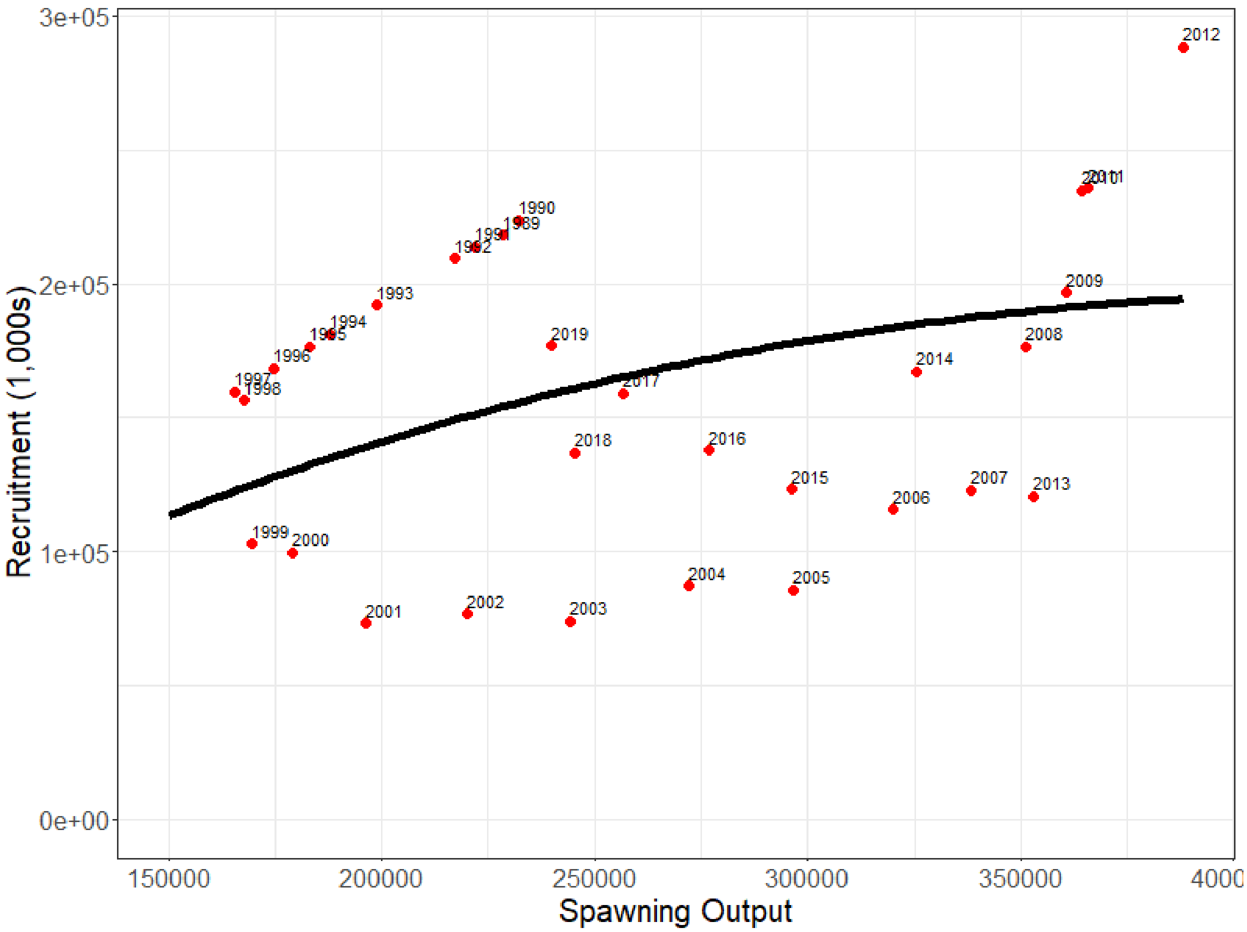


Figure 4.34. Fixed survivorship spawner-recruitment relationship, estimated age-0 recruitment (1,000s), and estimated spawning output by year for Atlantic spiny dogfish.

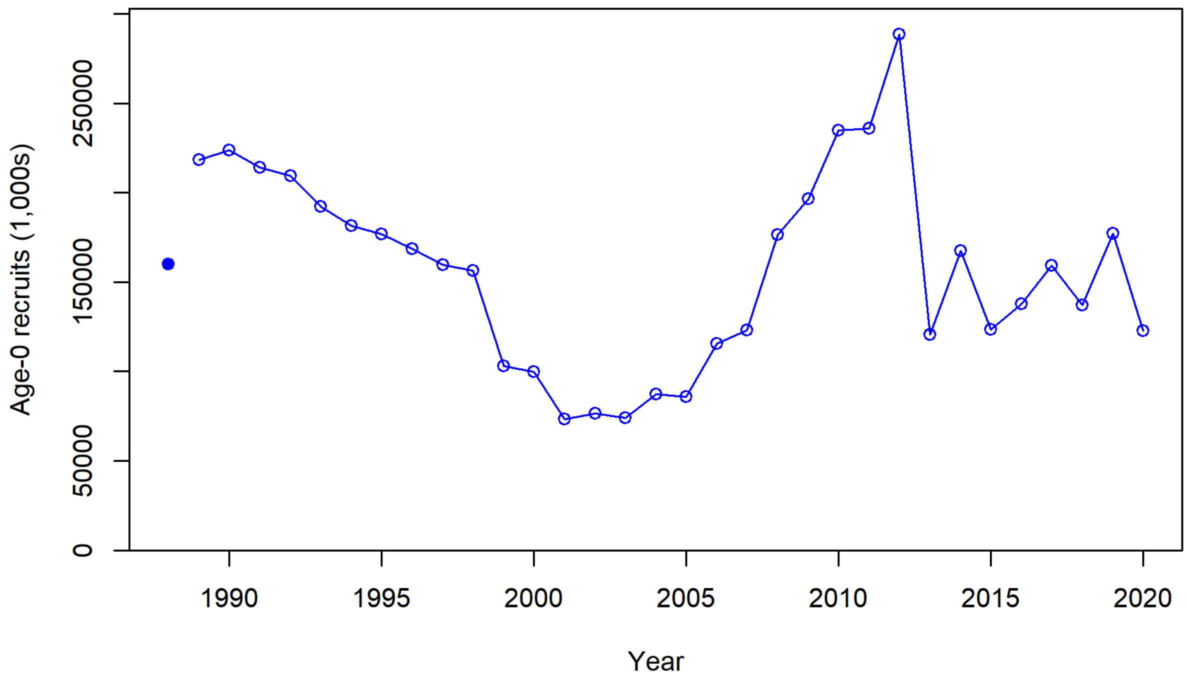


Figure 4.35. Estimated age-0 recruitment (1,000) by year for Atlantic spiny dogfish.

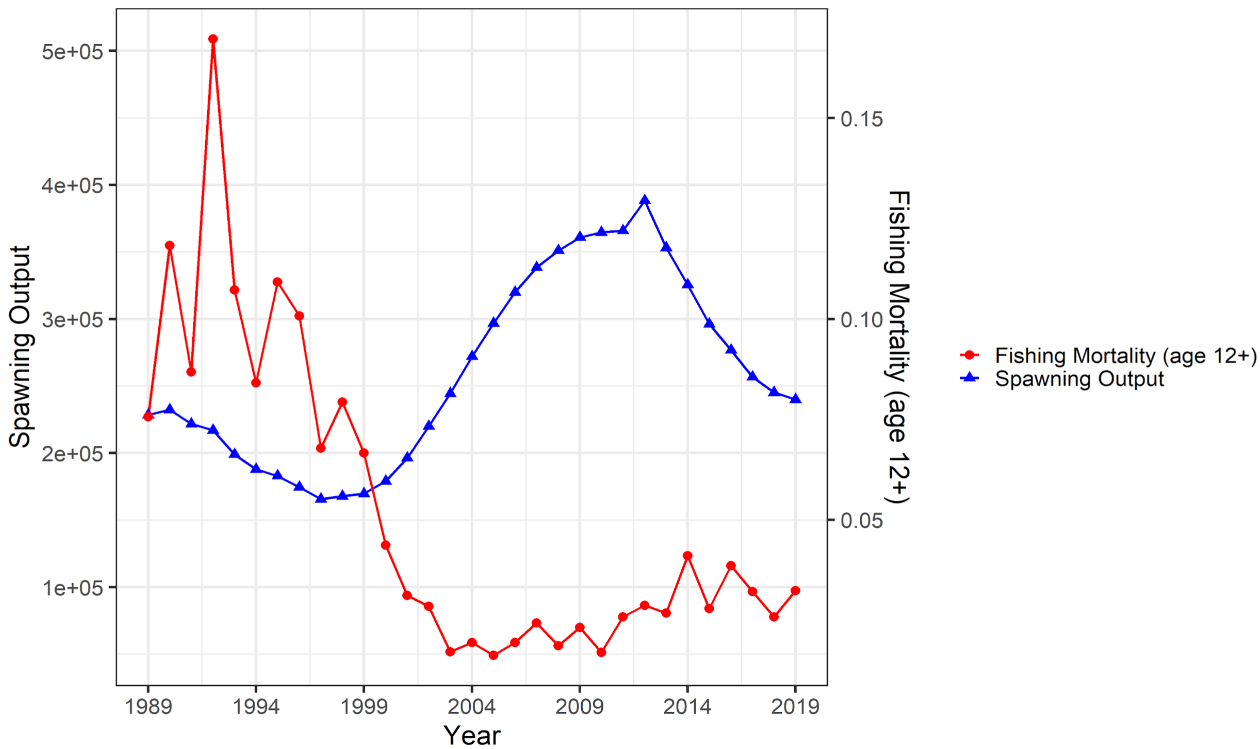


Figure 4.36. Estimated spawning output and fishing mortality (age 12+) by year for Atlantic spiny dogfish.

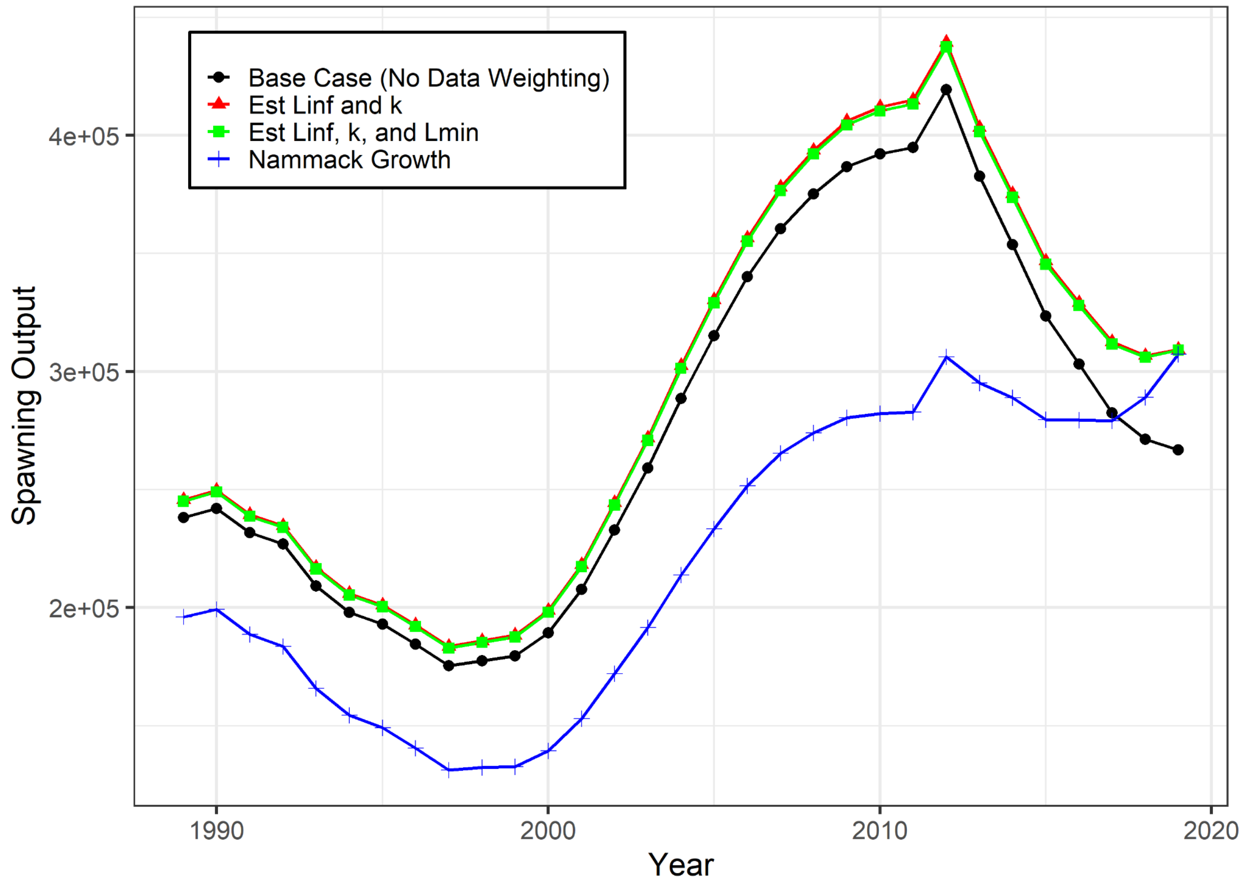


Figure 4.37. Spawning output estimated using different growth assumptions.

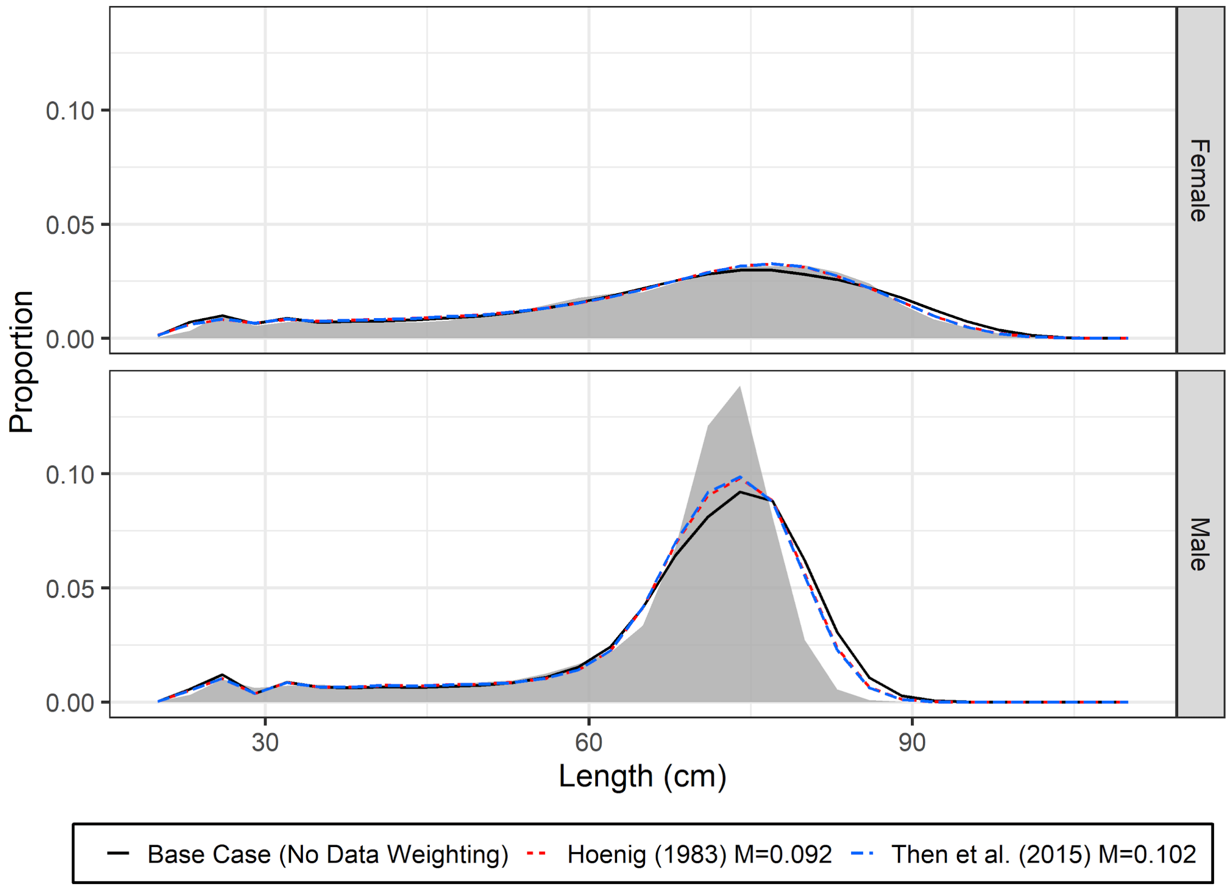


Figure 4.38. Observed (shaded) and model-predicted (line) length compositions by sex and natural mortality assumptions, aggregated across time.

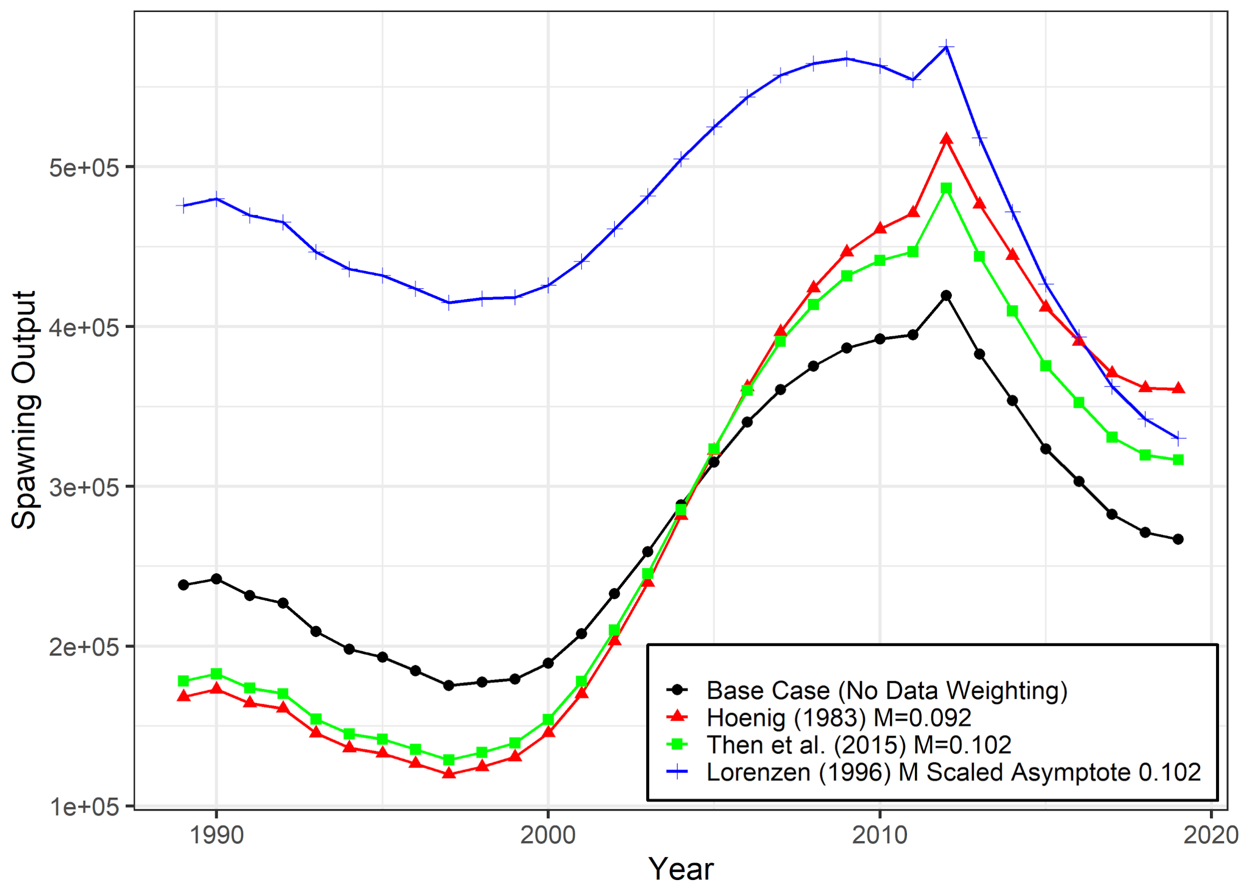


Figure 4.39. Spawning output estimated using different natural mortality assumptions.

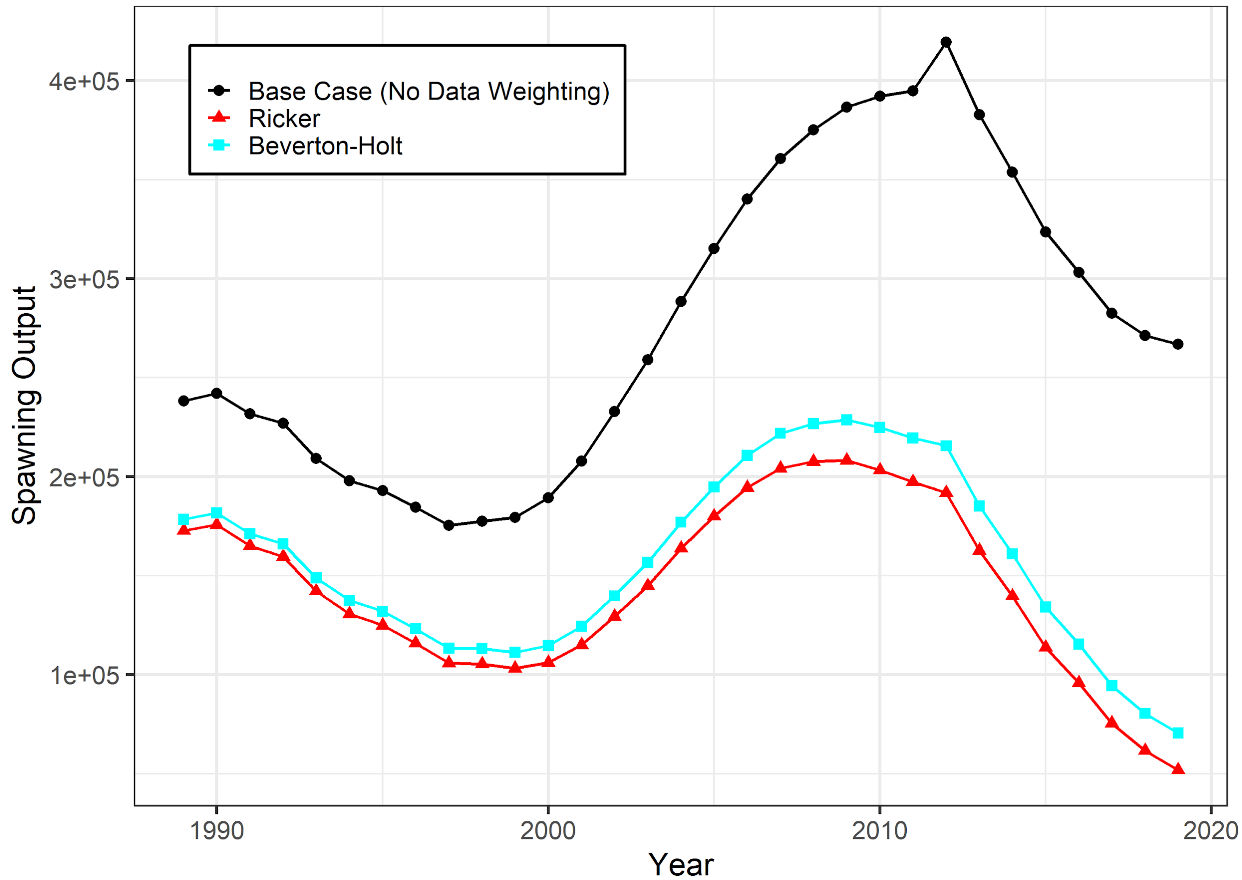


Figure 4.40. Spawning output estimated using different spawner-recruitment relationship assumptions.

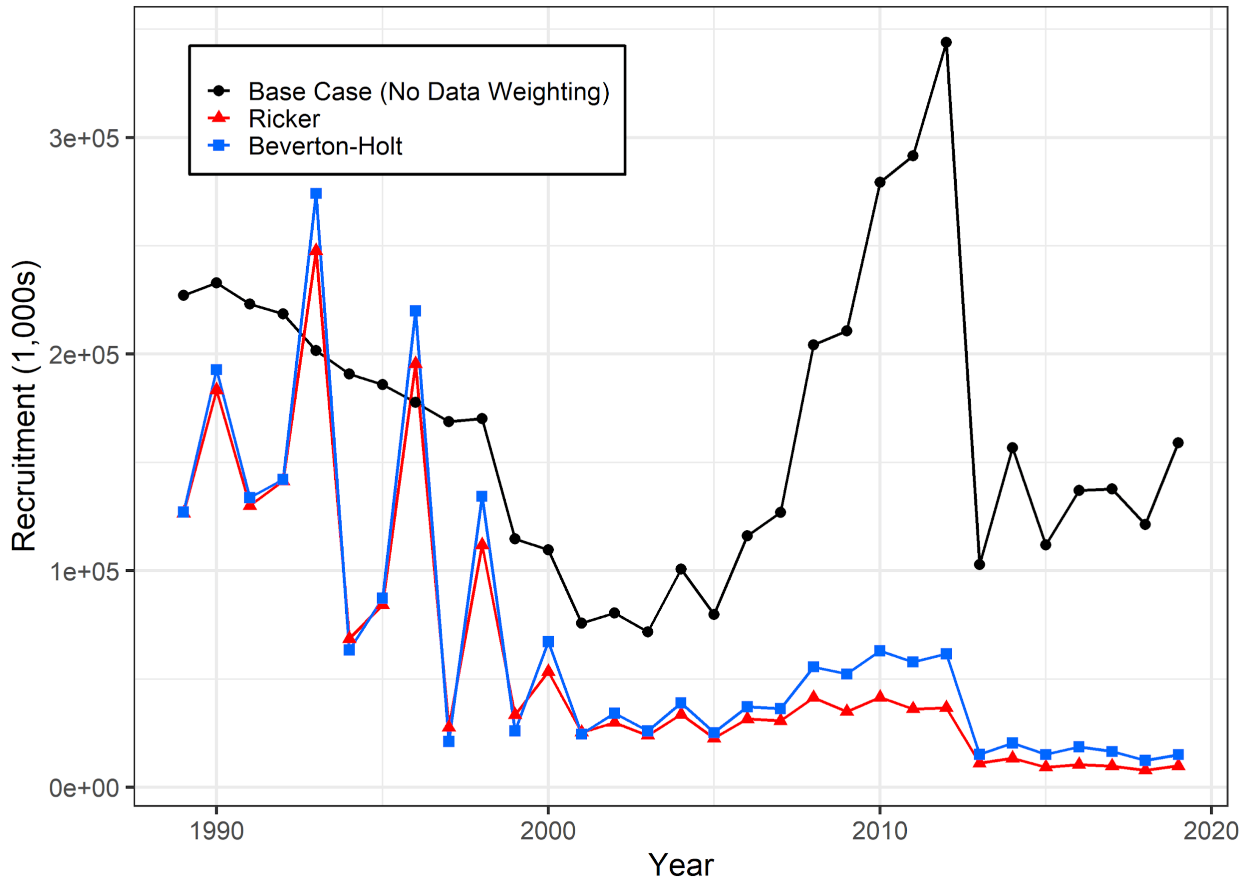


Figure 4.41. Recruitment (1,000) estimated using different spawner-recruitment relationship assumptions.

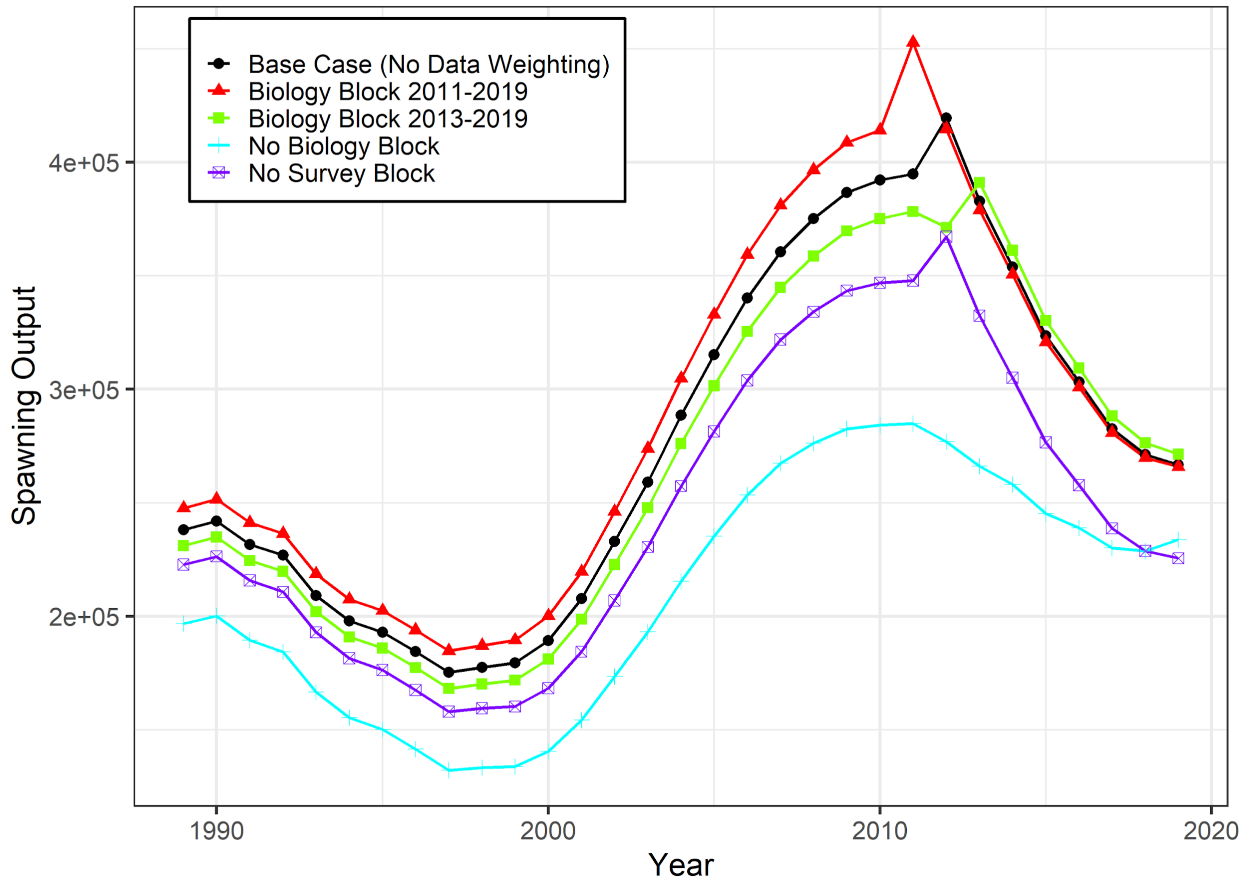


Figure 4.42. Spawning output estimated using different time block assumptions.

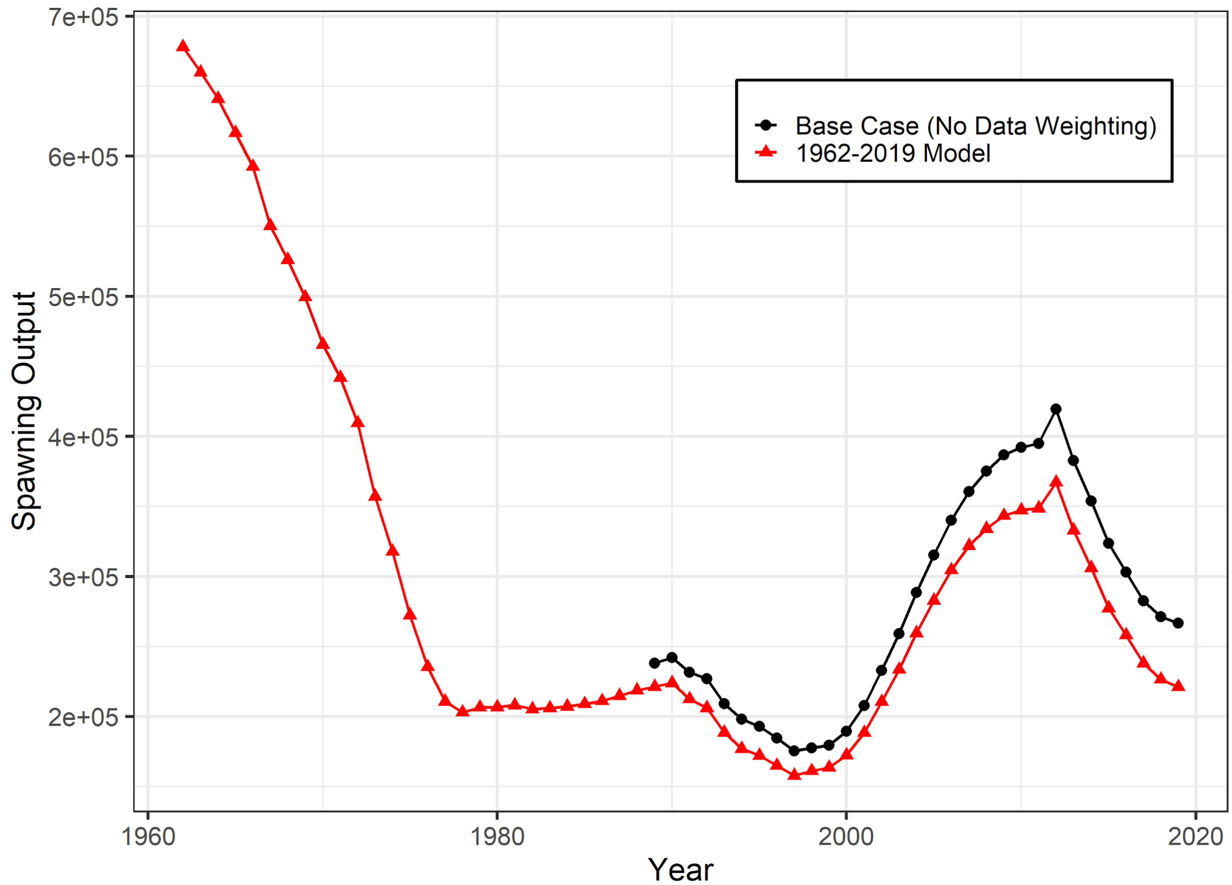


Figure 4.43. Spawning output estimated using different starting year assumptions.

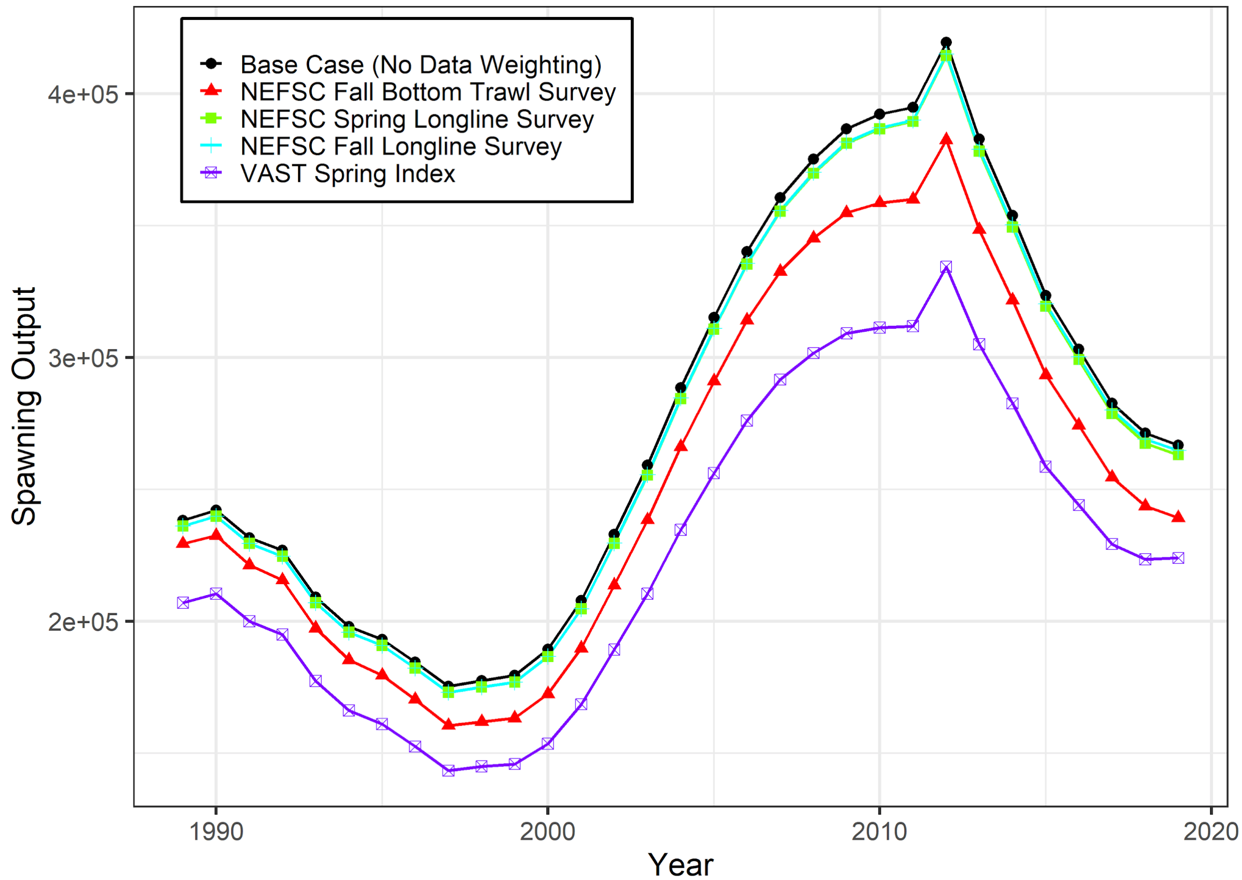


Figure 4.44. Spawning output estimated using different survey data.

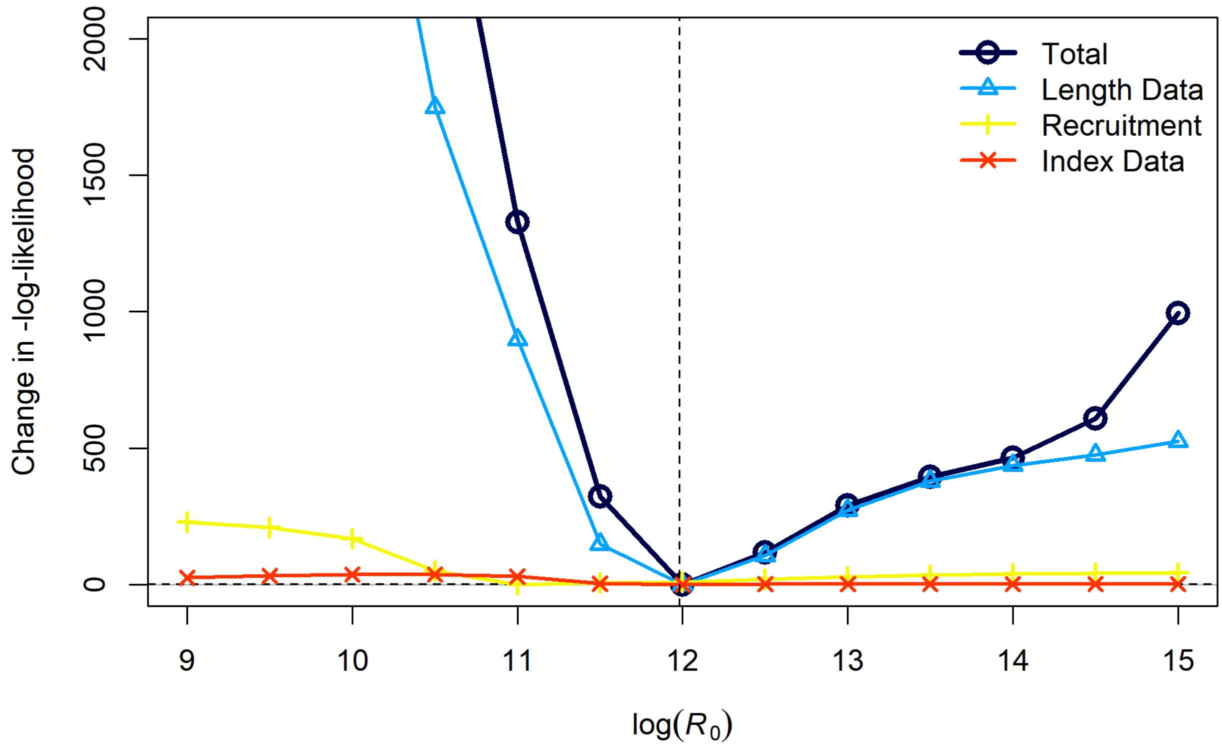


Figure 4.45. Log-likelihood profiles for R_0 for various data components.

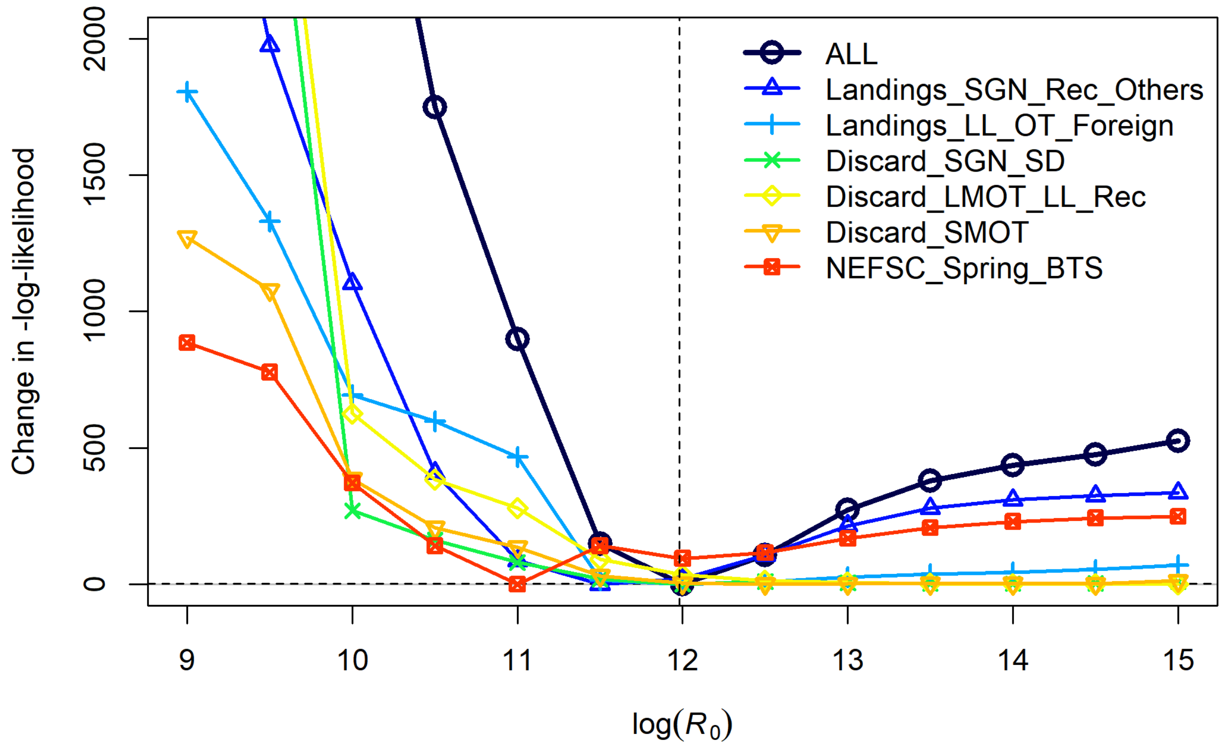


Figure 4.46. Log-likelihood profiles for R_0 for various source of length composition data.

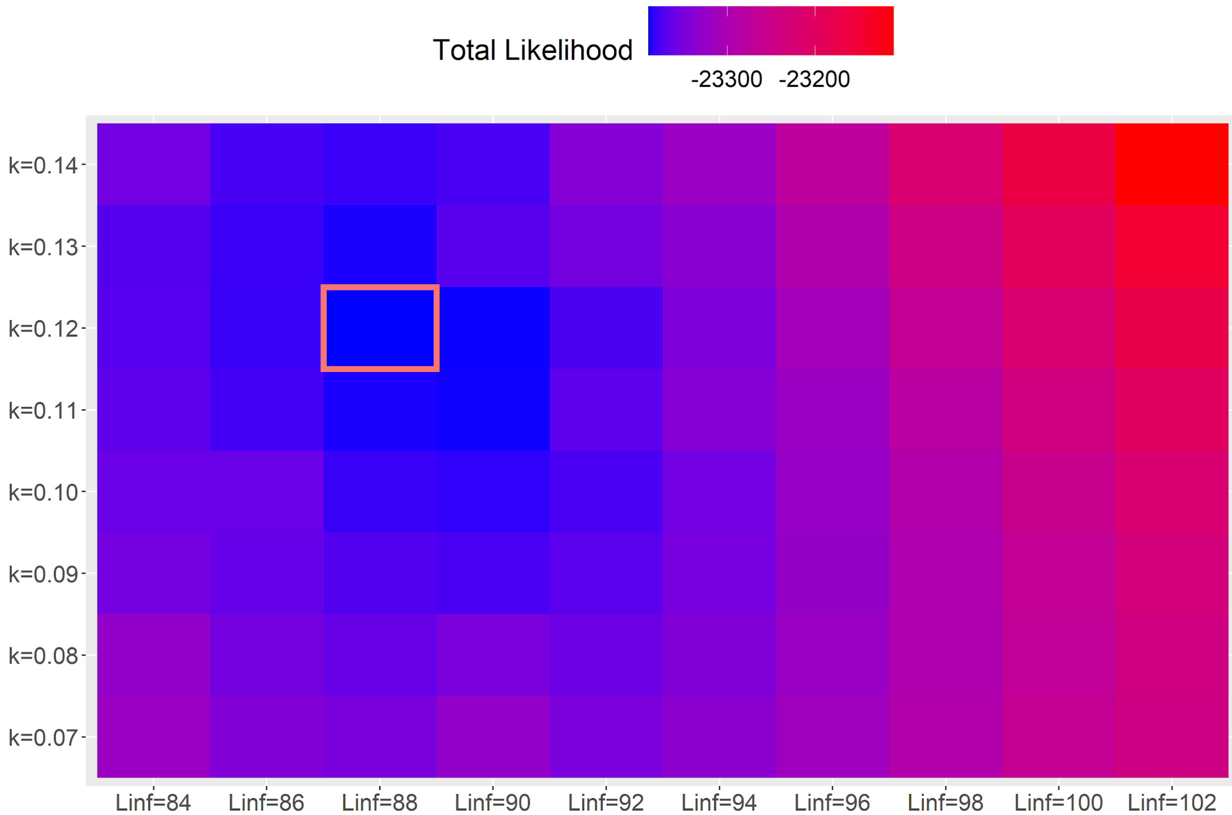


Figure 4.47. Total log-likelihood surface from profiling female L_∞ and k von Bertalanffy growth parameters. The box indicated the run with the smallest total likelihood.

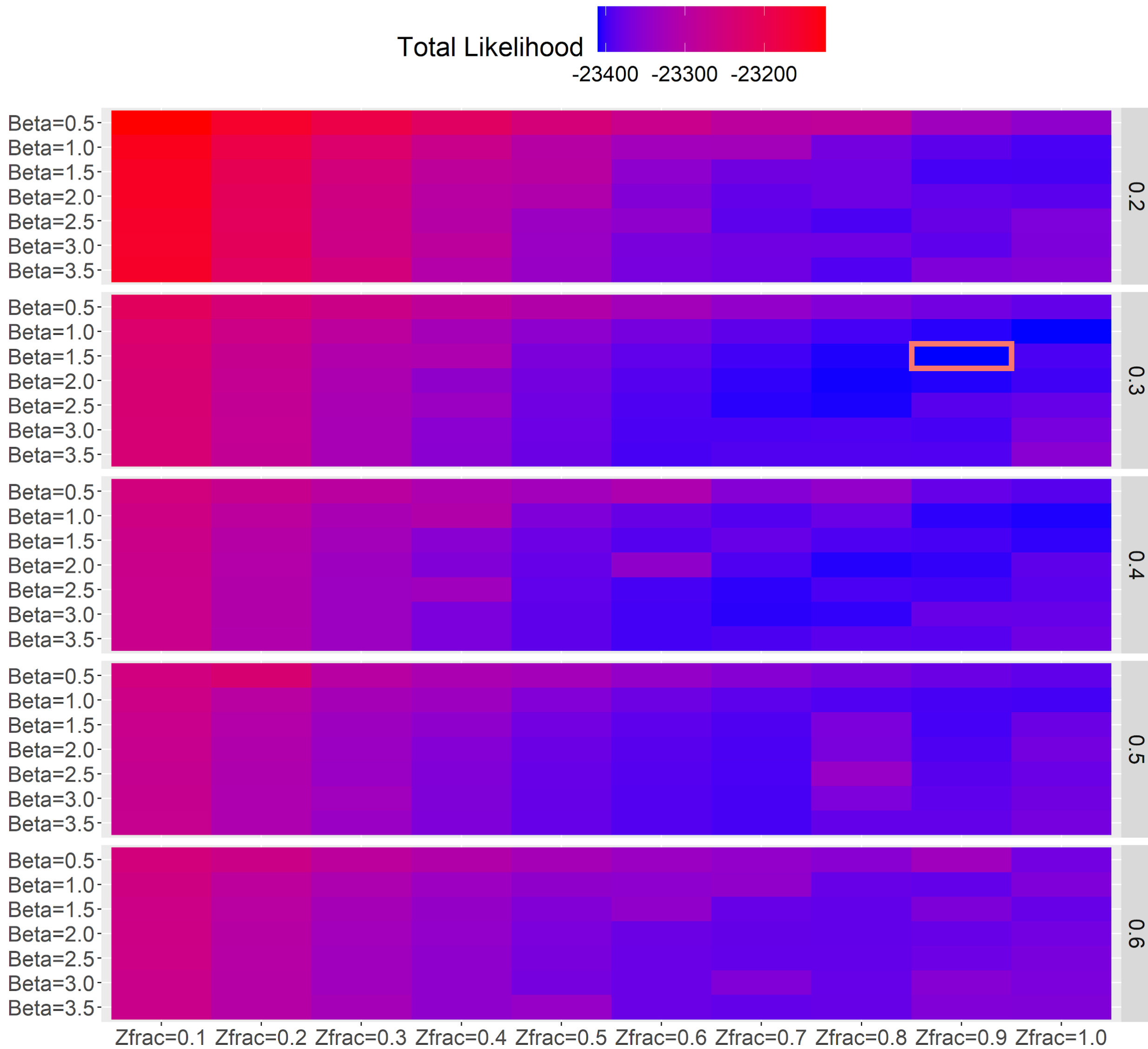


Figure 4.48. Total log-likelihood surface from profiling survivorship spawner-recruitment parameters Z_{frac}, β, and σ_R. The box indicated the run with the smallest total likelihood.

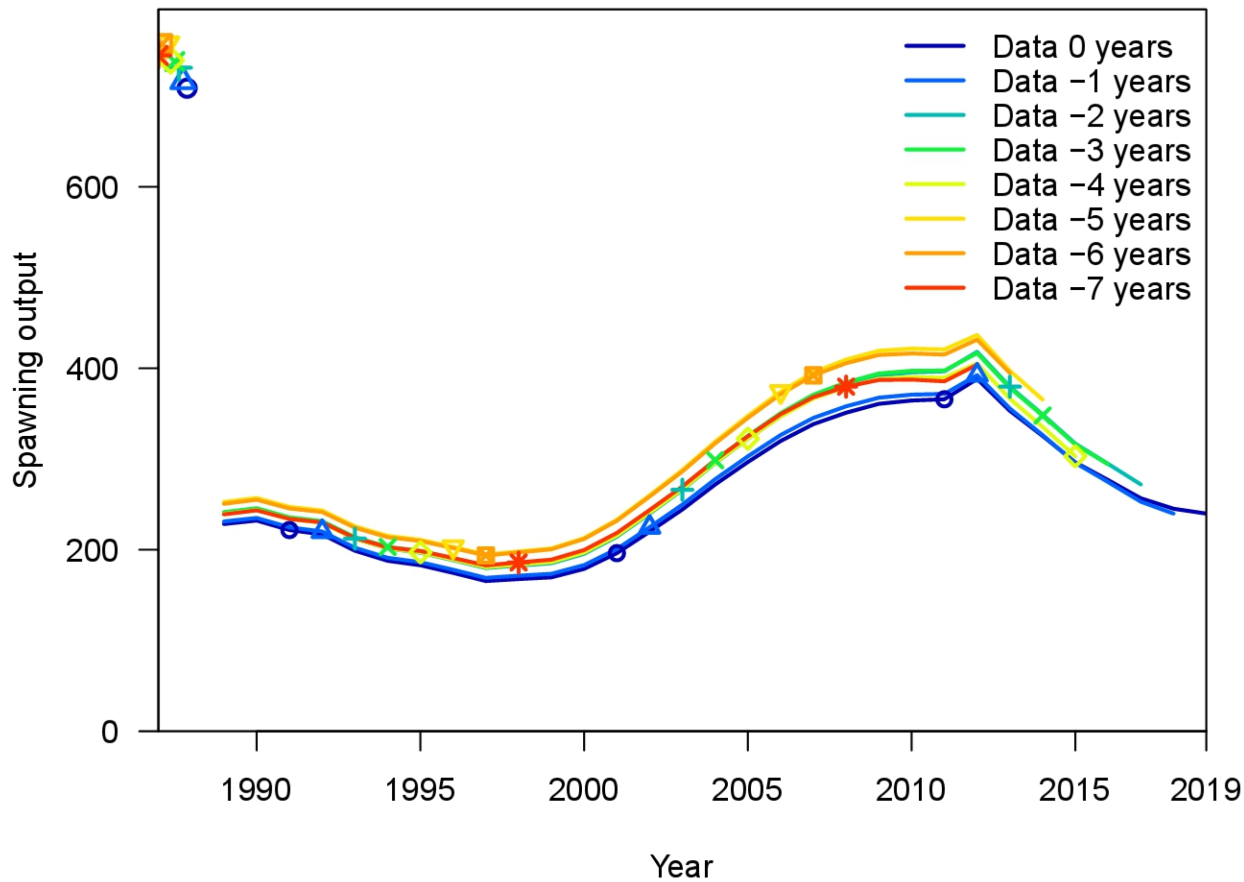


Figure 4.49. Retrospective plot for spawning output.

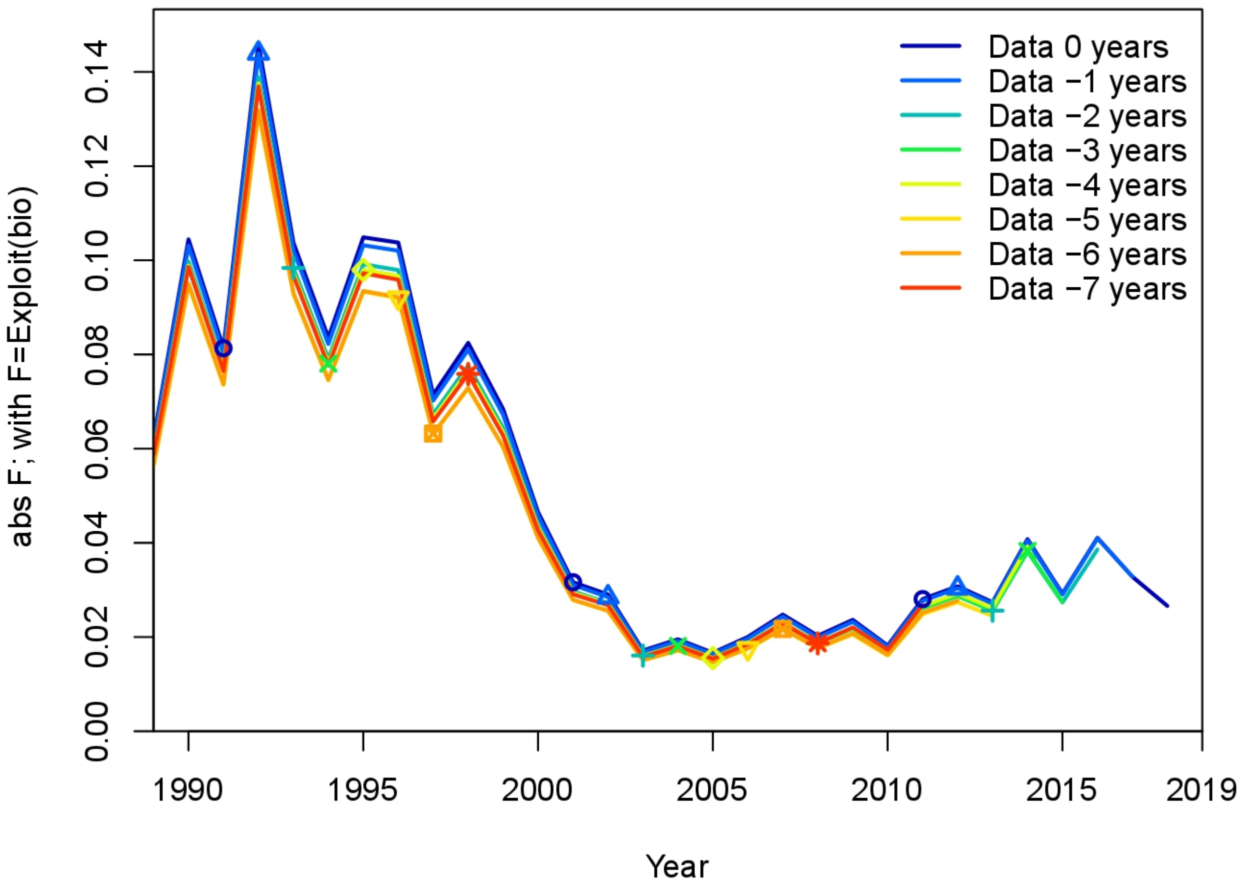
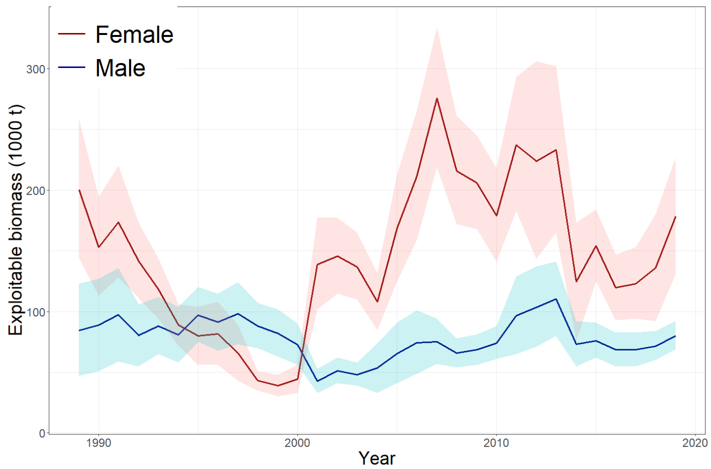
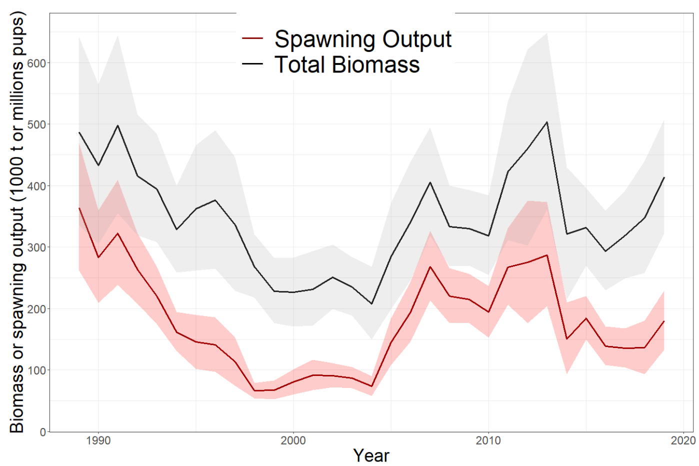


Figure 4.50. Retrospective plot for fishing mortality (age 12+).





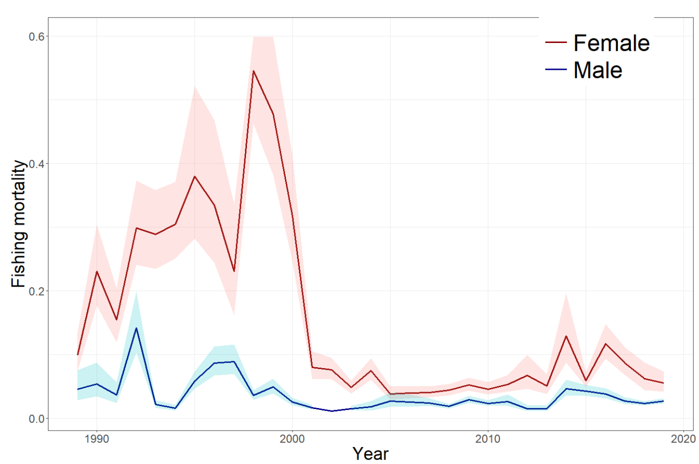


Figure 4.51. Exploitable biomass by sex (top), total biomass and spawning output (middle), and fishing mortality by sex (bottom), from the Stochastic Estimator with 90% confidence intervals.

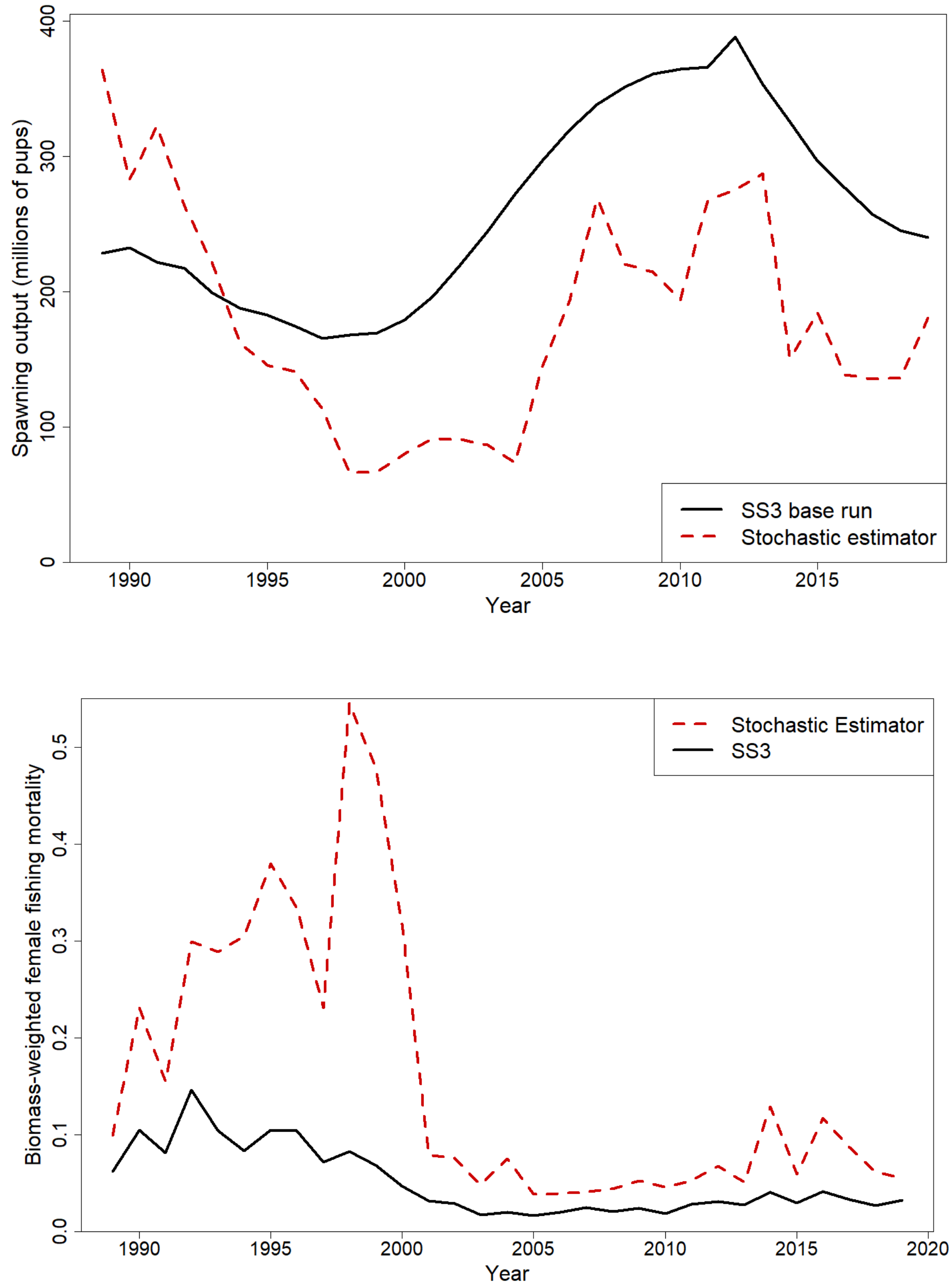


Figure 4.52. Comparison of estimates for spawning output (top) and fishing mortality (bottom) from the SS3 base run and the Stochastic Estimator.

TOR5: STATUS DETERMINATION CRITERIA

"*Update or redefine status determination criteria (SDC; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.*"

Per Recruit Analysis

A length-based per recruit analysis was performed for spiny dogfish; methods and details can be found in Hart and Chang (2022). Figure 5.1 shows yield- and pups-per-recruit at three different L_∞ values, and otherwise with parameters as in the SS3 base run. The fishing mortality that maximizes yield-per-recruit, F_{MAX}, ranges between 0.148 to 0.158. Pups-per-recruit is more constraining, and for the base case growth (L_∞ = 89.24 cm), $F > 0.03$ results in less than two pups-per-recruit. Therefore, fishing mortality needs to be below 0.03 to be sustainable.

Figure 5.2 shows the equilibrium fraction female in the population at F from the per recruit analysis. This can be compared to the observed fraction females observed on the spring trawl survey (Figure 5.3). This fraction, both by numbers and biomass, started out in 1980 at about its unfished equilibrium value, and then declined due to the relatively heavy fishing in the 1990s. The numbers-based fraction then leveled off at about 0.35, whereas the fraction by biomass continued to decline, likely as a reflection of slower growth.

SPR Reference Points

The Working Group examined three putative spawners per recruit (SPR) reference points from the SS3 base model: SPR50%, SPR60% and SPR70% (Table 5.1). The fishing mortality associated with SPR50% is 0.037, which produces less than two pups per recruit. Moreover, the mean fishing mortality between 2012-2019 was below 0.037, but nonetheless the stock rapidly decreased during that time (Figure 5.4).

By contrast, the fishing mortality associated with SPR60% (0.025) gives more than two pups per recruit. During the period when *F* was below this level, the stock increased, but it then decreased in 2012-2019 when *F* was above 0.025 (Figure 5.4). The SPR70% reference points would suggest that overfishing was occurring during the period that the stock was rapidly increasing (2000-2012) and thus is less credible than SPR60%. Based on the combination of theoretical and empirical evidence, the Working Group recommended the SPR60% reference points: a spawning output (analogous to spawning biomass) target of 370.8 million pups and *F* = 0.025. Based on these reference points, and assuming that the overfishing threshold is half the target, the stock was not overfished, but overfishing was occurring in 2019.

Comparison with Previous Reference Points

Previous assessments used a biomass target of SSB_{MAX}, the SSB that produces the maximum recruitment according to the Ricker stock recruit relationship (Rago and Sosebee 2010). A reanalysis of this approach (McManus et al. 2022) estimated SSB_{MAX} = 445,349 mt. Note that this analysis considered spawning biomass to be female biomass greater than 80 cm, consistent with Rago and Sosebee (2010), which is shifted to the right compared to the maturity curve for the latest period. Per recruit analysis indicates that 445,349 mt SSB corresponds to slightly under *F* = 0.03 (Hart and Chang 2022, Table 2), similar to the recommended SPR60% reference points, and would lead to the same status determination. However, both the updated SSB_{MAX} and SPR60% reference points are much greater than those calculated in Rago and Sosebee (2010), who estimated SSB_{MAX} to be 159,288 or 189,553 mt, depending on the assumed area swept by the survey trawl.

The evidence that Atlantic spiny dogfish follows a Ricker model has weakened, based on the updated stock-recruit fits. Additionally, the survival stock-recruit relationship (Taylor et al. 2013) produced a superior fit and more credible results in the SS3 model. The Working Group therefore concluded that the Ricker-based SSB_{MAX} may not be the most appropriate proxy reference point.

Table 5.1. SPR reference points: Target spawning output (thousands of pups), target fishing mortality, and equilibrium catch (mt, including dead discards).

Target type	SPR50%	SPR60%	SPR70%
Eq. Spawning Output	268,707	370,799	457,116
F	0.037	0.025	0.017
Eq. Catch	18,876	16,792	12,657

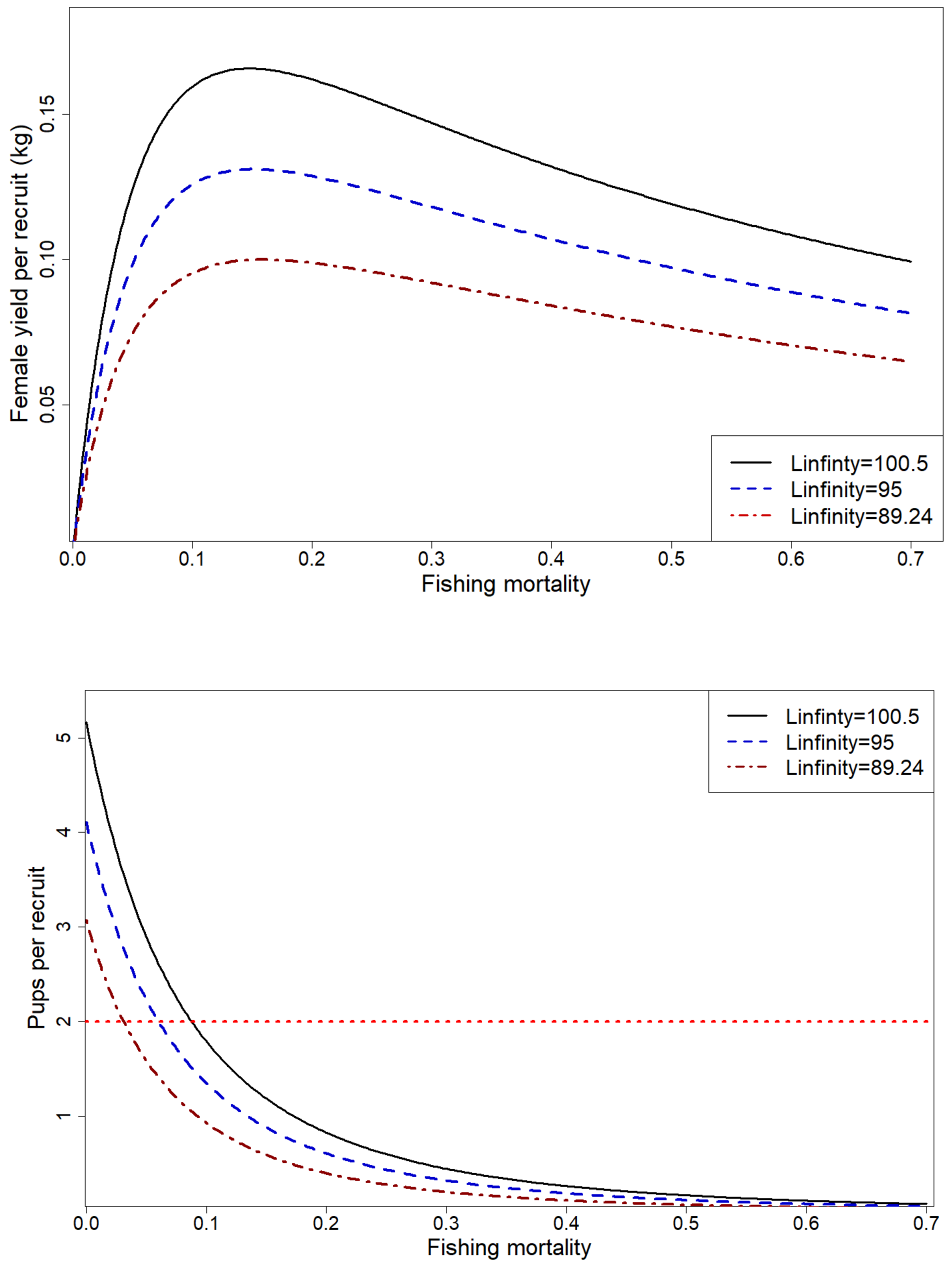


Figure 5.1. Yield and pups per recruit (top and bottom, respectively) assuming Lorenzen natural mortality at three different growth rates.

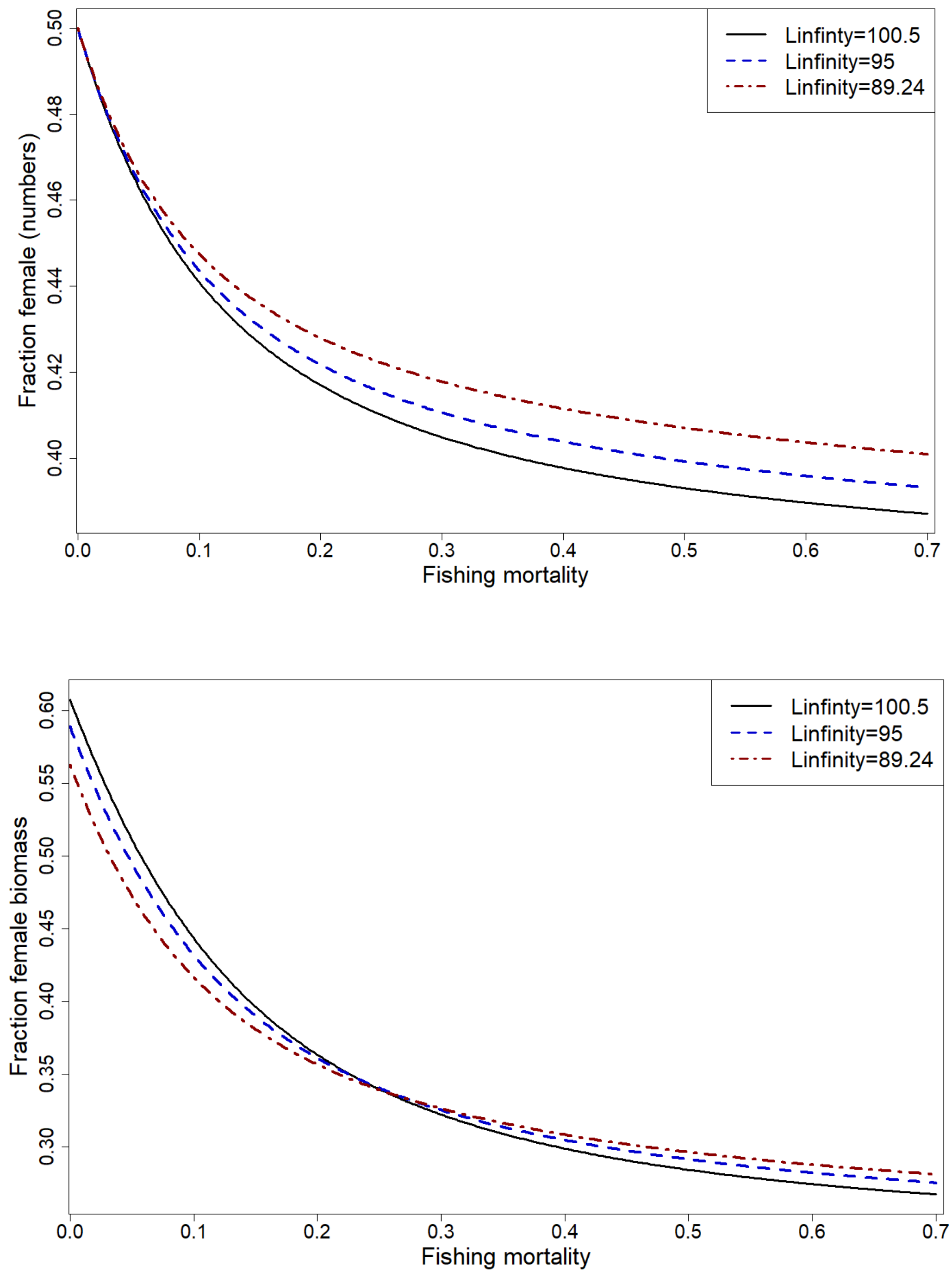


Figure 5.2. Fraction of spiny dogfish that are female (top) and the fraction of female biomass (bottom) from per recruit analysis.

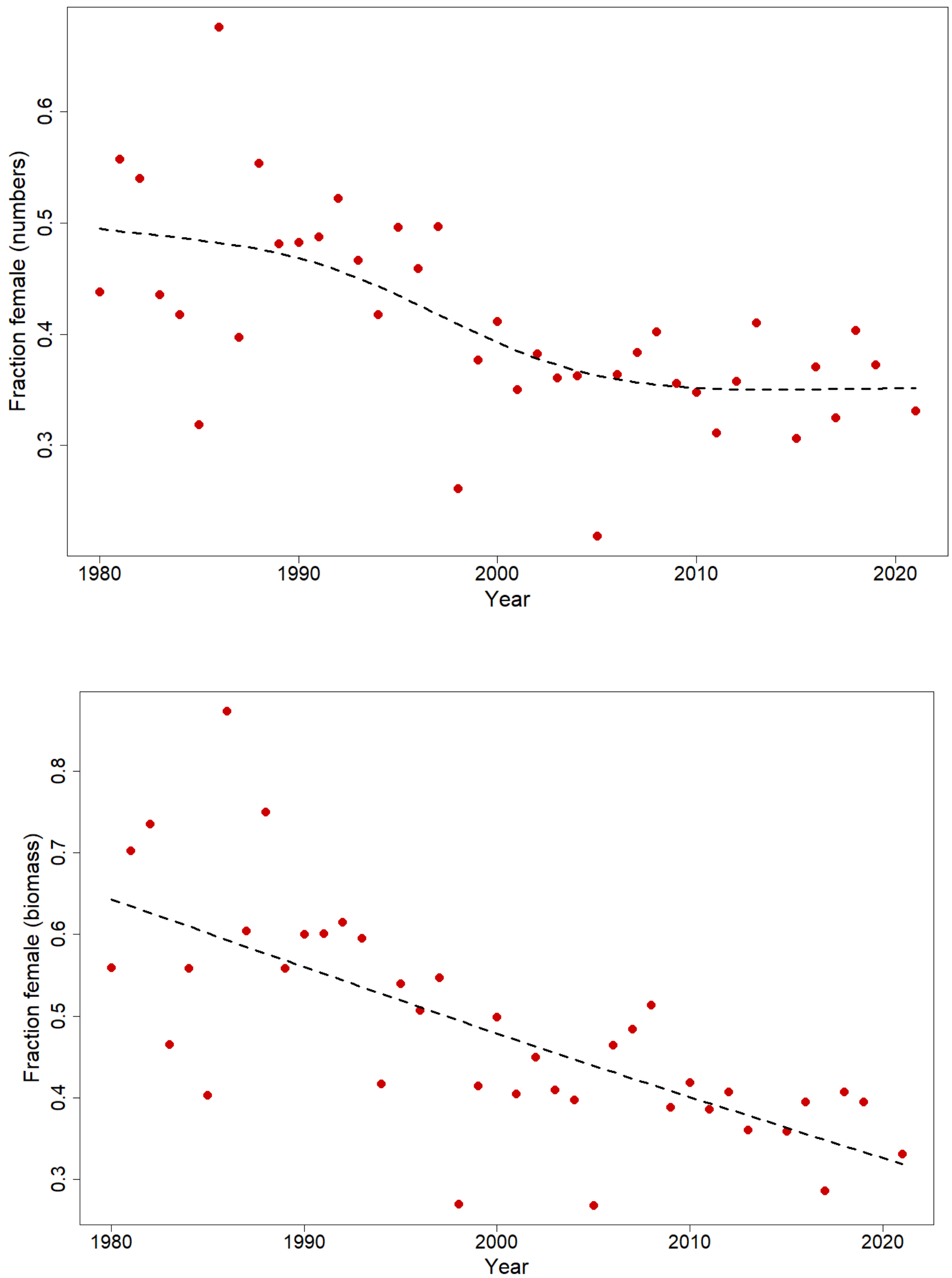


Figure 5.3. Fraction females in terms of numbers (top) and biomass (bottom) from the spring trawl survey. Dashed lines represent GAM smoothers.

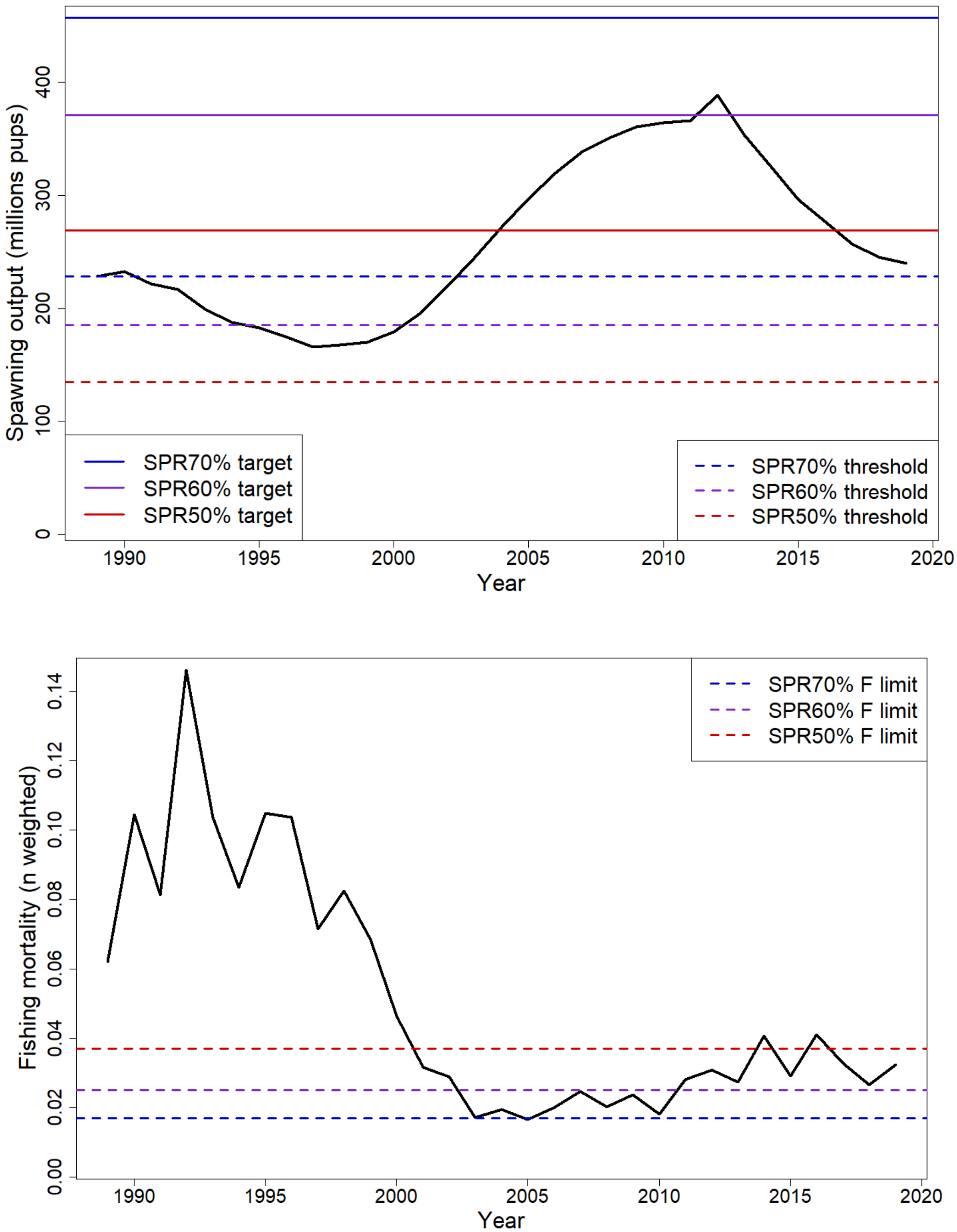


Figure 5.4. Time series of spawning output (top) and fishing mortality (bottom), from the SS3 base model, together with biomass and fishing mortality reference points at SPR50%, SPR60% and SPR70%.

TOR6: PROJECTION METHODS

"Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions"

The Working Group used SS3 as the preferred projection tool for this assessment. The continuity of both the assessment model and projections being conducted with the same software allowed for effective and efficient application of the projection tool. The Working Group conducted three-year projections (2020-2022) under four different fishing mortality rates: $F = 0$, 0.017, 0.025, and 0.037. The latter three figures are the F values associated with the spawner per recruit reference points SPR70%, SPR60% and SPR50%, respectively (Table 6.1, Figure 6.1). Spawning output is projected to decrease between 2019-2020, then increase under all four alternatives, likely due to ongoing maturation of the large 2009-2012 year classes.

These projections use the same biological and fishery assumptions employed in the SS3 estimation model for the 2012-2019 period of reference points, such as fishery fleet selectivity, maturity-at-age, natural mortality, and length compositions. The greatest uncertainties are assumptions regarding growth. Other uncertain assumptions include the amounts of discards, discard mortality, and the selectivities of the various fleets, as well as the uncertainties associated with the terminal year (2019) estimate.

Since the current dogfish fishery is female-targeted, the forecasted catch from SS3 is female-targeted as well. To get a potential male catch for a hypothetical male-targeted fishery, a reasonable potential removal rate for males will have to be assumed and applied to the forecasted male population. Time constraints precluded exploring this possibility during this assessment.

Table 6.1. Projected spawning output (thousands of pups) and catch (mt) under four potential *F* values (0, 0.017, 0.025, 0.037) for years 2020-2022.

Year	Quantity	*F*=0	*F*=0.017	*F*=0.025	*F*=0.037
2020	Spawning Output	165,541	165,541	165,541	165,541
2021	Spawning Output	185,599	181,608	179,460	176,500
2022	Spawning Output	211,191	202,404	197,805	191,618
2020	Catch	0	6,034	9,291	13,790
2021	Catch	0	6,649	10,156	14,905
2022	Catch	0	7,323	11,099	16,122

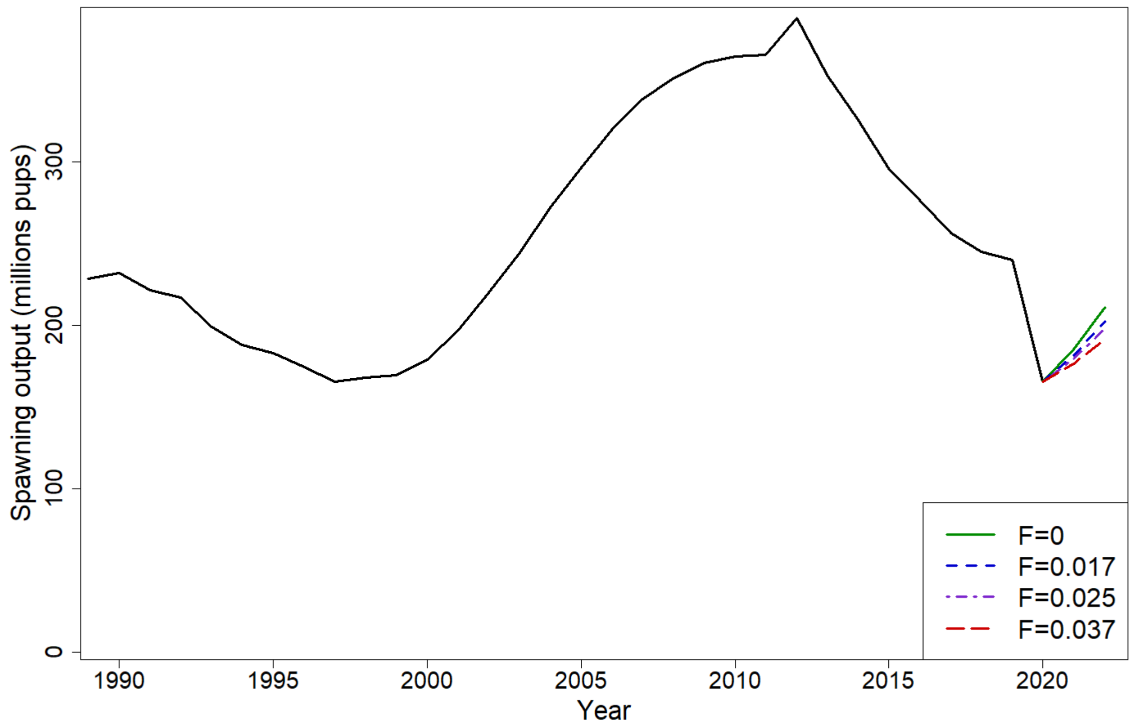


Figure 6.1. Estimated spawning output (1989-2019) from the SS3 base case model, with projected spawning output from 2020-2022 at four different values of *F*: *F*=0, 0.017, 0.025,0.037.

TOR7: RESEARCH RECOMMENDATIONS

"*Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.*"

Status of Previous Research Recommendations

Most recent research recommendations were evaluated by the Working Group, with responses provided.

43rd SAW Stock Assessment Report (NEFSC 2006)

1. "*Incorporate Canadian commercial fishery sample data into the assessment when it is made available (expected in 2007).*" – While commercial landings from Canada are included in the assessment, fishery sampling data were not incorporated into this assessment. Several attempts were made to retrieve available data from Canadian scientists, but such efforts were not successful. It appears that few samples, if any, are collected in Canada given the landings are at very low levels.
2. "*Conduct an ageing workshop for spiny dogfish, encouraging participation by NEFSC, NCDMF, Canada DFO, other interested state agencies, academia, and other international investigators with an interest in dogfish ageing (US and Canada Pacific Coast, ICES).*" – While a workshop was not conducted as part of this assessment, extensive work was conducted between NEFSC and Washington Department of Fish and Wildlife (WDFW) scientists on ageing spines for new growth estimates. There has been communication between the entities on methodologies, quality assurance and control of samples, and data analysis.

3. "*Examine observer data to calculate a weighted average discard mortality rate based on an assumption that the rate increases with catch size.*" – The Working Group did not address this during the assessment.

4. "*Develop experimental estimates of discard mortality in the New England and Mid-Atlantic commercial fisheries.*" – Experimental estimates of discard mortality were not developed during this assessment. Because there has not been research advancements in discard mortality since the last benchmark assessment, discard mortality assumptions were maintained for this assessment. The Working Group is aware of several proposals in recent years to do such work, but those were not funded.

5. "*Conduct a coast-wide tagging study for spiny dogfish to explore stock structure, migration patterns, and mixing rates.*" – Although a coast-wide study has not been funded, there was a new conventional tagging study conducted on a commercial platform since the last assessment that focused on distribution and movements by sex and life stage and movements between the US and Canada from southern New England, the Gulf of Maine, and Georges Bank. The Working Group reviewed this data and results from previously published conventional and high technology tagging studies.

Mid-Atlantic Fishery Management Council Comprehensive Five Year (2020–2024) Research Priorities (2019)

1. "*Integrate recent information on the efficiency of the NEFSC survey gear as it relates to: distribution of spiny dogfish beyond the current NEFSC trawl survey geographic footprint (including inter annual differences); gear efficiency; depth utilization within the footprint; distribution within the survey footprint under different environmental conditions.*" - VAST modeling allowed for the consideration of changing environmental conditions. While recent research has demonstrated presence in waters deeper than the survey, information to standardize survey indices for off-shelf habitat usage was not available for deeper waters. VAST models account for shifting spatial distributions of dogfish as well as produce standardized indices that account for changes in depth selection.

2. "*Explore model-based methods to derive survey indices for spiny dogfish.*" – Model-based methods (VAST) were used to derive spring relative abundance spiny dogfish indices that considered spatiotemporal changes and incorporated multiple surveys into a single index.

3. "*Investigate alternative stock assessment modeling frameworks that evaluate: the effects of stock structure; distribution; updated biological information such as sex ratio and spiny dogfish productivity; state-space models; and sex-specific models.*" – A new length-based assessment model (SS3) was developed for peer review consideration. Additionally, multiple data-limited tools new to spiny dogfish consideration were included (DCAC, DBSRA, PlanB Smooth).

4. "*Evaluate the utility of the study fleet information as it relates to issues identified under priority (1) above.*" – Fishery-dependent data from the Study Fleet and Observer Programs were integrated into deriving catch-per-unit-effort indices from the fishery using model-based approaches that tested environmental data (such as depth, year, month, area) for inclusion in the models. These analyses also suggest that there is a substantive overlap between the survey and Observer/Study Fleet Program. Recent research has also supported this overlap (Sagarese et al. 2015).

5. "*Research opportunities to increase domestic and/or international market demand.*" – Work regarding this recommendation was not conducted as part of this assessment because it is outside of the scope of its terms of reference.

6. "*Expand information on the efficiency of the NEFSC survey gear as it relates to: distribution of spiny dogfish beyond the current NEFSC trawl survey geographic footprint (including inter annual differences); gear efficiency; depth utilization within the footprint; distribution within the survey footprint under different environmental conditions.*"
See the response to recommendation #1 above.

7. "*Continue ageing studies for spiny dogfish age structures (e.g., fins, spines) obtained from all sampling programs (include additional age validation and age structure exchanges), and conduct an ageing workshop for spiny dogfish, encouraging participation by NEFSC, Canada DFO, other interested state agencies, academia, and other international investigators with an interest in dogfish ageing (US and Canada Pacific Coast, ICES).*" – New ageing analyses were conducted as part of the assessment to understand how growth has

changed in recent years, with the hope of incorporating this new growth information into the assessment model.

8. "*Evaluate ecosystem effects on spiny dogfish acting through changes in dogfish vital rates*" – Environmental variables were incorporated into the VAST modeling to understand environmental drivers on the stock. While this work did not directly address specific vital rates, the modeling is done under the theory that the relationships reflect spiny dogfish habitat needs to maintain vital rates.

Mid-Atlantic Fisheries Management Council Scientific and Statistical Committee Research Recommendations (2020)

1. "*Revise the assessment model to investigate the effects of stock structure, distribution, sex ratio, and size of pups on birth rate and first year survival of pups.*" The development of the SS3 model allowed for improved population dynamics modeling of the species. The SS3 model allows for estimating sex ratios and pups (age-0) stock abundance. The SS3 framework also allows for incorporating stock structure and distribution changes in various ways, although this was not including the present model. Spiny dogfish is currently managed as one stock and distributional changes were evaluated in TOR1, with data products from this informing the model through VAST indices. Pup size is not a model input into SS3 at this time.
2. "*Explore model-based methods to derive survey indices for spiny dogfish.*" Model-based methods (i.e., VAST) were used to derive relative abundance indices for use in the assessment model. Additional model indices were explored for the NEAMAP survey.
3. "*Consider development of a state-space assessment model.*" New stock assessment modeling frameworks have been built to allow for state-space modeling of biological processes within the model, most notably used in the region being the Woods Hole Assessment Model (WHAM, Stock and Miller 2021). However, WHAM is presently an age-structured model that requires annual ageing information. Because this information is lacking for spiny dogfish, this recommendation was not possible to pursue at this time, and developing a SS3 model was prioritized instead. Additionally, there is a State-Space Modeling Research Track Assessment Working Group underway that will explore the

application and use of state-space models across a wide range of stocks in the Greater Atlantic Region. The State-Space Research Track is also working on creating a model framework that fits to length composition data, which could be explored on dogfish in the future.

4. "*Compile and examine the available data from large scale (international) tagging programs, including conventional external tags, data storage tags, and satellite pop-up tags, and evaluate their use for clarifying movement patterns and migration rates.*" A synthesis of tagging information currently available was conducted as part of the assessment, including a review of new NEFSC tagging results since the last assessment.

5. "*Investigate the distribution of spiny dogfish beyond the depth range of current NEFSC trawl surveys, possibly by using experimental research or supplemental surveys.*" The Working Group reviewed available fishery independent data that may be able to address this question, but none were identified. Analyses from the Study Fleet and Observer Program datasets (specifically modeled CPUE from covariates) indicated depth was negatively correlated to CPUE, including some data points that were at depths greater than 200m. VAST models did not indicate a significant shift to deeper water or outside the survey range.

6. "*Continue ageing studies for spiny dogfish age structures (e.g., fins, spines) obtained from all sampling programs (include additional age validation and age structure exchanges), and conduct an ageing workshop for spiny dogfish, encouraging participation by NEFSC, Canada DFO, other interested state agencies, academia, and other international investigators with an interest in dogfish ageing (US and Canada Pacific Coast, ICES).*" New ageing analyses were conducted as part of the assessment.

7. "*Evaluate the ecosystem context of spiny dogfish including quantifying their role as predator and prey, and effects of climatic factors such as changes in temperature and salinity on the distribution, growth and survival, as they impact both population dynamics and reference points.*" A new study on the effects of groundfish on the spiny dogfish population was recently published and was reviewed by the Working Group. The Working Group also conducted a literature review of spiny dogfish diet. The VAST modeling also considered several environmental variables to understand their impact on spiny dogfish abundance.

2022 Research Track Stock Assessment Working Group Recommendations

1. Develop a consistent sampling program for ageing Atlantic spiny dogfish. Sampling should occur at minimum annually, and ideally include samples from both spring and fall seasons. Fish over the species' entire size range should be sampled. This includes near-term embryos, in order to assess timing, identification criteria, and spine base diameter at first annulus deposition to better inform ageing of young fish. It is also imperative to ensure that large spiny dogfish are obtained to get a better sense of maximum ages and inform parameterization (e.g., L_∞ estimates). Lacking appropriate growth information will result in increased uncertainty in the assessment model's estimates of stock size and mortality rates. Such growth investigations should include size at birth and maturity, as those are intricately related to growth. Investigation into alternate ageing methods should continue, owing to the large uncertainty inherent in ages estimated from worn spines using current methods. Finally, improve routine cleaning protocols for spine sampling in order to reduce potential damage to spine enamel and enable more accurate ageing.
2. Continue exploration into the spatial distribution of spiny dogfish. Such work should expand upon the analyses discussed and presented herein regarding the environmental drivers on spiny dogfish movement by sex and size, and whether such relationships have resulted in changes in distribution over time. Directed research should also be conducted on the seasonal or intra-annual movement of spiny dogfish. Questions remain regarding what component of the spiny dogfish population exists outside of the federal trawl survey bounds off the shelf, and whether such biomass varies seasonally or interannually. Such knowledge will allow for informing survey catchability. If possible, exploring environmental correlations to the degree of on- and off-shelf distribution may allow for predicting this dynamic over time, and provide a catchability time series for stock assessment model use.
3. Further explore the sensitivity of the SS3 model parameterization and configuration.
4. Conduct directed studies that estimate discard mortality rates for spiny dogfish by commercial and recreational harvesting gear type.
5. Develop state-space models that can tune to lengths. Such a model is worth considering if/when the tools are developed within SS3. When available, a review of results from the

State-Space Research Track Working Group should be conducted to evaluate the efficacy of developed tools for spiny dogfish.

6. Investigate prospective contributors to the decline in maturity over time for female spiny dogfish. Analyses could include but are not limited to assessing environmental drivers and harvest effects.

7. Coordinate a biological sampling program targeting spiny dogfish from additional locations and habitats outside those sampled by the NEFSC trawl surveys to understand the various factors that influence their life history (e.g., growth, maturity, fecundity)

8. Continue developing the VAST models presented to assess additional environmental variables that may influence abundance and distribution, and better predict the size composition for models that include multiple datasets.

9. Investigate datasets enumerating the abundance or diet of known spiny dogfish predators for comparison to natural mortality assumptions, and as potential proxies for dogfish natural mortality rates.

TOR8: BACKUP ASSESSMENT APPROACH

"Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment."

In the event the proposed assessment method fails peer review in the research track process or subsequently fails peer review in the routine management track process, the Working Group explored several assessment methods to serve contingency plans. The Working Group recommends that if the proposed assessment approach (SS3) does not pass peer review or is rejected in a future management track assessment, that the Stochastic Estimator approach be used in its place. The Stochastic Estimator has been used as the primary assessment tool for Atlantic spiny dogfish and has been previously considered sufficient for guiding catch advice. However, the Stochastic Estimator has many limitations that were noted previously, as it lacks the inclusion of life history information, and ability to use multiple fleets and surveys. If the Stochastic Estimator is considered a preferable modeling approach by the peer-reviewers, future application of the Stochastic Estimator should consider additional advancements to the data inputs that would address previous concerns. For example, future applications of the Stochastic Estimator should consider using model-based fisheries-independent indices that can integrate multiple surveys and address concerns regarding missed or incomplete sampling for a given year. Specifically, the Working Group recommends testing the VAST indices or a modified version presented as part of this assessment (Hansell and McManus 2022).

Other data-limited approaches were also evaluated as part of the research track. The Depletion-Corrected Average Catch (DCAC; MacCall 2009) model was applied to female Atlantic spiny dogfish to calculate a sustainable yield (Anstead 2022b). DCAC adjusts the average catch over the available time series based on an assumed depletion in the stock relative to its unfished biomass. Depletion-Based Stock Reduction Analysis (DB-SRA; Dick and MacCall 2011) was also implemented, which uses a flexible production model with a lumped biomass population dynamics model (Anstead 2022b). Both methods require similar input parameters (e.g., natural mortality, ratio of fishing mortality at maximum sustainable yield to natural mortality) and user-specified distributions (e.g., lognormal, uniform). Monte Carlo resampling is used to sample from the input parameters. DCAC recommends a sustainable yield whereas DB-SRA solves for the initial biomass that fits the specified inputs and, using the catch history, calculates catch limits and reference points. Catch advice for female spiny dogfish from both the DCAC and DB-SRA were consistent with each other (Figures 8.1 - 8.3), and both recommend a female harvest that is somewhat below the current coastwide total quota. When reviewing the methods, the Working Group believed the DCAC method provided more realistic catch advice over the DB-SRA method given it does not rely on a production function. However, the overall consensus of the Working Group was that since these methods ignore the size and age structure of the population, they did not provide a greater benefit over the Stochastic Estimator. Additionally, the biomass estimates and increasing trend from DB-SRA in recent years was not consistent with the results derived from SS3 and the Stochastic Estimator.

The Working Group also applied the Ismooth method (formerly known as the PlanB Smooth method) for Atlantic spiny dogfish to evaluate its performance (NEFSC 2020). The Ismooth method uses a LOESS-smoothed average index of abundance. A log linear regression on the last three years of the LOESS-smoothed index is conducted to derive the slope, which is used as a multiplier on recent catch to provide revised catch advice. This tool was evaluated as the backup model due to its performance during the Index Based Model Working Group (NEFSC 2020), and its use as the primary assessment model for stocks within the region. Simulation testing of the Ismooth method has indicated that it can be useful in the absence of an age-structured assessment depending on the given stock's biomass and exploitation in relation to its current status relative to its reference points (Legault et al. 2022). Application of the Ismooth highlights the decline that is observed in the standard NEFSC spring bottom trawl indices, and can be run to provide sex-specific catch multipliers (Figure 8.4). Given it also does not include the population dynamics or age information, the Working Group does not recommend this method as a contingency plan over the stochastic estimator.

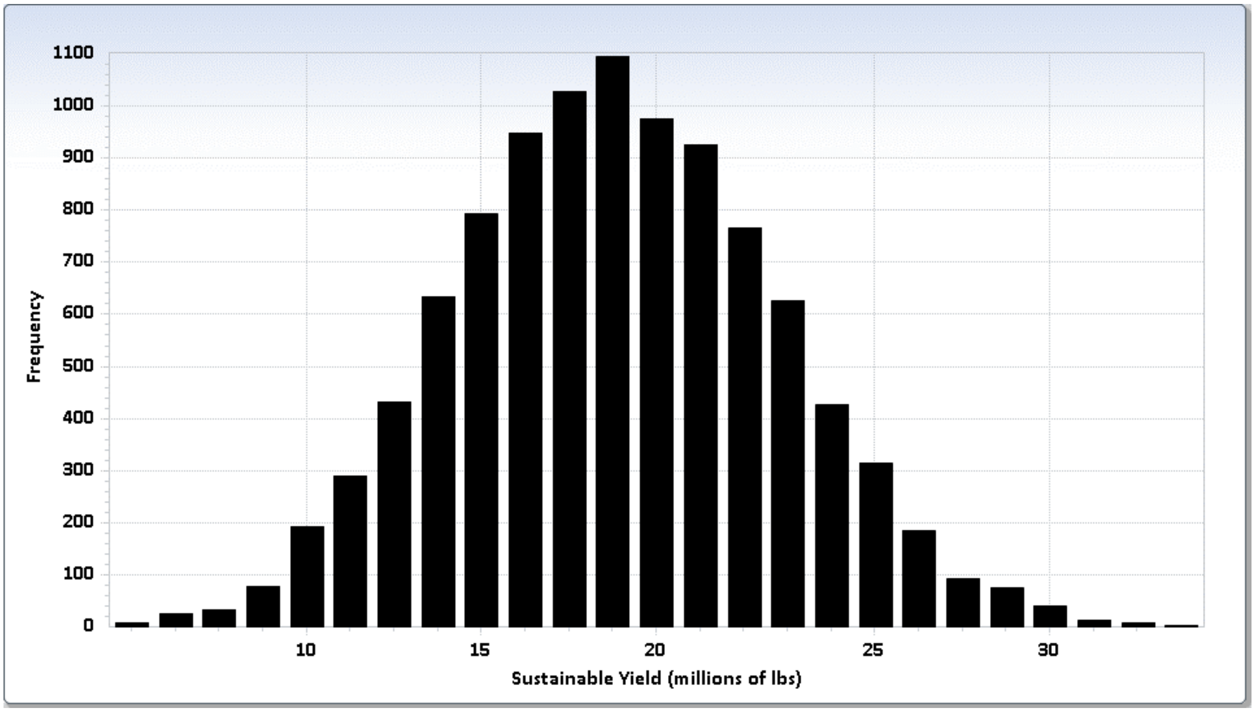


Figure 8.1 Distribution of sustainable yield estimates from the DCAC base configuration for female spiny dogfish.

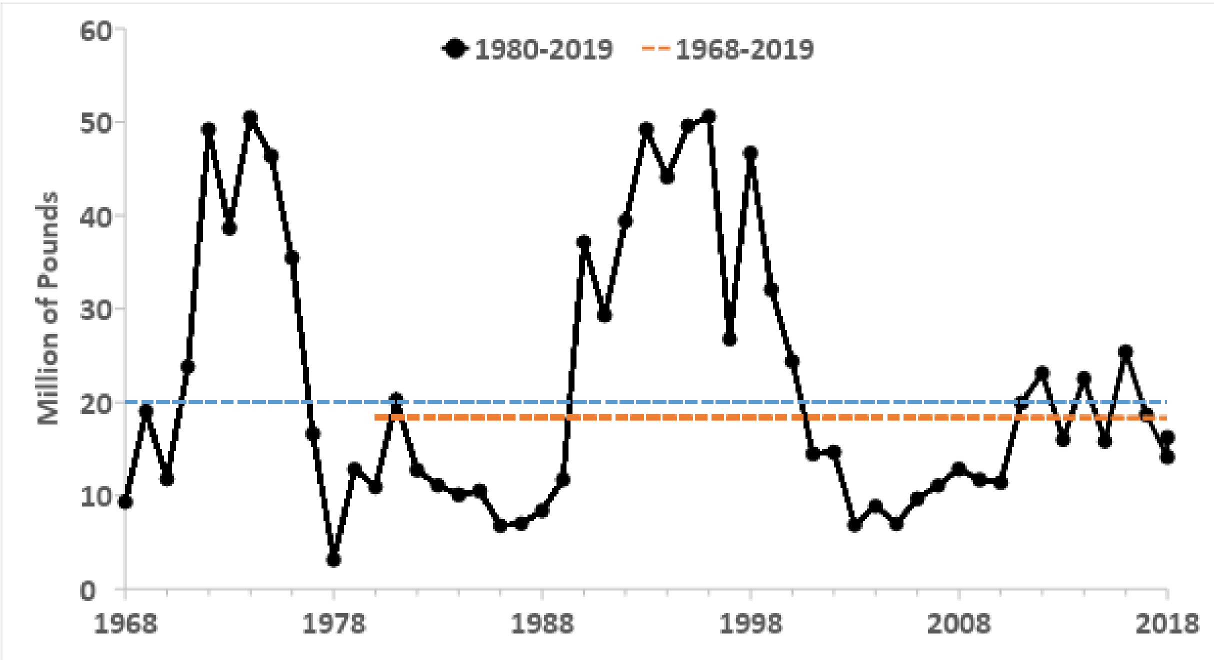


Figure 8.2. Female spiny dogfish removals in millions of pounds (black line) and the median sustainable yield estimate (18.36; dashed orange line) from the DCAC base configuration. A longer time series was also explored and the median sustainable yield from that sensitivity run was included (20.02; dashed blue line).

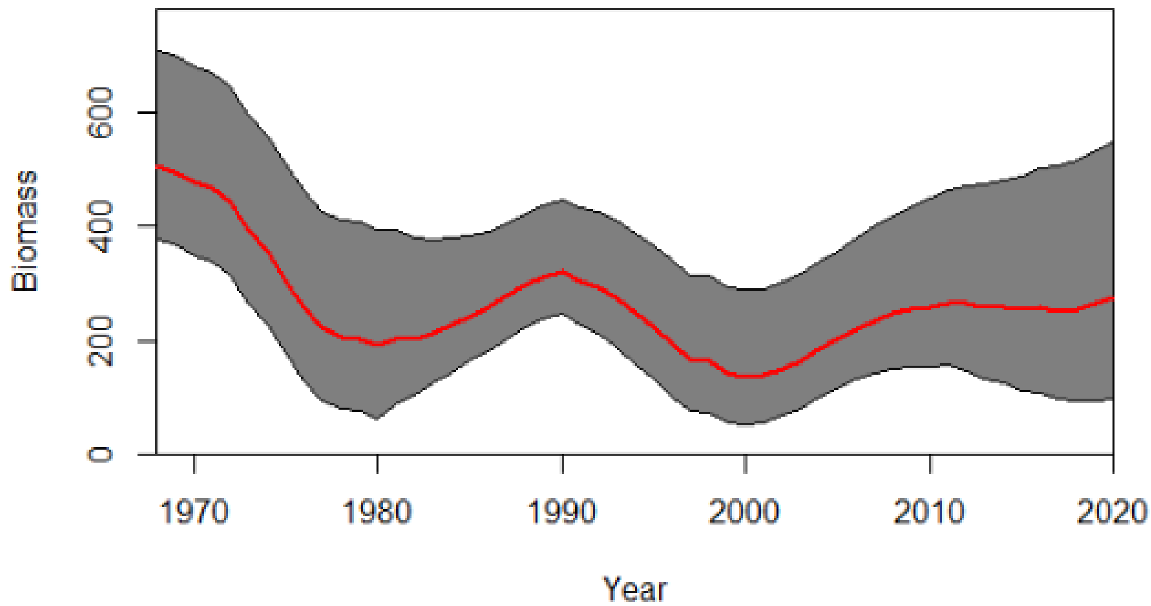


Figure 8.3. Estimated female spiny dogfish biomass (millions of pounds) from the DB-SRA model with 95% confidence intervals. Red line indicates the median biomass values.

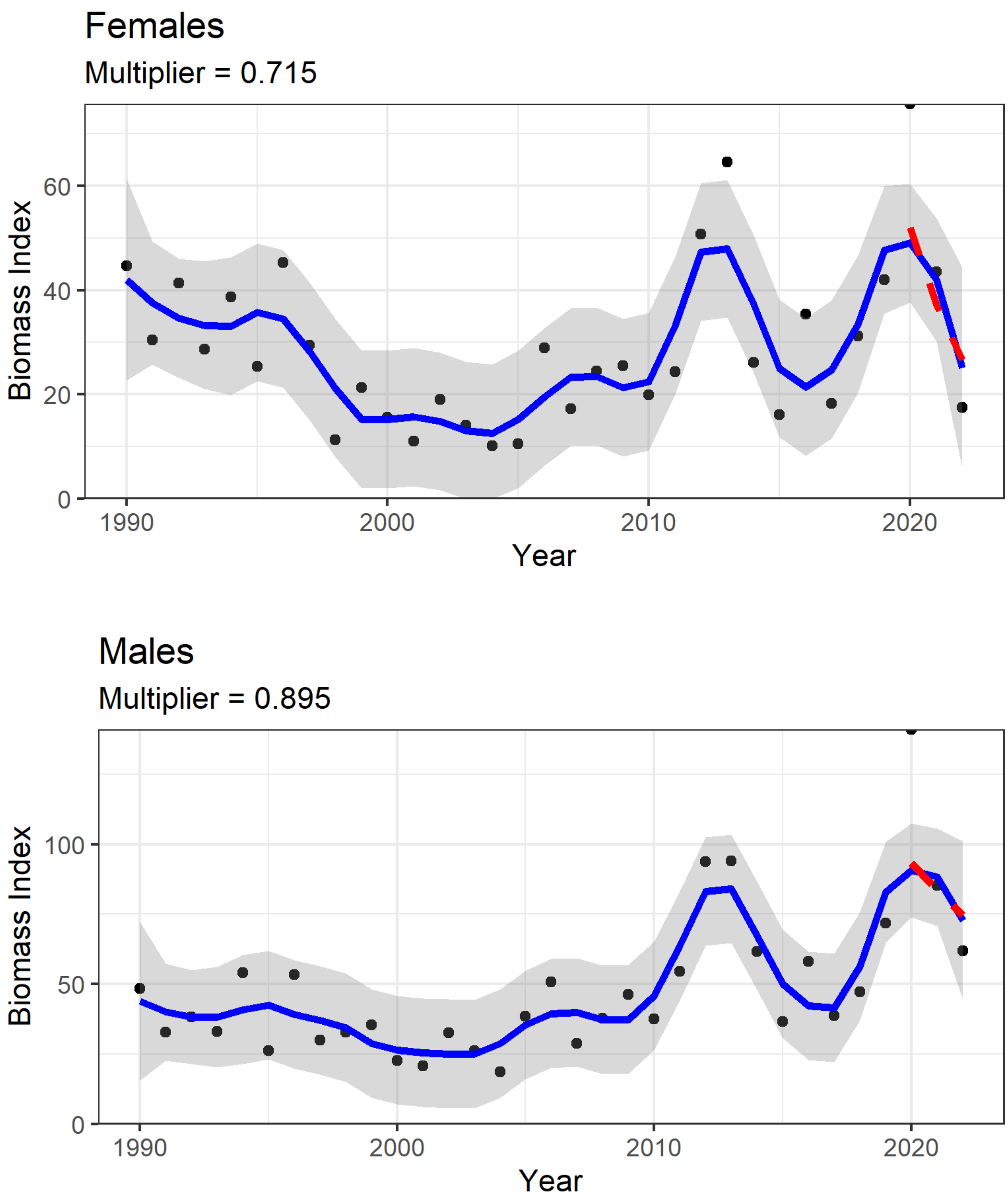


Figure 8.4. Female (top) and male (bottom) spiny dogfish NEFSC bottom trawl survey indices with the Ismooth approach applied. Loess fits (blue lines) and the associated confidence interval (gray areas) and the fit to the terminal years for deriving the multiplier (dashed red lines) are presented. Indices were derived from Stock SMART.

ACKNOWLEDGMENTS

This report reflects the work from a large group of people, including Working Group members and participants. Michele Traver and Alexander Dunn provided administrative support. Vladlena Gertseva and Ian Taylor provided insight into the development of the spiny dogfish SS3 models. Michelle Passerotti and scientists from the State of Washington provided technical and analytical support for ageing samples as part of this research track assessment. Andy Jones assisted with developing CPUE indices and spatial descriptions of catch information. Julie Neiland and Dave McElroy provided a synthesis of the NEFSC Bottom Long Line Survey. ASMFC Spiny Dogfish Technical Committee members provided information and guidance for consideration as part of the assessment.

REFERENCES

Alverson DL, Carney MJ. 1975. A graphic review of the growth and decay of population cohorts. ICES Journal of Marine Science 36(2): 133-143.

Anstead K. 2022a. Natural mortality estimates for spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Anstead K. 2022b. Two data poor methods applied to spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Anthony VC, Murawski SA. 1985. Rapporteur's Summary: First Northeast Fisheries Center Stock Assessment Workshop, 8-l2 July 1985, Falmouth, Massachusetts. 29 p.

Breidt J, Liu T, Opsomer JA. 2017. Small Area Estimation Approach for Reconciling Mode Differences in Two Surveys of Recreational Fishing Effort. Available at https://www.st.nmfs.noaa.gov/recreational-fisheries/MRIP/FES-Workshop/index.html.

Bubley WJ, Kneebone J, Sulikowski JA, Tsang PCW. 2012. Reassessment of spiny dogfish *Squalus acanthias* age and growth using vertebrae and dorsal-fin spines. Journal of Fish Biology 80(5): 1300-1319.

Burns TS, Shultz R, Bowman BE. 1983. The commercial catch sampling program in the northeastern United States. In: Doubleday, W.G. and Rivard, D. (eds.) Sampling commercial catches of marine fish and invertebrates. Canadian Special Publication of Fisheries and Aquatic Sciences 66:82-95.

Cadrin S, Frede R, Keiley E, Linton B, Maguire J-J, Rago P, Bell R, Giacalone V, Demarest C, Brown C, Gibson M, O'Keefe C, DeCelles D, Wright B, Hansell A, McGuire C, Hennen D. 2020. Fishery Data for Stock Assessment Working Group Report. New England Fishery Management Council Report.

Campana SE, Gibson JF, Marks L, Joyce W, Rulifson R, Dadswell M. 2007 Stock structure, life history, fishery and abundance indices for spiny dogfish (*Squalus acanthias*) in Atlantic Canada. Canadian Science Advisory Secretariat, Research Document 2007–2089.

Campana SE, Joyce W, Kulka DW. 2009. Growth and reproduction of spiny dogfish off the eastern coast of Canada, including inferences on stock structure. Biology and management of dogfish sharks: 195-208.

Carlson AE, Hoffmayer ER, Tribuzio CA, Sulikowski JA. 2014. The use of satellite tags to redefine movement patterns of spiny dogfish (*Squalus acanthias*) along the U.S. east coast: implications for fisheries management. PLoS ONE 9(7): e103384.

Carvalho F, Winker H, Courtney D, Kapur M, Kell L, Cardinale M, Schirripa M, Kitakado T, Yemane D, Piner KR, Maunder MN. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research 240: 105959.

Cao J, Thorson JT, Richards RA, Chen Y. 2017. Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. Canadian Journal of Fisheries and Aquatic Sciences 74: 1781-1793.

Chang J-H, Hart D, McManus MC. 2022. Stock synthesis for Atlantic spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Compagno LJV. 1984. FAO species catalogue. Vol. 4. Sharks of the world. Part 1.Food and Agriculture Program of the United Nations, Rome, Italy. p. 111-113.

Courtney D, Mathers A. 2019. An Updated Literature Review of Post-release Live-discard Mortality Rate Estimates in Sharks for use in SEDAR 65. SEDAR65-DW20. SEDAR, North Charleston, SC. 20 p.

Courtney D, Mathers A, Kroetz A. 2021. An Updated Literature Review of Post-Release Live-Discard Mortality Rate Estimates in Sharks for use in SEDAR 77. SEDAR77-DW25. SEDAR, North Charleston, SC. 47 p.

Dell'Apa A , Bangley CW, Rulifson RA. 2015. Who let the dogfish out? A review of management and socio-economic aspects of spiny dogfish fisheries. Reviews in Fish Biology and Fisheries 25: 273–295.

Dell'Apa A, Pennino MG, Bonzek C. 2017. Modeling the habitat distribution of spiny dogfish (*Squalus acanthias*), by sex, in coastal waters of the northeastern United States. Fisheries Bulletin 115: 89–10.

DFO [Fisheries and Oceans Canada]. 2020. Maritimes Research Vessel Survey Trends on the Scotian Shelf and Bay of Fundy. DFO Canadian Science Advisory Secretariat Science Research. 2020/019.

Dichmont, CM, Deng RA, Dowling N, Punt AE. 2021. Collating stock assessment packages to improve stock assessments. Fisheries Research 236: 105844.

Dichmont CM, Dengalt CM, Methot RD Jr, Porch CE, Prager MH. 2016. A review of stock assessment packages in the United States. Fisheries Research 183: 447-460.

Dick EJ, MacCall AD. 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110: 331-341.

Ebert DA, White WT, Goldman KJ, Compagno LJ, Daly-Engel TS, Ward RD. 2010. Resurrection and redescription of *Squalus suckleyi* (Girard, 1854) from the North Pacific, with comments on the *Squalus acanthias* subgroup (Squaliformes: Squalidae). Zootaxa, 2612: 22-40.

Fowler GM, Showell MA. 2009. Calibration of bottom trawl survey vessels: comparative fishing between the Alfred Needler and Teleost on the Scotian Shelf during the summer of 2005. Canadian Technical Report of Fisheries and Aquatic Sciences 2824: iv + 25 p.

Francis RC, Hilborn R. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68(6): 1124–1138. https://doi:10.1139/f2011-025

Friedland KD, Langan JA, Large SI, Selden RL, Link JS, Watson RA, Collie JS. 2020. Changes in higher trophic level productivity, diversity and niche space in a rapidly warming continental shelf ecosystem. Science of the Total Environment 704(20): 135270. https://doi.org/10.1016/j.scitotenv.2019.135270

Garrison LP. 2000. Spatial and dietary overlap in the Georges Bank groundfish community. Canadian Journal of Fisheries and Aquatic Sciences 57: 1679-1691.

Gertseva V, Taylor I, Wallace J, Matson SE. 2021. Status of the spiny dogfish shark resource off the continental U.S. Pacific Coast in 2021. Pacific Fishery Management Council, Portland, OR. Available from http://www.pcouncil.org/groundfish/stock-assessments/.

Haist V, Fowler M, Campana S. 2010. A Length-Based, Spatially Explicit Population Model for Spiny Dogfish (*Squalus acanthias*) in the Northwest Atlantic. TRAC Reference Document 2009/02.

Hansell A, McManus C. 2022. Spatio-temporal dynamics of spiny dogfish (*Squalus acanthias*) in US waters of the northwest Atlantic. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Hare JA, Morrison WE, Nelson MW, Stachura MM, Teeters EJ, Griffis RB, et al. 2016. A vulnerability assessment of fish and invertebrates to climate change on the northeast US continental shelf. PLoS ONE 11(2): e0146756. doi:10.1371/journal.pone.0146756.

Hart DR, Chang J-H. 2022. Per recruit modeling and reference points for spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Hart DR, Sosebee K. 2022. Length/Weight/Fecundity relationships for Atlantic spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Haugen JB, Curtis TH, Fernandes PG, Sosebee KA, Rago PJ. 2017. Sexual segregation of spiny dogfish (*Squalus acanthias*) off the northeastern United States: Implications for a male-directed fishery. Fisheries Research 193: 121-128.

Holden, MJ. 1965. The stocks of spurdogs (*Squalus acanthias* L.) in British waters, and their migrations. Fisheries Investigations. Series 2. Sea Fisheries. Ministry of Agriculture, Fisheries, and Food, Great Britain: 244: 20p.

Hurlbut T, Clay D. (eds) 1990. Protocols for Research Vessel Cruises within the Gulf Region (Demersal Fish) (1970-1987). Canadian Data Report of Fisheries and Aquatic Sciences 2082: 143p.

Hoenig JM. 1983. Empirical use of longevity data to estimate mortality rates. Fishery Bulletin 82: 898-903.

Jensen AL. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Canadian Journal of Fisheries and Aquatic Sciences 53: 820-822.

Johnston R, Sosebee K. 2014. History of the United States bottom trawl surveys, NAFO subareas 4–7. Northwest Atlantic Fisheries Organization SCR Document, 14-024.

Jones AW. 2022. Exploring vessel trip report and observer based fishery information for spiny dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Jones AW, Didden JT, McManus MC, Mercer AJ. 2022. Exploring commercial CPUE indices for the spiny dogfish in the northeast U.S. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Kerr L, Barajas M, Wiedenmann J. 2022. Coherence and potential drivers of stock assessment uncertainty in Northeast US groundfish stocks. ICES Journal of Marine Science 79: 2217–2230.

Legault CM, Restrepo VR. 1998. A flexible forward age-structured assessment program. ICCAT SCRS/1998/58.

Legault CM, Wiedenmann J, Deroba JJ, Fay G, Miller TJ, Brooks EN, Bell RJ, Langan JA, Cournane JM, Jones AW, Muffley B. 2022. Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. Canadian Journal of Fisheries and Aquatic Sciences e-First https://doi.org/10.1139/cjfas-2022-0045

Link, JS and Ford, MD. 2006. Widespread and persistent increase of Ctenophora in the continental shelf ecosystem off NE USA. Marine Ecology Progress Series 320:153-159.

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: 627-642.

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

Mandelman JW, Farrington MA. 2007a. The estimated short-term discard mortality of a trawled elasmobranch, the spiny dogfish (*Squalus acanthias*). Fisheries Research 83: 238–245.

Mandelman JM, Farrington MA. 2007b. The physiological status and mortality associated with otter-trawl capture, transport, and captivity of an exploited elasmobranch, *Squalus acanthias*. ICES Journal of Marine Science, 64: 122–130.

Maunder MN, Punt AE. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70(2-3): 141-159.

McAllister MK, Ianelli JN. 1997. Bayesian stock assessment using catch-age data and the sampling - importance-resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences 54: 284–300.

McCandless C. 2022. Preliminary spiny dogfish movements and growth estimates from NEFSC mark recapture data. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

McElroy WD, O'Brien L, Blaylock J, Martin MH, Rago PJ, Hoey JJ, Sheremet VA. 2019. Design, implementation, and results of a cooperative research Gulf of Maine longline survey. NOAA Technical Memorandum 2014-2017. NMFS-NE-249. 161p.

McManus MC, Sosebee K, Rago P. 2022. Biological reference points for spiny dogfish: Revisiting Rago and Sosebee (2010). Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Methot RD Jr, Wetzel CR. 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86–99.

Methot RD Jr, Taylor IG. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68(10): 1744-1760.

Methot RD, Wetzel CR, Taylor IG, Doering KJ. 2021. Stock Synthesis User Manual: Version 3.30.18.

Miller TJ, Das C, Politis PJ, Miller AS, Lucey SM, Legault CM, Brown RW, Rago PJ. 2010. Estimation of Albatross IV to Henry B. Bigelow calibration factors. Northeast Fisheries Science Center Reference Document 10-05.

Miller TJ, Brooks EN. 2021. Steepness is a slippery slope. Fish and Fisheries 22(3): 634-645.

Nammack MF, Musick JA, Colvocoresses JA. 1985. Life history of spiny dogfish off the Northeastern United States. Transactions of the American Fisheries Society 114(3): 367-376.

NEFSC [Northeast Fisheries Science Center]. 1990. Report of the Eleventh NEFC Stock Assessment Workshop. 121p.

NEFSC [Northeast Fisheries Science Center]. 1994. Report of the 18th Northeast Regional Stock Assessment Workshop: Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NOAA/NMFS/NEFSC: Woods Hole, MA. Northeast Fisheries Science Center Reference Document 94-22.

NEFSC [Northeast Fisheries Science Center]. 1998. Report of the 26th Northeast Regional Stock Assessment Workshop (26th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fisheries Science Center Reference Document 98-03; 283p.

NEFSC [Northeast Fisheries Science Center]. 2000. Report of the 30th Northeast Regional Stock Assessment Workshop: Stock Assessment Review Committee (SARC) Consensus

Summary of Assessments. NOAA/NMFS/NEFSC: Woods Hole, MA. Northeast Fisheries Science Center Reference Document No. 00–04.

NEFSC [Northeast Fisheries Science Center]. 2003. Report of the 37th Northeast Regional Stock Assessment Workshop (37th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fisheries Science Center Reference Document 03-16; 597p.

NEFSC [Northeast Fisheries Science Center]. 2006. 43rd Northeast Regional Stock Assessment Workshop (43rd SAW): 43rd SAW assessment report. Northeast Fisheries Science Center Reference Document 06-25; 400p.

NEFSC [Northeast Fisheries Science Center]. 2020. Draft report of the Index Based Methods Working Group, submitted for peer review November 20, 2020. 128p.

Nieland JL, McElroy WD. 2022. NEFSC Gulf of Maine Bottom Longline Survey Data and Analyses for Spiny Dogfish. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

NOAA Fisheries. 2022a. DisMAP data records. Retrieved from apps-st.fisheries.noaa.gov/ dismap/DisMAP.html. Accessed 9/8/2022.

NOAA Fisheries. 2022b. Fisheries Habitat in the Northeast U.S. Shelf Ecosystem. https:// www.fisheries.noaa.gov/new-england-mid-atlantic/ecosystems/fisheries-habitat-northeast-us-shelf-ecosystem#spiny-dogfish. Accessed 10/6/2022.

Nye JA, Link JS, Hare JA, Overholtz WJ. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Marine Ecology Progress Series 393: 111–129.

O’Brien L, Worcester T. 2010. Proceedings of the Transboundary Resources Assessment Committee (TRAC) Spiny Dogfish Review. Proceedings 2010/01. 57p.

Overholtz, WJ and Link JS. 2006. Consumption impacts by marine mammals, fish, and seabirds on the Gulf of Maine - Georges Bank Atlantic herring (*Clupea harengus*) complex during the years 1977-2002. ICES Journal of Marine Science, 64: 83–96.

Papacostas KJ, Foster J. 2021. The Marine Recreational Information Program. Survey Design and Statistical Methods for Estimation of Recreational Fisheries Catch and Effort.

Parkins C, Olszewski SD. 2021. Assessment of recreationally important finfish stocks in Rhode Island waters: Coastal Fishery Resource Assessment Trawl Survey 2021 Performance Report F-61-R Segment 21 Jobs 1 and 2. Rhode Island Department of Environmental Management Division of Marine Fisheries.

Passerotti MS, McCandless CT. 2022. Updated age and growth estimates for spiny dogfish *Squalus acanthias*. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Perretti CT, Fogarty MJ, Friedland KD, Hare JA, Lucey SM, McBride RS, Miller TJ, Morse RE, O'Brien L, Pereira JJ, Smith LA, Wuenschel MJ. 2017. Regime shifts in fish recruitment on the Northeast US Continental Shelf. Marine Ecology Progress Series 574: 1-11.

Punt AE, Dunn A, Elvarsson BÞ, Hampton J, Hoyle SD, Maunder MN, Methot RD, Nielsen A. 2020. Essential features of the next-generation integrated fisheries stock assessment package: a perspective. Fisheries Research 229: 105617.

Rago PJ, Sosebee KA, Brodziak JKT, Murawski SA, Anderson ED. 1998. Implications of recent increases in catches on the dynamics of Northwest Atlantic spiny dogfish (*Squalus acanthias*). Fisheries Research 39: 165-181.

Rago PJ, Wigley SE, Fogarty MJ. 2005. NEFSC Bycatch Estimation Methodology: Allocation, Precision, and Accuracy. NOAA/NMFS/NEFSC. Northeast Fisheries Science Center Reference Document 05-09.

Rago PJ, Sosebee KA. 2010. Biological Reference Points for Spiny Dogfish. Northeast Fish Sci Cent Ref Doc. 10-06; 52 p. Available from: National Marine Fisheries Service, 166 Water

Street, Woods Hole, MA 02543-1026, or online at: http://www.nefsc.noaa.gov/nefsc/publications/

Rideout RM, Ings DW. 2020. Temporal and spatial coverage of Canadian (Newfoundland and Labrador Region) Spring and Autumn Multi-Species RV Bottom Trawl Surveys, with an emphasis on surveys conducted in 2019. NAFO SCR Doc. 20/002.

Rulifson RA. 2007. Spiny Dogfish mortality induced by gill-net and trawl capture and tag and release. North American Journal of Fisheries Management 27:279–285.

Rulifson RA, Pratt M, Bell TJ, Parente I, Cudney-Burch J, Dell'Apa A. 2012. Is Cape Cod a natural delineation for migratory patterns in US and Canadian spiny dogfish stocks? Final Report for the Commercial Fisheries Research Foundation, Southern New England Collaborative Research Initiative, Saunderstown, RI. p 44.

Sagarese SR, Frisk MG, Miller TJ, Sosebee KA, Musick JA, Rago PJ. 2014a. Application of generalized additive models to examine ontogenetic and seasonal distributions of spiny dogfish (*Squalus acanthias*) in the Northeast (US) shelf large marine ecosystem. Canadian Journal of Fisheries and Aquatic Sciences 71: 847–877.

Sagarese SR, Frisk MG, Miller TJ, Sosebee KA, Musick JA, Rago PJ. 2014b. Influence of environmental, spatial, and ontogenetic variables on habitat selection and management of spiny dogfish in the Northeast (US) shelf large marine ecosystem. Canadian Journal of Fisheries and Aquatic Sciences 71: 567–580

Sagarese SR, Frisk MG, Cerrato RM, Sosebee KA, Musick JA, Rago PJ. 2015. Spatiotemporal overlap of spiny dogfish (*Squalus acanthias*) and commercial fisheries in the northeast U.S. shelf large marine ecosystem. Fishery Bulletin 113(2): 101-120.

Sagarese SR, Frisk MG, Cerrato RM, Sosebee KA, Musick JA, Rago PJ. 2016. Diel variations in survey catch rates and survey catchability of spiny dogfish and their pelagic prey in the Northeast U.S. Continental Shelf Large Marine Ecosystem. Marine and Coastal Fisheries 8:1, 244-262.

Shelton AO, Thorson JT, Ward EJ, Feist BE. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences 71(11): 1655-1666.

Sherman SA, Stepanek K, Sowles J. 2005. Maine - New Hampshire Inshore Groundfish Trawl Survey Procedures and Protocols. ME DMR Research Reference Document 05/01.

Sosebee KA. 2005. Are density-dependent effects on elasmobranch maturity possible? Journal of Northwest Atlantic Fishery Science 35: 115-124.

Sosebee KA, Jacobson L, Rago P, Traver M. 2010. Description of Data and Exploratory Stock Assessments of Spiny Dogfish, *Squalus acanthias*, using Stock Synthesis. TRAC Reference Document 2010/03.

Sosebee KA. 2019. Update of Landings, Discards and Survey Indices for Spiny Dogfish in 2018-2019. Report to the Mid Atlantic Fishery Management Council Scientific and Statistical Committee 2019: 1-31.

Sosebee KA. 2022a. Maturity of spiny dogfish in US waters from 1998-2021. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Sosebee KA. 2022b. Spiny dogfish catch summary and derivation of catch at length and sex. Working Paper Submitted as part of the 2022 Spiny Dogfish Research Track Assessment.

Stehlik LL. 2007. Essential fish habitat source document. Spiny dogfish, *Squalus acanthias*, life history and habitat characteristics. NOAA technical memorandum NMFS-NE-203. 52p.

Stock BC, Miller TJ. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research 240: 105967.

Stone HH, Gross WE. 2012. Review of the Georges Bank Research Vessel Survey Program, 1987-2011. Canadian Technical Report of Fisheries and Aquatic Sciences 2988: xiii + 95p.

Taylor IG, Gertseva V, Methot Jr RD, Maunder MN. 2013. A stock–recruitment relationship based on pre-recruit survival, illustrated with application to spiny dogfish shark. Fisheries Research 142:15-21.

Templeman W. 1976. Transatlantic migrations of spiny dogfish (*Squalus acanthias*). Journal of the Fisheries Research Board of Canada 33:2605-2609.

Terceiro M. 2003. Stock assessment of summer flounder for 2003. Northeast Fisheries Science Center Reference Document 03-09; 179 p.

Then AY, Hoenig JM, Hall NG, Hewitt DA, and handling editor: Jardim E. 2015. 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.

Thorson JT, Shelton AO, Ward EJ, Skaug HJ. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science 72: 1297–1310.

Thorson, JT, Johnson KF, Methot RD, Taylor IG. 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research 192: 84-93.

TRAC [Transboundary Resources Assessment Committee]. 2010. Northwest Atlantic Spiny Dogfish. TRAC Status Report 2010/02.

Van Voorhees DA, Witzig JA, Osborn MF, Holliday MC, Essig RJ. 1992. Marine recreational fishery statistics survey, Atlantic and Gulf coasts, 1990-1991. National Marine Fisheries Service Current Fisheries Statistics Number 9204, 275 p. Silver Spring, MD, USA.

Verissimo A, McDowell JR, Graves JE. 2010. Global population structure of the spiny dogfish *Squalus acanthias*, a temperate shark with an antitropical distribution. Molecular Ecology 19: 1651–1662.

Wigley SE, McBride HM, McHugh NJ. 2003. Length-weight relationships for 74 fish species collected during NEFSC research vessel bottom trawl surveys, 1992-99.

Wigley SE, Rago PJ, Sosebee KA, Palka DL. 2007. The Analytic component to the standardized bycatch reporting methodology omnibus amendment: sampling design, and estimation of precision and accuracy (Second Edition) NOAA/NMFS/NEFSC. Northeast Fisheries Science Center Reference Document 07-09.

APPENDIX A. STAKEHOLDER SESSION MEETING NOTES

Stakeholder Session Meeting Summary
Tuesday, February 15, 2022
3:02PM-5:06PM

*Please note that **public questions/comments/input is bolded** to distinguish from working group presentation summaries and responses*

An introduction to the research track and terms of reference

Presenter: Conor McManus-RIDEM
Provided background on the Spiny Dogfish Research Track Stock Assessment process and Work Group.

Discussion summary: No discussion

The assessment model

Presenter: Jui-Han Chang-NEFSC
Provided background on Stock Synthesis 3 (SS3), a stock assessment-modeling program, which will be evaluated for use in the stock assessment.

Discussion summary

- **What will be the terminal year for the research track?**
    - As of now, the assessment will run through 2019 as 2020-2021 catch data is still under analysis.
    - Terminal year is considered so that GARFO and NEFSC will use the same set of data.

- **Clarification of the current working model and the final model**
    - The current working model is not an SS3 model.
    - As part of this research track assessment, the WG has developed a SS3 model for spiny dogfish using data through 2008, but will be updating data through 2019 for the assessment. At this point the WG will evaluate the model's performance.

Ecosystem drivers and influences

Presenter: Alex Hansell-NEFSC
Provided background and efforts addressing Terms of Reference 1 regarding environmental drivers for the stock.

Discussion summary

- Ecological and/or climate influences on abundance and the effects of environmental influences on catch?
    - **The water is now warmer, seal population is increasing which could be influencing numbers.**

- **Are seals eating dogfish and if so, how does this impact the dogfish population? The relationship between dogfish and groundfish needs to be considered.**

- Shifts in spatial distribution over time
  - **Fishers must travel further offshore to encounter and get the same biomass as in the past**
  - **Pockets of no fish up to 20 miles offshore**
  - **Travel distance has increased every year.**
  - **In 2008-2010, trip limits were 2000lbs and they used smaller skiffs. These limits were easier to fill the boat, gaffing the fish from the boat a half mile from the beach. Now traveling 10-20 miles offshore to catch that same biomass. Fishers are traveling much greater distances offshore. Within that subset of time, within 3-4 years ago they had state permits within state waters, sometime in 2015-2016, fishers were required to go outside state waters and with each subsequent year, fishers travel distance increased. This previous year was extra challenging because the fish were not even offshore at the fishing grounds. The fish are pushing offshore, and they seem to have pockets of no fish whatsoever. Groups of fishers would go up and down the coast searching for spiny dogfish.**
  - **Travel distance has increased every year; as of 3 years ago "no fish inshore", especially this past year 0 catch.**

- Timing of location over time
  - **In 2014-2015 they were reliable to predict, at the end of June spiny dogfish would move inshore and into November, they would move offshore again. The start and end dates have moved "closer". Instead of June moving inshore, they were moving inshore in July, the catchability window of inshore dogfish has shortened.**
  - **The time frame from maxing out on trip limit, the shoulders are getting larger where folks aren't maxing out, July-August guys are maxing out. The time is there but the biomass isn't in the Gulf of Maine.**
  - **Because the start and end of the availability has moved closer together, there is a shorter window of economic viability.**

- Changes in the size of fish or sex ratios
  - **This year, spiny dogfish were smaller, and it was more difficult to catch big fish.**
  - **Lower yields because fishermen couldn't get out of the small fish, the "art" of fishing is finding the big fish within smaller fish schools.**
  - **This year was one of the first that females were in the 4-5lb range (medium large females) something they rarely see. Normally females are 6+lbs.**
  - **Four years ago gear switched from 6.5" to 7" gear, this year he couldn't catch a fish because they all swam through the mesh.**

- Prey

- **Squid, sand eels, and herring sometimes**

Movement ecology as related to tagging data

Presenter: Cami McCandless-NEFSC

Provided background on all tagging data that is being reviewed and/or analyzed by the Working Group.

Discussion summary

- **Why is a recapture rate of 3% considered good for sharks? Is this from tag retention or a catchability index?**
  - It is probably a little bit of both.
  - Catchability: the size of the ocean and their highly migratory spatial dynamics means few are encountered and of those not all are reported
  - The highest recapture rate is about 13% (shortfin mako) and 9% (blue).

- **We release spiny dogfish without tagging them with the NMFS M-dart tags from the Cooperative Shark Tagging Program. It looks like you're not in strong need for tagging of spiny dogfish with these tags. Is that a correct interpretation?**
  - That is correct. Spiny dogfish are typically tagged with the roto tags instead of the dart tags due to concerns with tag retention. [Coopertative Shark Tagging Program M-dart tags are designed for sharks 3 feet and larger]
  - **Over the years 42,000 spiny dogs tagged with Stainless Steel single barbed shark tags from floy tag company, and we have had tag retention for decades. One of the reasons we didn't use roto tags was we were concerned about the catch in gillnets. The other thing is that we saw most of our tag returns coming from the inner continental shelf where most of the fishery is taking place, most of those came within a year, but there was something very strange, we started getting tags returned 10 years later, those had been released from NC waters and showing up in MA fisheries 10 years later, I think they went offshore from any harvest and stayed there for a long time before they returned to the continental shelf. I haven't published that but that's what we are seeing in the data.**

- **How is the study fleet data being used?**
  - Vessels used for tagging were part of the study fleet.
  - Tagging data is not regularly received from the study fleet currently, only a handful a year. When tags were originally put out there was a high return.
  - NOAA is investigating further/increased use of the study fleet data for answering directed questions.

- **Where do the off the shelf tag returns come from?**
  - All but 1 are summer recaptures.
  - **There was one recapture of a stainless-steel barbed tag from Iceland.**
  - **We see more male fish(undesirable) every summer off cape cod and less large females. We also see a lot of mixed sized fish of both sexes.**

Survey and catch information

Presenter: Kathy Sosebee-NEFSC

Provided background on the NEFSC Bottom Trawl Survey data being used in the assessment currently, and reviewed the other survey data that will be evaluated as part of the assessment. Landings information was also reviewed.

Discussion summary

- **Where does the catch data come from?**
  - Refer to vessel/trip reports to derive landings information.

- **Are the surveys indicating that numbers are going down?**
  - The trawl survey numbers are down slightly with 2017 being a strange year. The numbers are not down to the extent they were 20 years ago. Currently, 2020 and 2021 are not included.

- **What is the method and format of the recreational landing survey?**
  - The method has changed overtime, used to be a combination of phone survey and mail survey. Now it is a mail survey combined with an intercept survey where they get information, just not many spiny dogfish encountered.
  - The success of the survey has been reviewed by the National Academy of Sciences.
  - Participation in the intercept survey is high and the mail survey is higher than the phone survey. Jason Didden can be contacted for more information (jdidden@mafmc.org).

- **Is catch taking into account effort metrics? Seeing a decline in landing component but Gulf of Maine probably ever contributed much in the first place, but the effort has decreased? Is this being accounted for? Decreases in catch from Gulf of Maine could lead to erroneous conclusions if decrease in effort is not considered.**
  - CPUE issue when there is only one processor which is in an early stock assessment report; unsure if CPUE can be looked at if there is only 1 processor; the Working Group will investigate further.
  - CPUE must account for changes in the number of participants.

- **What is the rush in raising the trip limits? Concern that raising trip limits is a push from processors and could negatively impact stock size and push mid-size vessels out of the fishery. Industry participation has already declined significantly.**
  - The trip limit is a management question rather than a question that will be addressed as part of the assessment. The working group does not have final estimates or inferences on current stock size. When the assessment is complete with updated modeling, we will have a better understanding of the stock size, and assuming it passes peer-review, will be available for future management considerations.

    - In terms of current management, GARFO is working on trip limit revisions for May of 2022. The reason we are changing this year is from industry requests and looking at the data. Trip limits limit the people's ability to run their operations as they would like to, and it's causing many discards and they aren't able to meet the quotas they have by the end of the year. It's not a significant increase either, so it's largely by industry request. It's related but not directly involved with the research assessment.

- **What efforts have been made to increase the value of dogfish?**
    - Several projects have been funded to increase dogfish market value.
    - Varying degrees of success in these efforts.

Closing

The next Working Group meeting is scheduled for February 24 from 10-1PM. The link to the meeting can be found on the assessment's homepage: https://www.fisheries.noaa.gov/event/research-track-working-group-2022-improving-assessments-spiny-dogfish.

2022 BLUEFISH RESEARCH TRACK ASSESSMENT

PARTICIPANTS

Working Group

NAME	AFFILIATION
Michael Celestino	NJFW
Karson Cisneros	MAFMC
Katie Drew	ASMFC
Sam Truesdell	MADMF
Abigail Tyrell	NEFSC/Ocean Associates Inc.
Jessica Valenti	NOAA/Rutgers
Samantha Werner	NEFSC
Tony Wood	NEFSC

Chair-invited analytical participants

Sarah Gaichas	NEFSC
Jim Gartland	VIMS
Tim Miller	NEFSC
Joe Myers	ACCSP

Working Group meeting attendees

Alan Bianchi	NC-DEQ
Greg DiDomenico	Lund's Fisheries
Cynthia Ferrio	GARFO
James Fletcher	Unk
Jesse Hornstein	NYDEC
Nathan Jackson	Unk
Cynthia Jones	ODU
Mike Waine	American Sportfishing Association

ACKNOWLEDGEMENTS

The Working Group thanks: Sarah Gaichas and her collaborators on development of the forage fish index including Jim Gartland, Brian Smith, Elizabeth Ng, James Thorson. We thank Tim Miller for assistance with WHAM exploration, development, and support. We also thank Joe Myers, ASMFC's Bluefish Technical Committee and Fishery Management Plan Coordinator, Dustin Colson Leaning, for their data collection and analysis support as well as participating in the ASMFC State Data Review Workshop for bluefish. Michele Traver and Alexander Dunn provided helpful administrative support. Matt Seeley participated in the early WG meetings before moving on to the Western Pacific Fishery Management Council; José Moñtanez provided helpful interim support after M Seeley's departure. Eric Robillard contributed important information on bluefish ageing and oversaw or conducted the majority of the bluefish ageing for this assessment. Lou Carr-Harris assisted with development of a research recommendation.

EXECUTIVE SUMMARY

Term of Reference (TOR) #1: Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

Temperature and photoperiod are the principal factors directing activity, migrations, and distribution of adult bluefish. Based on this mechanistic connection, quantitative indicators of optimal temperature were developed to better understand temperature trends during the bluefish spawning season. Sources of uncertainty are discussed. Analyses suggested that the spawning season may now extend later in the year compared to historical periods, though it is unclear how these changes in potential spawning season may affect bluefish recruitment. On the other hand, the amount of habitat in the optimal temperature range during the peak spawning month of July has not changed over time, indicating stability in spawning conditions and therefore possibly also in recruitment. A Vector Autoregressive Spatiotemporal (VAST) model was developed from the fall NEFSC bottom trawl survey to determine the fall centers of gravity of three bluefish size groups over time; analyses suggested systematic trends in large and medium bluefish, but not small bluefish. Temperature was tested as a covariate in the VAST model, but resulting poor model diagnostics were beyond the scope of the present working group to address.

Using a VAST framework, we also developed a forage fish index to evaluate changes in bluefish prey over time and space that could be used to inform survey and/or fishery availability in the bluefish stock assessment to inform annual deviations in catchability. Small pelagic forage species are difficult to survey directly, so we developed a novel method of assessing small pelagic fish aggregate abundance using predator diet data. The forage fish indices based on fall, spring, and annual datasets all show fluctuations in forage fish biomass, alternating between multiple years or decades with higher and lower levels.

Variability in bluefish life history processes was modeled by splitting life history data by semesters of the year, by decade, by geographic region, and by sex; results and sources of uncertainty are discussed. Natural mortality was updated for this assessment from one based on a “rule of thumb” estimate of 0.2 for all ages to Lorenzen weight-based age-varying estimates. Our findings were considered and/or incorporated into several subsequent TORs, including: spatial domain of the stock (TOR2), estimates of seasonal and regional catch weights (TOR2), development of survey indices of abundance with environmental covariates (TOR3), incorporation of the forage fish index into a companion assessment model (TOR4), updating natural mortality for use in the assessment model (TOR4), and informed several research recommendations (TOR7).

Term of Reference #2: Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

The majority of commercial landings over the time series (1950-present) have been taken in the Mid-Atlantic region (New York, New Jersey, and North Carolina). The majority of recreational activity occurred from May to October, with specific seasonal patterns varying by state.

Recreational offshore (3-miles, or 4.8-km, or more from shore) areas account for only about 7% of total catch.

Total bluefish removals (total dead catch) have declined since the beginning of the time series. There was a slow increase from 1996 to 2010, but the declining trend has continued to the lowest values in the time-series in recent years. On average, commercial landings account for 14% of the total removals with commercial discards averaging only 0.2%. Dead commercial discards have not contributed to total removals in previous assessments, but since they have been identified as a source of uncertainty, they were included in this assessment. Total removals are dominated by the recreational fishery with recreational landings accounting for 71% of total removals, and recreational dead releases averaging 15% of total removals. The recreational dead release mortality rate was updated for this assessment through reexamination of the methods used in the previous assessment, and an updated literature review; the value changed from 15% to 9.4%. The recreational dead discard component of the catch was calculated using the season/region length frequency distributions developed from all of the recreational biological sampling data for released fish; this is a change from previous assessments to account for regional differences in fish size.

> *Term of Reference #3: Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.*

The WG participated in an ASMFC Bluefish Technical Committee workshop to review available state datasets. The WG explored standardizing fishery independent indices of abundance using environmental covariates in a GLM framework. However, the standardization process did not notably affect index trends or reduce interannual variability or index coefficients of variation, so the WG did not use the standardized indices in the base run and instead used the stratified arithmetic mean for surveys with a stratified random design and the geometric mean for surveys with a fixed station design. Bayesian hierarchical modeling was used to combine YOY indices into a single composite index, using the method developed by Conn (2010) that represents the coast wide recruitment dynamics of bluefish. Surveys included in the composite index were from NH Juvenile Finfish Seine Survey, RI Narragansett Bay Juvenile Finfish Beach Seine Survey, NY Western Long Island Seine Survey, NJ Delaware River Seine Survey, MD Juvenile Striped Bass Seine Survey, and VIMS Juvenile Striped Bass Seine Survey. In addition, the bluefish working group decided on 8 additional representative indices of bluefish abundance for the assessment:

1. NEFSC Fall inshore strata: 1985-2008 (age-0 – age-6+)
2. NEFSC Fall outer inshore strata (FSV Bigelow): 2009-2021 (age-0 – age-6+)
3. NEAMAP Fall Inshore trawl survey: 2007-2021 (age-0 – age-6+)
4. ChesMMAP trawl survey: 2002-2018 (age-0-3)
5. Pamlico Sound Independent Gillnet Survey; 2001-2021 (age-0 – 6+)
6. Marine Recreational Information Program CPUE: 1985-2021 (age-0 – age-6+)
7. SEAMAP Spring Inshore trawl survey: 1989-2021 (age-1)
8. SEAMAP Fall Inshore trawl survey: 1989-2021 (age-0)

Calculation of the MRIP CPUE was updated for this assessment. Bluefish trips were defined using a guild approach where a trip was considered a bluefish trip if it caught either bluefish or a species that was significantly positively associated with bluefish. This was a change from the previous benchmark assessment where effort was described using "directed trips," which describe trips where bluefish were considered a target species.

Multinomial age length keys were also explored as part of this assessment. Seasonal multinomial age length keys (ALKs) reduced retrospective trends and improved convergence diagnostics in statistical catch at age models relative to alternative ALKs; additionally, the WG did not believe data were sufficient for higher resolution (e.g., regional) ALKs, and so seasonal multinomial ALKs were selected for use in the assessment.

> *Term of Reference #4: Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.*

The Woods Hole Assessment Model (WHAM), a state-space, age structured stock assessment model, was used as the base model to estimate annual fishing mortality, recruitment, stock biomass, and associated estimates of uncertainty, with data updated through 2021. A suite of model fit diagnostic plots were examined for each model of interest and model fits were examined using conventional residual diagnostics, as well as one-step ahead residual diagnostics. Retrospective patterns in model results were evaluated using Mohn's rho values.

The final model configuration included a number of notable model and data changes since the previous peer reviewed model, including: a state-space model, updated natural mortality estimate, addition of new indices, including a newly estimated MRIP CPUE index, and addition of several selectivity blocks. Spawning stock biomass from the final base model starts in 1985 high and declines through the late 1990s, remains stable for several years before rising to a localized peak in 2008, declining through 2018, and rising in the years since. This pattern broadly reflects trends from the previously accepted model, albeit with differences in scale. Fishing mortality from the base model starts low in 1985 and rises quickly, then declines and varies without trend over much of the timeseries; fishing mortality reached a high in 2017, and has declined to timeseries lows since. The trend from the previously accepted model is broadly similar, albeit again, with some differences in scale, primarily in estimates of recruitment.

WHAM allows for incorporation of environmental covariates on the catchability of survey indices, and we explored a companion model that leveraged this capability. The companion model that used the forage fish index as a covariate on catchability of the MRIP index showed promise for continued development. The covariate led to an overall decreasing trend in catchability over time.

Term of Reference #5: Update or redefine status determination criteria (SDC; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

Existing status determination criteria from the 2021 management track assessment (data through 2019) were F_{MSY} proxy = $F_{35\%}$ = 0.181 and SSB_{MSY} = 201,729 MT (1/2 SSB_{MSY} = $SSB_{THRESHOLD}$ = 100,865 MT). Updated reference points from the ASAP continuity run are F_{MSY} proxy = $F_{35\%}$ = 0.176 and SSB_{MSY} = 190,771 MT (1/2 SSB_{MSY} = $SSB_{THRESHOLD}$ = 93,386 MT).

Both $F_{35\%}$ and $SSB_{35\%}$ were calculated in WHAM using average recruitment over the time series (1985-2021), and 5-year averages for fishery selectivity, maturity and weights-at-age for SSB per recruit calculations. Reference points from the final model (BF28W_m7) were F_{MSY} proxy = $F_{35\%}$ = 0.248 (95% CI: 0.209 – 0.299) and SSB_{MSY} proxy = $SSB_{35\%}$ = 91,897 MT (95% CI: 66,219–127,534 MT); $SSB_{THRESHOLD}$ =1/2 SSB_{MSY} proxy = 45,949 MT (95% CI: 33,110–66,768 MT). The retrospectively adjusted values of terminal year F and SSB were within the 90% confidence bounds of the unadjusted values, indicating a retrospective adjustment was not necessary to determine stock status. The terminal year SSB was 55,344 MT (95% CI: 35,185 – 87,052 MT) which was above the $SSB_{THRESHOLD}$ and 60% of SSB_{MSY}. Full fishing mortality was 0.166 (95% CI: 0.103 – 0.268) in 2021, which was 67% of the $F_{35\%}$ reference point. Stock status determination based on the final model indicates that there is an 87% chance that the bluefish stock is currently not overfished and over-fishing is not occurring.

Status determination criteria	2021 Management track assessment	2022 research track assessment (continuity run)	2022 research track assessment (WHAM)
F_{MSY} proxy = $F_{35\%}$	0.181	0.176	0.248
SSB_{MSY}	201,729 MT	190,771 MT	91,897 MT
½ SSB_{MSY}	100,865 MT	93,386 MT	45,949 MT

Term of Reference #6: Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

Short-term projections were conducted in WHAM, and incorporated model uncertainty and auto-regressive processes in recruitment and numbers-at-age. The projections used 5-year averages for natural mortality, maturity, fishery selectivity and weights-at-age. Removals in 2022 were assumed to be equal to the 2022 ABC (11,460 MT), and projections were carried forward for years 2023-2025 with different fishing mortality and harvest assumptions: F = 0, $F_{status\ quo}$ = 0.166, $F_{35\%}$ = 0.248, and that harvest in each year is equal to the acceptable biological catch (ABC) in each year. The probability of SSB in 2025 being above the SSB threshold is > 80% for all scenarios explored. Catch advice will be updated as part of the 2023 Management Track assessment, but catch advice from WHAM under the most likely scenario explored for this research track assessment (MAFMC risk policy assuming CV = 100%) is expected to be stable, but lower, relative to 2022.

> *Term of Reference #7: Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 1 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.*

The SAW 60 WG reviewed the status of previous research recommendations and proposed new ones to address issues raised during WG meetings. Notable accomplishments relative to past research recommendations include: development of an MRIP index using a species-association method to identify bluefish trips, updating the estimate of natural mortality used in the assessment model, evaluating model results that aggregated all model input data at a seasonal and regional level of resolution, multiple fishery independent surveys were combined using VAST as part of this assessment, examination of differences in the calibrated and uncalibrated MRIP estimates of bluefish catch, spatial stratification of recreational release length frequencies when calculating the weight of dead recreational releases, and the migration to the WHAM framework will allow for continued exploration and testing of covariates influencing time-varying catchability and selectivity.

The WG proposed several new research recommendations to better understand bluefish dynamics and assessing the population through the current or future models. These include the following: expand collection of recreational release length frequency data, continue development and refinement of the forage fish / availability index as well as incorporation of this index in to a base model for bluefish management advice, initiate additional fisheries-independent surveys or fishery-dependent sampling programs to provide information on larger, older bluefish, continue coastwide collection of length and age samples from fishery-independent and -dependent sources, refinement and development of indices of abundance, and develop a recreational demand model.

> *Term of Reference #8: Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.*

A backup assessment approach is required to be in place as a hedge against a scenario where the primary catch-at-age model is not suitable for providing management advice. The bluefish Working Group chose the index-based method Ismooth (previously known as PlanBSmooth) as the backup model due to its performance in the analyses performed by the Index Based Model Working Group (NEFSC 2020) and because it has a history of application at the NEFSC as an approach that has been used to develop ABCs (e.g., Georges Bank cod, Gulf of Maine / Northern

Georges Bank and Southern Georges Bank / Mid-Atlantic monkfish). Briefly, this approach applies recent trends in an index or indices to recent dead catch to generate ABC advice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 ECOSYSTEM AND CLIMATE INFLUENCES

> *Term of Reference (TOR) #1: Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.*

1.1 Ecosystem and Socioeconomic Profile (ESP)

An Ecosystem and Socioeconomic Profile (ESP) was used as a framework to address TOR1 in this research track assessment. ESP is a standardized framework to facilitate the inclusion of ecosystem and socioeconomic information in the stock advice process; it leverages existing information to understand the ecological and socioeconomic drivers of stock dynamics and to incorporate this diverse information into the stock advice process through the creation of ecosystem and socioeconomic indicators. This standard framework also facilitates the interpretation of data and allows future working groups to update the existing indicators in addition to creating and assessing new indicators.

The ESP process begins with (1) a systematic review of existing ecosystem and socioeconomic literature and identification of problem statements for the stock, followed by (2) development of conceptual models to outline the major drivers on the stock, (3) creation of indicators relevant to stock performance, (4) analysis of select indicators, and, lastly, (5) reporting out scientific advice. The scientific advice provided by an ESP can inform the stock assessment in multiple ways, ranging from providing additional context and research recommendations, to suggesting new covariates to include that can inform dynamic processes within assessment modeling.

The bluefish ESP includes a comprehensive literature review of bluefish life history and related ecosystem considerations relating to bluefish habitat, distribution, diet, predators, competitors, growth, and survival at each life stage. It also served as a review of the history of the bluefish stock assessment and relevant biological information that is used to make decisions relating to the assessment modeling. A conceptual model identifying the major drivers for different life stages of bluefish was developed from this review (Figure 1). Diet data collected from multiple scientific surveys were analyzed to determine the major prey and predators of bluefish, supplementing the literature review on this topic with the most recent data. Distributional and environmental data from multiple state and federal surveys were analyzed to understand where, when, and under what conditions bluefish of different life stages and size classes were found. Ecosystem and socioeconomic indicators were developed to better understand the current status of bluefish in the context of each of these dimensions, as well as to begin to probe potential mechanistic linkages between the environment and the status of the bluefish stock. Relevant results are summarized below; see Working Paper 1 (Tyrell et al. 2022) for the detailed report.

1.1.1 Diet

In the Northeast Fisheries Science Center (NEFSC) bottom trawl data, anchovies, butterfish, and squid were important prey items in all years. Sandlance, herring, bluefish, scup, and drum were important prey in some years in the NEFSC bottom trawl. Bay anchovy, butterfish, and striped anchovy were important prey species in the Northeast Area Monitoring and Assessment Program (NEAMAP) bottom trawl. Bay anchovy, spot, and menhaden were important prey species in the

Chesapeake Bay Multispecies Monitoring & Assessment Program (ChesMMAP) bottom trawl. Overall, >80% of bluefish diet was composed of fish, both by weight and by abundance. There were few records of bluefish in the stomachs of other species captured and sampled in these surveys.

1.1.2 Environment, Spatial Distribution, and Cohorts

Adult and juvenile bluefish are found primarily in waters less than 20 meters deep along the Atlantic coast (Shepherd and Packer 2006). The 2022 bluefish research track assessment Working Group (referred to herein as WG) investigated whether Gulf of Mexico bluefish were part of the unit stock being assessed for the 2022 research track assessment. This investigation did not identify any known systematic studies (e.g., tagging, genetics) that demonstrated bluefish migrations into or out of the Gulf of Mexico. A review of the Florida Fish and Wildlife's acoustic receiver network and the American Littoral Society's volunteer angler tagging program indicated that no bluefish tagged on the Atlantic coast were ever recaptured in the Gulf of Mexico. Marine Recreational Information Program (MRIP) and commercial landings queries indicated that, on average, total bluefish harvest or removals (landings plus dead releases) in the Gulf of Mexico is 3-4% of combined Gulf and Atlantic coast bluefish removals. Finally, a query of recreational harvest length frequency suggested similarities between the two regions, which did not support a WG hypothesis that "missing" lengths in some observed periodic bimodal length frequency distributions on the Atlantic coast might reside in the Gulf of Mexico. Therefore, no data suggest that the Gulf of Mexico is an important habitat for Atlantic bluefish.

MRIP data and state and federal scientific surveys supported the seasonal migration pattern of bluefish, with fish observed in more southern locations in the winter and migrating northward in spring and summer. Spawning was also recorded in spring and summer, with eggs observed in some years in May through August. Length data from scientific surveys supported the presence of multiple sub-annual young-of-the-year cohorts in some years, although precise quantification of spring-spawned versus summer-spawned cohorts was generally not possible due to spatiotemporal variation in sampling effort and low sample sizes. Juveniles may be estuarine dependent (Munch 1997) although they also occur in nearshore ocean waters (Taylor et al. 2006); juvenile habitat use may vary by cohort (Taylor et al. 2007; Wuenschel et al. 2012). Adults use both estuarine and ocean environments and favor warmer water temperatures although they are found in a variety of hydrographic environments (Ross 1991; Shepherd and Packer 2006; Wuenschel et al. 2012). Small (≤30.3 cm) bluefish were generally found in the highest abundance along the Atlantic coast between Long Island and North Carolina. Medium (30.3-50.0 cm) bluefish were generally found in the highest abundance along the Atlantic coast as well as on Georges Bank. Large (≥50.0 cm) bluefish were generally found in the highest abundance in Southern New England and on Georges Bank.

In recent years, stakeholders have reported that larger bluefish are staying offshore and are less abundant inshore. The American Littoral Society Fish Tagging Program's bluefish data were analyzed to assess whether larger fish (> 18 inches or 46 cm) are being tagged and released or recaptured more frequently offshore. Analyses did not show that larger fish are being tagged and released or recaptured more frequently offshore in recent years (Working Paper 2, Valenti 2022a); however, very few large bluefish were tagged and released or recaptured in the last five years, so the sample size was notably small. The disparity between the stakeholder reports and the results of this analysis could be due to limitations inherent to volunteer fish tagging program data, including low sampling of bluefish and variability in angler effort and reporting.

A Vector Autoregressive Spatiotemporal (VAST; Thorson and Barnett (2017); Thorson (2019)) model was developed from the fall NEFSC bottom trawl survey to determine the fall centers of gravity of three bluefish size groups over time. Center of gravity analyses showed that medium bluefish are moving north and east at an average rate of 1.1 km/year, and large bluefish are moving north at an average rate of 0.2 km/year and east at an average rate of 0.5 km/year. The center of gravity of small bluefish did not have a trend. This distribution change may support anecdotes about large bluefish moving offshore in recent years. Further research is needed to fully understand bluefish distribution, as the 2020 and 2021 NEFSC fall bottom trawl surveys did not catch enough bluefish to be included in this study. Additionally, temperature was tested as a covariate in the VAST model, but resulting poor model diagnostics were beyond the scope of the present working group to address. See Working Paper 3 (Tyrell 2022) for more details.

1.1.3 Ecosystem Indicators

1.1.3.1 Temperature

Bluefish can tolerate temperatures ranging from approximately 11°-30.4°C, however they exhibit stress, such as an increase in swimming speed, at both extremes (Olla and Studholme 1971; Klein-MacPhee 2002). The literature indicated temperature and photoperiod are the principal factors directing activity, migrations, and distribution of adult bluefish (Olla and Studholme 1971, Taylor et al. 2007). Based on this mechanistic connection, quantitative indicators of optimal temperature were developed to better understand temperature trends during the bluefish spawning season.

The spawning season may now extend later in the year compared to historical periods. Bluefish spawning has been recorded at 18-25.6°C (Norcross et al. 1974). In the greater Mid Atlantic Bight and Southern New England regions, the first day when 75% or more of the sea surface reaches 18°C has remained stable over time, while the last day when 75% of the sea surface is above 18°C has occurred later in the year over time, currently persisting into mid and late October (in contrast to the beginning of October in the 1980s); the total number of days with 75% or more of the sea surface above 18°C has increased over time. It is unclear how these changes in potential spawning season may affect bluefish recruitment. There were no notable correlations between first, last, or number of days above 18°C and the composite young-of-the-year index used in the model (Section 3.1.2) or modeled recruitment (NEFSC 2019). However, the surveys used to characterize young-of-the-year bluefish may not fully document spawning that occurs in the fall due to a mismatch in survey timing and the recruitment of fall-spawned bluefish to estuaries; furthermore, most surveys do not capture the smallest bluefish (<15 cm length) that have been spawned most recently.

In contrast to trends in the potential spawning season, the amount of habitat in the optimal temperature range during the peak spawning month of July has not changed over time, indicating stability in spawning conditions and therefore possibly also in recruitment; however, the amount of habitat with colder-than-optimal temperatures (<18°C) has decreased, while the amount of habitat with warmer-than-optimal temperatures (>25.6°C) has increased. The amount of area in the Central Atlantic with optimal bluefish spawning temperatures in July was marginally positively correlated with bluefish recruitment (modeled recruitment in the 2015 assessment), while the amount of area with warmer-than-optimal temperatures was negatively correlated with bluefish recruitment. Although the amount of area with optimal temperatures in July has remained consistent over time, future ocean warming may eventually decrease the proportion of the Central Atlantic with optimal bluefish spawning temperatures as more areas warm above 25.6°C and fewer or no areas are left below 18°C.

1.1.3.2 Natural Mortality

Proxies for natural mortality were investigated to the extent possible. Relative condition of the small, medium, and large bluefish size groups was determined over time. Condition of large bluefish was found to be increasing over time, while the condition of medium bluefish had no change over time. Relative condition of small bluefish decreased slightly in the spring only, but remained above one, indicating good condition. Condition could be considered as a proxy for natural mortality, and generally indicates that mortality sources other than fishing have not increased compared to historical conditions, supporting the use of time-invariant natural mortality in the assessment model.

Bluefish predators are not well sampled, but existing data suggest that bluefish are not currently experiencing higher predation risk relative to historical conditions.

1.1.3.3 Condition and Recruitment

An increasing trend in the relative condition of large bluefish may be beneficial for bluefish recruitment, as larger and fatter bluefish may produce more eggs and/or more high-quality eggs; however, further research is needed to quantify the relationships between fecundity and length, weight, and age, which are currently not well documented.

1.1.4 Socioeconomic Indicators

Despite lower catches in recent years, bluefish remains one of the top recreational fisheries on the U.S east coast in terms of total catch, and therefore likely helps support a robust recreational fishing industry.

Although management was fairly stable in terms of catch limits and trip limits until the bag limit reductions implemented in 2020, the recreational fishery has shifted to catch-and-release rather than catch for harvest. Recreational landings in weight have decreased over time, with landings in 2021 being less than 10% of landings in 1981. Over the same time period, the total recreational catch (harvest plus all released bluefish) has decreased from 76 million fish in 1981 to 30 million fish in 2021.

Neither commercial nor recreational catch typically exceed catch limits, though in both 2020 and 2021 there were recreational catch limit overages of 32% in 2020 and 41% in 2021. However, the Acceptable Biological Catch (ABC) has generally decreased each year since it was implemented in 2011 due to stock condition. Therefore, recent decreases in catch and landings may be attributable to management actions rather than lack of interest in the bluefish fishery.

1.2 Forage Fish Index

The objective of this work was to create a forage fish index to evaluate changes in bluefish prey over time and space that could be used to inform survey and/or fishery availability in the bluefish stock assessment to inform annual deviations in catchability. Changing distribution and abundance of small pelagic prey may drive changes in predator distributions, affecting predator availability to fisheries and surveys. However, small pelagic forage species are difficult to survey directly, so we developed a novel method of assessing small pelagic fish aggregate abundance using predator diet data.

We used piscivore diet data collected from multiple bottom trawl surveys within a Vector Autoregressive Spatio-Temporal (VAST) model (Thorson and Barnett 2017, Thorson 2019) to assess trends of small pelagic forage species on the Northeast US shelf. This approach uses survey-sampled predator stomach contents as observations to develop a survey index for forage fish, following Ng et al. (2021), which used predator stomach data to create a biomass index for a single prey, Atlantic herring.

We adapted the approach of Ng et al. (2021) to generate an index for bluefish prey in aggregate rather than a single prey species. Further, we include inshore and offshore regions by combining two regional bottom trawl surveys, the NEFSC survey and the NEAMAP survey, as was done previously for summer flounder biomass (Perretti and Thorson 2019). Finally, since bluefish themselves are sparsely sampled by the surveys, we aggregate all predators that have a similar diet composition to bluefish to better quantify bluefish prey biomass.

Methods and results are summarized below; for more detail, see Working Paper 4 (Gaichas et al. 2022).

1.2.1 Forage Fish in Bluefish Diets

Using NEFSC bottom trawl survey diet data from 1973-2021, 20 small pelagic groups were identified as major bluefish prey, with 10 or more observations in bluefish stomachs over the entire 48-year period. In descending order of observations, bluefish prey are: longfin squids (*Doryteuthis* formerly *Loligo* sp.), anchovy family (Engraulidae), bay anchovy (*Anchoa mitchilli*), Atlantic butterfish (*Peprilus triacanthus*), Cephalopoda, striped anchovy (*Anchoa hepsetus*), red eye round herring (*Etrumeus teres*), sandlance (*Ammodytes* sp.), scup (*Stenotomus chrysops*), silver hake (*Merluccius bilinearis*), shortfin squids (*Illex* sp.), Atlantic herring (*Clupea harengus*), herring family (Clupeidae), bluefish (*Pomatomus saltatrix*), silver anchovy (*Engraulis eurystole*), longfin inshore squid (*Doryteuthis pealeii*), Atlantic mackerel (*Scomber scombrus*), flatfish (Pleuronectiformes), weakfish (*Cynoscion regalis*), and Atlantic menhaden (*Brevoortia tyrannus*).

Prey categories such as “fish unidentified”, “Osteichthyes”, and “unidentified animal remains” were not included in the prey list. Although unidentified fish and Osteichthyes can comprise a significant portion of bluefish stomach contents, unidentified fish in other predator stomachs may not represent the same types of unidentified fish in bluefish stomachs.

1.2.2 Predators Feeding Similarly to Bluefish

All size classes of 50 fish predators captured in the NEFSC bottom trawl survey were grouped by diet similarity to identify the size classes of piscivore species with the most similar diet to bluefish in the region. Diet similarity analysis was completed using the Schoener similarity index (Schoener 1970; B. Smith, pers. comm.), and is available via the NEFSC food habits shiny app. The WG evaluated several clustering methods to develop the predator list (see this link with detailed cluster results).

The nineteen predators with highest diet similarity to bluefish from the NEFSC diet database (1973-2020) included Atlantic cod, Atlantic halibut, buckler dory, cusk, fourspot flounder, goosefish, longfin squid, shortfin squid, pollock, red hake, sea raven, silver hake, spiny dogfish, spotted hake, striped bass, summer flounder, thorny skate, weakfish, and white hake.

The NEAMAP survey operates closer to shore than the current NEFSC survey. The NEAMAP dataset includes predators sampled by the NEFSC survey and adds two species, Spanish mackerel and spotted sea trout, not captured by the NEFSC survey offshore but included as bluefish-like predators based on WG expert judgement of diet similarity to bluefish. Predator size classes included are listed in Table 2 of Working Paper 4 (Gaichas et al. 2022).

1.2.3 Datasets

The mean weight of forage fish per predator stomach at each location was calculated by combining weight across the 20 forage fish (bluefish prey) groups found in stomachs from all 22 piscivores (including bluefish) at each surveyed location. Data for each station included station ID, year, season, date, latitude, longitude, vessel, mean bluefish prey weight (g), mean piscivore length (cm), number of piscivore species, and sea surface temperature (degrees C). Because approximately 10% of survey stations were missing in-situ sea water temperature measurements, National Oceanic and Atmospheric Administration Optimum Interpolation Sea Surface Temperature (NOAA OI SST) V2 High Resolution Dataset (Reynolds et al. 2007) data provided by the NOAA PSL, Boulder, Colorado, USA, from their website at https://psl.noaa.gov were used to fill gaps. For survey stations with in-situ temperature measurements, the in-situ measurement was retained. For survey stations with missing temperature data, OI SST was substituted for input into VAST models.

Models were developed combining all data for the year ("Annual") and with separate data for "Spring" (collection months January - June) and "Fall" (collection months July-December) to align with seasonal stratification used in the bluefish stock assessment. Modeled years included 1985-2021 to align with other data inputs in the bluefish stock assessment.

1.2.4 VAST Modeling

VAST is structured to estimate fixed and random effects across two linear predictors, which are then multiplied to estimate an index of the quantity of interest. Following the methods of Ng et al. (2021), we applied a Poisson-link delta model to estimate expected prey mass per predator stomach. We used a higher resolution (500 knots, estimated by k-means clustering of the data) to define the spatial dimensions of each seasonal model. Two step model selection first compared whether the data supported estimation of spatial and spatio-temporal random effects, and then evaluated whether catchability covariates improved fits. Best fit models included spatial and spatio-temporal random effects, with predator mean length, number of predator species, and sea surface temperature as catchability covariates; that is, these covariates all influenced the observation process rather than the distribution or abundance of prey. Detailed results of model selection are available in Working Paper 4 (Gaichas et al. 2022).

Similar to findings of Ng et al. (2021), a vessel effect was not supported, but the inclusion of the predator length covariate may more directly account for vessel differences in predator catch that affect stomach contents than modeling a vessel catchability covariate directly. Similar to our results, Ng et al. (2021) found that predator length covariates were strongly supported as catchability covariates (larger predators being more likely to have more prey in stomachs). In our aggregate predator dataset, we also found strong support for including the number of predator species in a tow as a catchability covariate. The rationale for including number of predator species is that more species "sampling" the prey field at a particular station may result in a higher encounter rate (more stomachs with bluefish prey). Water temperature was also supported as a catchability covariate, perhaps because temperature affects predator feeding rate and fish distribution.

1.2.5 Spatial Forage Indices

Spring, fall, and annual prey indices were split into inshore and offshore areas to quantify changing prey availability over time in areas available to the recreational fishery and the bottom trawl survey. First, we define a partition that includes survey areas relevant to the bluefish assessment (Mid Atlantic and Georges Bank). Within this partition,

- To evaluate bluefish availability to the NEFSC bottom trawl survey, two inshore-offshore strata partitions were created to account for the NEFSC survey vessel change in 2008. Inshore and offshore strata partitions included:
    - Albatross inshore stations (historically included in the Albatross NMFS bottom trawl index developed for the bluefish assessment)
    - Bigelow inshore bluefish index stations (historically included in the Bigelow NMFS bottom trawl index developed for the bluefish assessment)
    - Offshore bluefish index stations (the same for both vessels, and considered for addition to the NMFS bottom trawl bluefish indices in 2022)
- To evaluate bluefish availability to the MRIP catch-per-unit-effort (CPUE) index, recreational fishery strata partitions included:
    - shoreline to 3 miles offshore (state waters)
    - offshore of 3 miles (federal waters)

NEFSC survey strata definitions are built into the VAST "northwest-Atlantic" extrapolation grid. We defined additional new strata to address the recreational inshore-offshore 3-mile boundary, and incorporated them into a custom extrapolation grid so that the forage indices could be calculated and bias corrected (Thorson and Kristensen 2016) for all strata within VAST.

Full VAST model results for Fall, Spring, and Annual models, along with diagnostics, are available in Working Paper 4 (Gaichas et al. 2022). Here we show the forage fish index for the Fall model. The index is calculated for several regions relevant to the bluefish assessment:

- Albatross New (AlbNew) includes all inshore and new offshore survey strata (largest area)
- Albatross Old (AlbOld) includes all inshore survey strata
- Bigelow New (BigNew) includes the subset of inshore survey strata that can be sampled by the R/V Henry Bigelow plus new offshore strata
- Bigelow Old (BigOld) includes the subset of inshore survey strata that can be sampled by the R/V Henry Bigelow
- StateWaters includes the coastline to 3 nautical miles offshore (smallest area)

1.2.6 WHAM model example covariates: forage index time series, fall

Comparison of inshore and offshore spatial forage indices shows higher abundance of forage fish in state waters than in the subset of inshore strata that can be sampled by the R/V Henry Bigelow (Figure 2). Highest forage abundance is in the largest area, which includes all inshore survey strata as well as new offshore strata proposed for use in the bluefish assessment. The forage fish indices based on fall, spring, and annual datasets all show fluctuations in forage fish alternating between multiple years or decades with higher and lower levels. In general, the fall forage indices were higher at the beginning of the time series (mid-1980s), dropping to lower levels in the mid- and late-1990s, then showing cyclical fluctuations up until 2021, but never returning to high levels observed in the mid-1980s.

1.3 Life History Parameters

A single dataset was created from life history data collected by fishery-independent and fishery-dependent sampling by NMFS and Atlantic coast states and agencies. These data included ages, lengths, weights, and maturity observations. Life history processes were modeled, including mean length- and weight-at-age, modeled age-length relationships (e.g., Figure 3), allometric growth (e.g., Figure 4) and maturity-at-age (Table 1).

1.3.1 Age-Length and Length-Weight Relationships

Parameter estimates from the different life history models and expected values for size-, weight- and maturity-at-age and weight-at-length were generally consistent with analyses performed during the 2015 benchmark as well as with other previous research. Variability in the life history processes was modeled by splitting the data by semesters of the year, by decade, by geographic region (north and south, defined as Maine-Virginia and North Carolina-Florida, respectively), and by sex. Seasonal differences in length-at-age, weight-at-length, and maturity-at-age were apparent from these data; consistent with first principles, bluefish tended to be larger, weigh more and were more likely to be mature for their age during the second semester of the year relative to the first semester. Inter-decadal changes and differences by sex were less evident from the data. See Working Paper 5 (Truesdell et al. 2022) for more information including figures and tables outlining these findings.

1.3.2 Maturity-at-Age

The bluefish maturity schedule, in combination with the estimated total weight by age class, is used to estimate spawning stock biomass. The WG examined a variety of approaches to calculate maturity-at-age; all were based on using logistic generalized linear models (GLMs). The WG surmised through fitting models using different iterations of the data that the 2015 benchmark assessment had most likely employed federal data only (i.e., NMFS survey and port sampling data) to determine this ogive. The 2022 WG computed ogives using only federal data through 2021 and using state and federal data combined. The ogives that used both state and federal data estimated a mid-year maturity schedule by including time of year in the GLM; see Working Paper 5 (Truesdell et al. 2022) for more details. Ultimately, the differences in the versions of maturity ogives were not dramatic (Table 1) and the WG decided to use the same schedule that was implemented during the 2015 benchmark for primary model runs as this had been previously reviewed.

1.3.3 Natural Mortality

In the absence of direct natural mortality estimates for bluefish, the WG evaluated life history and longevity-based estimators of natural mortality. These methods included: maximum-age based methods, life history methods (using von Bertalanffy parameters), and length- or weight-based methods. The length and weight data described in the life history working paper (Truesdell et al. 2022a) were used for these calculations. For a detailed comparison of methods and natural mortality (M) estimates, see Working Paper 6 (Tyrell and Truesdell 2022).

The WG decided not to rely on natural mortality methods based on von Bertalanffy parameters, following the reasoning of Then et al. (2015). Based on the updated analyses using the maximum-age based methods of Hewitt and Hoenig (2005) and Then et al. (2015), the WG agreed that the "rule of thumb" estimate of 0.2 for all ages used in the 2015 benchmark assessment was too low. Ultimately, the WG decided to proceed with the Lorenzen weight-based age-varying natural mortality method because these estimates were in line with both the Hewitt-Hoenig (2005) and Then et al. (2015) estimates, and furthermore retained biological realism. The Lorenzen (1996) estimates using empirical weight-at-age (M range = 0.27-0.85) were higher at all ages than the age-constant value of 0.2 used in the previous benchmark assessment (Figure 5).

1.4 Uncertainty

Some of the ecosystem indicators identified in the ESP were not developed due to uncertainty around the underlying mechanistic connection to bluefish and/or the data source, and are described in detail in the ESP (Working Paper 1 Tyrell et al. 2022). For example, the WG proposed a large predator index to inform natural mortality, but could not locate sufficient data to create this index. The WG also identified overwinter survival as a bottleneck on survival to age 1, but could not develop any indicators of overwinter survival due to uncertainty in the locations where juvenile bluefish overwinter. Furthermore, the main source of long-term, large-scale fishery independent data for bluefish are bottom trawls (e.g., the NEFSC bottom trawls, NEAMAP), which are not well suited to capturing a large, fast-moving pelagic species like bluefish. As a result, the available data on the spatio-temporal distribution and movement of larger, older bluefish and the associated environmental indicators is limited. All of these proposed indicators are documented in the ESP and can be revisited as further data become available during future stock assessments.

The ESP identified several mechanistic linkages between the environment and bluefish stock dynamics and developed several quantitative environmental indicators. The development of these indicators is a first step towards including ecosystem variability in the assessment model to reduce uncertainty. With the bluefish stock assessment model now in a Woods Hole Assessment Model (WHAM) modeling framework, these environmental linkages can be tested in the future by including environmental covariates in model sensitivity runs, informing processes such as catchability, selectivity, natural mortality, and recruitment.

While the fall forage fish indices are temporally aligned with bluefish assessment inputs and both temporally and spatially aligned with two trawl survey indices used in the assessment, improvements in spatial overlap with recreational fisheries and other survey indices could be considered in the future to reduce uncertainty in associating forage fish with the bluefish MRIP abundance index. The current forage index does not cover inland waters, aside from Narragansett

Bay and Buzzards Bay. Diet data are available for Chesapeake Bay from the ChesMMAP survey, which could be added to the VAST model in the future. Less diet information is available for the portion of the bluefish range south of Cape Hatteras, although some collections have taken place. Investigation of sources of diet information, or possibly direct forage fish surveys for inland and southern areas would be worthwhile to see whether data are adequate to cover the full range of bluefish.

A key recommendation for future treatment of the life history data is to account for variability in spatio-temporal observations, as numerous fishery-independent and fishery-dependent sampling programs contributed to the available life history information, each with different sampling intensity across the Atlantic coast and across seasons and years. The VAST model developed for the NEFSC fall bottom trawl (Working Paper 3 Tyrell 2022) is a first step towards addressing some of this uncertainty caused by the spatiotemporal variability, but more work is needed to resolve the issue.

1.5 Incorporating Findings into Impacted TORs

TOR2 (catch data): The WG elected to omit Gulf of Mexico bluefish catch data from the assessment, based on the review of movement and distribution data; this is consistent with previous assessments for bluefish (NEFSC 2015). To capture seasonal and regional variations in growth and availability/distribution of bluefish, the WG used a seasonal length-weight relationship and seasonal-regional length frequencies to describe the age structure of the commercial and recreational catch.

TOR3 (survey data): The WG explored standardizing survey indices using generalized linear models (GLMs) parameterized with several environmental covariates (depending on the data collected during the survey), such as temperature, salinity, and dissolved oxygen. Ultimately, the WG found that the trends and uncertainty in the standardized survey indices were similar to the trends and uncertainty in the nominal indices, and decided to use nominal indices because they are simpler to maintain and update in management track assessments (e.g., versus possible future update GLM convergence issues). See TOR 3 and Working Paper 7 (Celestino et al. 2022a) for more details.

TOR4 (fishing mortality, recruitment, spawning stock biomass), TOR5 (stock status), TOR6 (projections): The forage fish index was incorporated into a companion model run as a covariate for catchability associated with the MRIP index. This companion model had good diagnostics, but was not put forward as the primary model due to concerns that it did not capture forage fish trends in the South Atlantic Bight, among other issues (see Section 4). The companion model was used to generate population estimates for comparison with the primary model, and generally showed similar results to the primary model, and its continued exploration and development is a high priority research recommendation (see TOR 7). The age-structured primary model used time-varying size-at-age from the observed average weight-at-age by year and fleet to calculate total and spawning stock biomass, to reflect the observed interannual variability in growth and condition. The primary assessment model also used the age-specific natural mortality schedule that was developed under TOR 1. Additionally, the assessment model was shifted into the WHAM platform in part due to WHAM's ability to incorporate environmental covariates in future model updates.

TOR7 (research recommendations): The ecosystem information compiled for TOR1 was used to inform several research recommendations under TOR7, most notably suggesting further sampling to resolve spring-spawned and summer-spawned cohorts, associated environmental drivers of relative cohort strength, and possible effects on the bluefish population, as well as testing additional environmental covariates in the WHAM modeling framework.

2 CATCH

> *Term of Reference #2: Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.*

For more detailed information on commercial and recreational data collection and analysis, see Working Papers 8-10.

2.1 Commercial Removals

2.1.1 Commercial Landings Data Collection

Commercial landings (1950 to present) for all species on the Atlantic coast are maintained in the Atlantic Coastal Cooperative Statistics Program (ACCSP) Data Warehouse. The Data Warehouse is an online database of fisheries dependent data provided by the ACCSP state and federal partners. The Data Warehouse was queried on 31 May 2022 for all commercial bluefish landings (monthly summaries by state, gear and market category) from 1985-2021 for Florida (Atlantic coast), Georgia, South Carolina, North Carolina, Virginia, Maryland, Delaware, New Jersey, New York, Connecticut, Rhode Island, Massachusetts, New Hampshire, and Maine.

2.1.2 Commercial Landings

Over the last approximately 40 years, commercial landings from the bluefish fishery ranged from a high of 7,162 MT (15.8 million pounds) in 1988 and have steadily declined to a low of 1,090 MT (2.4 million pounds) in 2021 (Figure 6). During this time, commercial landings have been consistently lower than the recreational catch and accounted for on average approximately 14% of the total removals in weight (Table 2, Figure 6). Amendment 1 to the bluefish Fishery Management Plan (FMP) was implemented in the year 2000 and the commercial fishery has been regulated by quota since this time. Gill nets are the dominant commercial gear used to target bluefish and average approximately 50% of the bluefish commercial landings from 1982 to 2021; this gear is fished primarily in the Mid-Atlantic and Florida. Other commercial gears, including hook & line, pound nets, seines, and trawls, collectively account for approximately 50% of the commercial landings.

Regional variations in commercial fishing activity are linked to the seasonal migration of bluefish. The majority of commercial fishing activity in the North and Mid-Atlantic occurs from late spring to early fall when bluefish are most abundant in these areas. As water temperatures decrease in late fall and winter, bluefish migrate south. Peak landings in the South Atlantic occur in late fall and winter. The majority of commercial landings over the time series (1950-present) have been taken in the Mid-Atlantic region (New York, New Jersey, and North Carolina); approximately 65% of the coast-wide total landings have been taken by these three states since

1982. Florida accounted for a larger percent of commercial catch historically (early 1980s) but has accounted for a diminishing proportion of landings over time.

2.1.3 Commercial Biological Sampling

Commercial fisheries from Maine to Virginia were sampled as part of the NEFSC data collection program (1985-2021). In addition, Virginia, North Carolina, and Florida have collected age and length data from their commercial fisheries to characterize the catch from the late 1980s onward. Since 2012, states that account for more than 5% of total coastwide landings have been required to collect 100 paired age and length samples, although those samples may come from any combination of commercial, recreational, or fishery independent sources. Sampling details were modified in 2020; see ASMFC 2020 and 2021 for more detail on state sampling programs. Length frequency data for Maine – North Carolina were expanded according to total landings in weight by market category and quarter. Biological data collection for the bluefish fishery south of North Carolina was sparse. Florida landings were characterized by North Carolina length frequencies from 1985-1991 due to lack of sampling in Florida; from 1992-2021, Florida samples were expanded by half-year (hereafter referred to as "season"). Landings from South Carolina and Georgia were generally negligible across the time series; when they occurred, they were pooled with Florida landings and characterized using the length frequency data used for Florida.

2.1.4 Commercial Length Frequency Distribution

The length frequency distribution from the commercial fisheries is characterized by a bi-modal distribution for much of the time-series. In the more recent years (2012-2021), the larger mode is reduced, leading to a skewed distribution with a peak around 35 cm. This pattern in bluefish length frequency has been observed in some years of the recreational harvest length frequencies, and the recreational discard length frequencies. The bi-modal pattern is likely a result of low availability to the fisheries of age 3 to age 4 bluefish. Bluefish are known to school by size class and it is speculated that movement dynamics at this age/size range affects availability of these fish. Much of this size cohort could be staying in the south (SC-FL) or offshore in certain years, and since the dominant fisheries for bluefish are coastal and north of Cape Hatteras, North Carolina, this would account for a reduced availability of this size/age class.

2.1.5 Commercial Discards

Previous bluefish technical committees (TCs) and working groups have concluded that commercial discards for bluefish along the Atlantic coast were insignificant, and historically this portion of the commercial catch has been ignored. The 2022 research track WG concluded that although commercial discards are a small fraction of the total catch, they should still be estimated and included in the commercial catch totals. To estimate commercial discards for bluefish, the Standardized Bycatch Reporting Method (SBRM) approach (Wigley et al. 2007) was applied, using the combined (D2) estimator. Commercial discard rates from 1989-2021 were calculated by half-year, gear, mesh, and region. A commercial discard mortality estimate of 32% was estimated via a literature review and meta-analysis based on the relevant gear types for the bluefish fishery and applied to the annual discards. See Appendix I to Working Paper 8 (Wood 2022a) for more details. Commercial landed lengths were used to characterize the size structure of the dead discards in all years due to the absence of adequate discard length samples.

Commercial bluefish dead discards have ranged from a high of 166 MT in 1996, to a low of 7 MT in 2017 (Table 2, Figure 6). Trawl and gillnet fisheries account for almost all of the discards, with small contributions from handline, longline, and midwater fisheries. Observed trips where bluefish was a primary target averaged around 1,800 trips per year over the time series. Commercial bluefish discards average 1.5% of the commercial catch by weight, and 0.2% of the total catch. While this portion of the catch is insignificant, the inclusion of these data will allow future shifts in magnitude to be monitored and accounted for in the assessment and more closely represent commercial allocations in catch accounting.

2.1.6 Commercial CAA and WAA

Seasonal length-weight parameters (Figure 4; Working Paper 5 Truesdell et al. 2022) were used to calculate numbers at length for the commercial catch. Final commercial catch-at-age (CAA) and weight-at-age (WAA) matrices were calculated using the annual seasonal multinomial age length keys (Section 3.3.1; Working Paper 14 Celestino et al. 2022b). The commercial catch is predominately comprised of age-1 and age-2 bluefish.

2.2 Recreational Removals

2.2.1 Recreational Data Collection

Estimates of recreational harvest and live releases for bluefish come from the NOAA Fisheries Marine Recreational Information Program, which uses a combination of effort surveys and angler-intercept surveys to develop those estimates (Papacostas and Foster 2018). This program was historically known as the Marine Recreational Fishery Statistics Survey (MRFSS), but was renamed in 2013 as NOAA Fisheries began making improvements to the survey design and estimation methods to address concerns identified by a National Academies review of the program (NRC 2006).

In 2018, MRIP transitioned from the Coastal Household Telephone Survey (CHTS) of effort to a mail-based survey, the Fishing Effort Survey (FES), following three years of side-by-side benchmarking. The CHTS and the FES only estimate effort for the private angler mode; the for-hire mode is covered by a separate survey, the For-Hire Survey (FHS). The FES produced consistently higher estimates of effort than the CHTS, so MRIP calibrated the historical estimates of catch and effort from the CHTS to the new scale of the FES estimates to provide a consistent time series (Papacostas and Foster 2018). The calibration model included fixed annual and seasonal effects as well as random effects and included information on trends in state-specific population size for the full time series and the prevalence of wireless/cell phone only households by state from 2007-2014. The calibration process also included the 2013 changes to the angler-intercept survey and corrections for the historical inconsistencies in the MRFSS intercept survey design (Papacostas and Foster 2018).

This increase in effort translated into an increase in total catch for bluefish, in both harvest and live releases. The overall trends in harvest and live releases were generally the same between the calibrated and uncalibrated time series, but the calibrated estimates were consistently higher. For a more detailed review of MRIP changes over time and the impacts of the calibration, see Working Paper 9 (Drew 2022a).

2.2.2 Recreational Harvest

Recreational harvest estimates of bluefish have averaged around 20,000 MT (44.1 million pounds) annually since 1985. From the 1980s to the early 1990s, recreational harvest declined by about 60%. The 1985-1989 average harvest was 52,064 MT (114.8 million pounds), while the 1990-1994 harvest averaged 22,285 MT (49.1 million pounds). Recreational harvest estimates continued to decline at a somewhat slower rate until reaching a low of 10,695 MT (23.6 million pounds) in 1999, increasing to 21,269 MT (46.9 million pounds) in 2010, and steadily decreasing since then to a value of 5,471 MT (12.1 million pounds) in 2021 (Table 2, Figure 6). In 2021, recreational anglers along the Atlantic coast caught 6.2 million bluefish, a 34% decrease from 2020.

The majority of recreational activity occurred from May to October, with the peak activity in July and August and almost 70% of the bluefish harvest being taken between July and October. The seasonal pattern varies by state, however, with more northern states seeing a peak in the summer and more southern states seeing peaks at the beginning and end of the year, reflecting both differing effort patterns by state and differing availability to states as bluefish migrate.

MRIP assigns catch to three fishing areas based on where anglers report doing the majority of their fishing: inland (which includes bays and estuaries like Long Island Sound, Chesapeake Bay, and Albemarle Sound), near-shore ocean (state waters less than three 3 miles from the shore), and offshore ocean (federal waters three miles or more from shore). About 51% of the catch of bluefish on a coast-wide basis came from inland waters, followed by near-shore ocean (42%) (Figure 7). Offshore ocean is only about 7% of the total catch. The inland portion of the harvest has been decreasing in recent years, with a concurrent increase in near-shore ocean harvest (Figure 7). For a detailed analysis of the spatial distribution of bluefish based on MRIP catch information see Working Paper 10 (Drew 2022b).

The majority of recreational harvest comes from the private boat and shore-based fishing modes (Figure 8). Less than 10% of the catch came from for-hire boats over the time-series.

2.2.3 Recreational Discards/Dead Releases

MRIP estimates of bluefish released alive have ranged from a low of 5.2 million fish (1988) to a high of 42.5 million fish (2001) from 1985-2021. Recreational release estimates have generally increased in proportion to harvested fish over the time series, increasing from approximately 19% of the total coast-wide catch in 1985 to over 80% in 2021. These releases represent both regulatory discards as well as voluntary releases by anglers practicing catch-and-release fishing.

About 48% of recreational bluefish releases on a coast-wide basis came from inland waters, and 48% from nearshore waters (Figure 7). Offshore ocean is only about 4% of the total releases. For a detailed analysis of the spatial distribution of bluefish harvest and releases based on MRIP data see Working Paper 10 (Drew 2022b).

The majority of recreational live releases comes from the private boat and shore-based fishing modes (Figure 8). Less than 10% of the releases came from for-hire boats over the time-series.

2.2.3.1 Recreational Release Mortality Rate

Estimating recreational catch-and-release mortality of bluefish is an important component of the stock assessment process given the popularity of catch-and-release angling in this fishery and the direct influence of release mortality on the total allowable catch. The literature reviews and analyses completed for the 2015 benchmark assessment (NEFSC 2015) were updated to re-assess the appropriateness of the 15% bluefish recreational release mortality estimate. From the updated literature reviews, no additional bluefish-specific release mortality papers were discovered, and one additional release mortality review paper (which was used for a meta-analysis) was discovered. Eleven exclusion criteria were applied to each bluefish-specific study and the studies within the review paper to determine which studies were suitable for inclusion in the bluefish-specific analysis and the meta-analysis. Three bluefish-specific studies passed the exclusion criteria. The individual mortality estimates from these three studies were used to calculate the mean (± standard error) bluefish-specific release mortality estimate, which was 9.4% ± 0.6%. From the review paper literature tables, 19 studies passed the exclusion criteria. The 22 individual mortality estimates from these 19 studies were used to calculate the mean (± standard error) meta-analysis release mortality estimate, which was 9.7% ± 1.9%.

The bluefish-specific release mortality estimate of 9.4% was used for this assessment. See Working Paper 11 (Valenti 2022b) for the full review and analysis.

2.3 Recreational Biological Sampling

2.3.1 Recreational Harvest

Recreational landings are sampled for length as part of the MRIP program. The MRIP length samples were used to expand recreational harvest per half-year season. In some years of the time series bluefish harvest lengths exhibit a bi-modal distribution, with a peak of fish around 35 cm, and a smaller peak around 70 cm. This trend has diminished in recent years but is consistent with trends seen in the commercial length frequency distributions. The bi-modal pattern is a result of an apparent low availability to the fisheries of age-3 to age-4 bluefish. Bluefish are known to school by size class and it is likely that unobserved movement dynamics at this age/size range affects availability of the population. It is possible a larger portion of the population at these sizes are staying south or offshore each year. Since the dominant fisheries for bluefish are coastal and north of Cape Hatteras, North Carolina, this would account for reduced availability of this size/age class.

The size of bluefish harvested by the recreational fishery varied by state and mode, with more northern states harvesting a wider range of sizes with a higher proportion of large bluefish than more southern states, which rarely harvest bluefish larger than 50 cm in fork length (Figure 9). In addition, bluefish harvested by the shore mode in states from Massachusetts through New York had a distinct peak of smaller fish around 15-20 cm that were not harvested by the private and for-hire boat modes. Young-of-the-year “snapper” bluefish are typically found inshore and are often targeted by shore-based anglers in the northern states. From New Jersey southward, that peak of smaller fish disappeared and the shore and boat mode length frequencies almost completely overlapped each other (Figure 9).

2.3.2 Recreational Dead Releases

MRIP conducts limited at-sea observing on headboat trips to collect lengths of fish released alive. To characterize the length frequency of the dead releases, the MRIP observer data were supplemented with lengths from the American Littoral Society (ALS) volunteer angler tagging program (by definition released fish), volunteer angler logbook programs in RI, CT, and NJ, and a volunteer angler tagging program in SC. See Atlantic States Marine Fisheries Commission (ASMFC) (2021) for more details on the state volunteer angler programs.

The recreational dead discard component of the catch was calculated using the season/region length frequency distributions developed from all of the recreational biological sampling data for released fish. For each year, expanded lengths were calculated by season/region and summed to get a seasonal total length distribution. Seasonal length-weight parameters (see above) were then used to calculate total seasonal weight and summed for a total annual release weight. A discard mortality estimate of 9.4% (Section 2.2.3.1) was applied to calculate the weight of dead discards for the total catch.

When the samples were pooled across season and region without weighting by removals, as was done for SARC 60, the harvest and release length frequency distributions appeared fairly distinct, with harvested fish centered around a smaller mean size than released fish (Figure 10). However, when stratified by region and season, the length frequencies for the harvested and the released fish were generally similar, with the exception of the northern region in the second half of the year, which had a peak of smaller fish in the harvest and a peak of larger fish in the releases (Figure 11). The majority of the release lengths were from the northern region; the southern region was not well sampled, particularly in recent years (see Working Paper 8 Wood 2022a for more details on sample size), and the differences in length frequency by region made it important to stratify the releases by region as well as season. Of note, in season/region/year cells where n < 30 fish, the cumulative length frequency of released alive fish was used as a proxy, instead of borrowing from another region or season.

Recreational releases/discards in 2021 were estimated at 14,792 MT, and after adjusting for a 9.4% mortality rate, the resulting discard loss was 1,391 MT. Recreational discard loss in weight has ranged from a low of 905 MT in 1988, to a high of 7,271 MT in 2001 (Table 2, Figure 6).

2.3.3 Recreational CAA and WAA

Final recreational harvest-at-age, dead releases-at-age, and weight-at-age matrices were calculated using the annual seasonal multinomial age length keys (Section 3.3.1; Working Paper 14 Celestino et al. 2022) applied to the harvest and dead release length frequencies. The recreational harvest-at-age and dead-releases-at-age were summed to calculate the total recreational dead catch-at-age. The recreational catch is predominately comprised of age-0, age-1 and age-2 bluefish.

2.4 Total Removals

Total bluefish removals (total dead catch) by component are presented in Table 2 and Figure 6. Overall, total removals have declined since the beginning of the time series. There was a slow increase from 1996 to 2010, but the declining trend has continued to the lowest values in the time-series in recent years (Figure 6). On average, commercial landings account for 14% of the total removals with commercial discards averaging only 0.2%. Total removals are dominated by the recreational fishery with recreational landings accounting for 71% of total removals, and recreational dead releases averaging 15% of total removals.

3 SURVEY DATA

> *Term of Reference #3: Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.*

The ASMFC Bluefish TC held a workshop in November 2021 to review the available state datasets for bluefish with the goal of evaluating their utility for this assessment, including fishery independent surveys. Metrics used to evaluate the datasets included the length of the time series, the geographic coverage, the quality and consistency of the survey design, and the prevalence of bluefish in the dataset, as measured by the percent positive tows or hauls for bluefish. Detailed descriptions of the surveys considered and the TC evaluations are available in the State Data Review Workshop Report (ASMFC 2021). The WG participated in the workshop and reviewed the final recommendations of the TC as to which datasets to include, exclude, or explore further. The WG's final decisions on which indices to include are summarized in Table 3 and Table 4. The surveys covered the majority of the bluefish range on the Atlantic coast, ranging from the Gulf of Maine in the north to Cape Canaveral, Florida in the south (Figure 12).

The suite of indices used in the base model was similar to what was used in SARC 60 (NEFSC 2015). Two new indices were added: the SEAMAP age-1 index (Section 3.1.1) and the ChesMMAP age-0+ index (Section 3.2.3). Two indices were dropped: the New Jersey Ocean Trawl Survey (NJ OT) and the Connecticut Long Island Sound Trawl Survey (CT LISTS). The NJ OT survey was dropped on the recommendation of the state data providers, as it was dominated by age-0 fish and did not seem to be adequately tracking age-1+ abundance (ASMFC 2021). The CT LISTS survey was removed for similar reasons: it covered a smaller spatial area than other trawl surveys in the model and was dominated by age-0 fish with little information on age-1+ fish. In addition, inclusion of the index resulted in worse model diagnostics without significantly affecting population estimates (Section 4.3.1).

The WG explored standardizing the fishery independent indices of abundance using environmental covariates in a GLM framework. However, the standardization process did not significantly affect index trends or reduce interannual variability or index coefficients of variation (CVs), so the WG did not use the standardized indices in the base run and instead used the stratified arithmetic mean for surveys with a stratified random design and the geometric mean for surveys with a fixed station design; see Working Paper 7 (Celestino et al. 2022a) for a detailed write-up of the process and results. The exception to this decision was the SEAMAP survey; see Section 3.1.1 below.

3.1 Recruitment Indices

For detailed descriptions of survey methods, see the ASMFC State Data Review Workshop Report (ASMFC 2021).

3.1.1 SEAMAP Age-0 and Age-1 Indices

The Southeast Area Monitoring and Assessment Program (SEAMAP) Coastal Trawl Survey has sampled the coastal zone off the southeast U.S. between Cape Hatteras, North Carolina and Cape Canaveral, Florida with a standardized protocol since 1990. A stratified random sampling design is used, with strata based on latitude and water depth. The SEAMAP survey encounters both age-0 and age-1 bluefish, with the age frequency varying by season. The spring survey is dominated by age-1 bluefish, while the fall survey is dominated by age-0 bluefish (Figure 13). Therefore, separate indices were developed for age-0 (fall-caught) and age-1 (spring-caught) bluefish.

SEAMAP used a GLM to calculate the indices for bluefish. The GLM standardization was able to smooth an exceptionally large value in the nominal index for age-0 bluefish which improved the correlation between the SEAMAP age-0 and lagged age-1 indices and the correlation between the SEAMAP age-0 and the composite age-0 indices (Section 3.1.2). In addition, due to vessel, weather, and funding issues, sampling in the northern-most strata of the survey has dropped off in recent years. Those strata have the highest abundance of bluefish in the SEAMAP survey, and the use of latitude in the standardization accounts for the decline in sampling in those strata. Therefore, the WG used the standardized age-0 and age-1 indices developed by SEAMAP for both indices (Zimney and Smart 2022).

The age-0 and age-1 indices have generally varied without trend over the time series; strong and weak year classes can be tracked from the age-0 to age-1 index in several years (Figure 14).

3.1.2 Composite Young-of-Year (YOY) Index

States from New Hampshire to Virginia conduct seine and trawl surveys for juvenile finfish that capture YOY bluefish. These surveys are noisy and cover small geographical areas, compared to the range of bluefish. Bayesian hierarchical modeling was used to combine these indices into a single composite index, using the method developed by Conn (2010), which represents the coast wide recruitment dynamics of bluefish. A composite index developed from state trawl YOY surveys (Table 4) was also explored, but it was not well correlated with the age-0 catch or any of the other indices and was not used in the assessment model. See Working Paper 12 (Drew 2022c) for details of the analysis. The surveys included in the composite index are described below.

Overall, the composite index did not show a strong trend over the time series; the early years were higher than later years, but also had more uncertainty around them (Figure 14).

3.1.2.1 New Hampshire Juvenile Finfish Seine Survey

The New Hampshire Juvenile Finfish Seine Survey samples at 15 fixed stations during June through November. The stations are spread throughout the New Hampshire coast, including the Hampton/Seabrook Estuary, Little Harbor, the Piscataqua River and Little Bay/Great Bay. Historical catches have ranged from 2.3 – 22 cm total length, all classified as YOY using a 25 cm size cutoff. Samples from November and December were removed from the analysis due to no positive catches. The survey has run from 1997 through the present. The nominal index was calculated as a geometric mean catch per tow with bootstrapping (n = 1000) to estimate the annual CVs. The index varied without trend (Figure 15).

3.1.2.2 Rhode Island Narragansett Bay Juvenile Finfish Beach Seine Survey

The Rhode Island Narragansett Bay Juvenile Finfish Beach Seine Survey currently samples 18 fixed stations throughout the bay; the survey began with 15 stations and added one additional station in each of 1990, 1993 and 1995. The survey began in 1988 and runs from June through October. A 25 cm size cutoff was used as the threshold to identify YOY bluefish. The nominal index was calculated as a geometric mean catch per tow with bootstrapping (n = 1000) to estimate the annual CVs. The early part of the time series was characterized by considerable variability. Catches were generally stable from 2010-2016, dropped during 2017 and 2018, and have since increased (Figure 15).

3.1.2.3 New York Western Long Island Seine Survey

The New York Department of Environmental Conservation Western Long Island Beach Survey has employed a consistent methodology since 1987 to sample sites at fixed stations within western Long Island bays: Little Neck and Manhasset Bay on the north shore of Long Island, and Jamaica Bay on the south shore (1984-present). Other bays have been sampled on a shorter time frame but were not included in this index. The nominal index was calculated as a geometric mean catch per tow with bootstrapping (n = 1000) to estimate the annual CVs. The index has generally varied without trend over the time series (Figure 15).

3.1.2.4 New Jersey Delaware River Seine Survey

The New Jersey Fish and Wildlife Delaware River Seine Survey is a fixed station beach seine survey conducted in three regions of the Delaware River. It targets age-0 striped bass, but bluefish are also captured in the brackish to tidal freshwater regions of the river. A 25 cm length cutoff is used to identify age-0 bluefish. The bluefish YOY index was reported as the geometric mean number of YOY bluefish per seine haul of samples collected from mid-June through September in region 1, with bootstrapping (n = 1000) to estimate the annual CVs; samples taken in October through November were excluded as YOY bluefish are rarely captured in those months. The index included data from 2002-2021, although 2020 was missing due to the COVID-19 pandemic. The index generally varied without trend, but the three lowest values in the time series occurred in 2016, 2018, and 2021 (Figure 15).

3.1.2.5 Maryland Juvenile Striped Bass Seine Survey

The Maryland Department of Natural Resources Juvenile Striped Bass Seine Survey is a fixed station survey that samples in major striped bass spawning areas in Maryland’s portion of the Chesapeake Bay from July – September. A subset of 13 sample sites was selected for the development of a juvenile bluefish index from 1981 to present. The nominal index was calculated as a geometric mean catch per tow with bootstrapping (n = 1000) to estimate the annual CVs. The index is variable but has shown a declining trend over time, with low catch rates and a low proportion of positive hauls in recent years (Figure 15).

3.1.2.6 Virginia Institute of Marine Science Juvenile Striped Bass Seine Survey

The Virginia Institute of Marine Science Juvenile Striped Bass Seine Survey is a fixed station survey that samples from July – September in the James, York, and Rappahannock Rivers, as well as in the main tributaries of these systems. The nominal index was calculated as a geometric mean catch per tow with bootstrapping (n = 1000) to estimate the annual CVs. The index showed a period of higher recruitment from the late 1980s to the late 1990s, followed by a period of lower recruitment from the early 2000s forward, although 2019-2021 have been higher (Figure 15).

3.2 Age 0+ Indices

3.2.1 Northeast Fisheries Science Center (NEFSC) Fall Inshore Trawl Survey

Since 1963, the NEFSC has conducted a standardized bottom trawl survey during the fall and spring along the northeastern continental shelf of the United States in the area comprising the Western Scotian Shelf of the Gulf of Maine, south to Cape Lookout, North Carolina. The survey uses a stratified random design. There was a vessel change in 2009 from the F/RV Albatross to the F/RV Bigelow, which resulted in the loss of historical inshore strata from the survey area, all of which are now sampled by the Northeast Area Monitoring and Assessment Program (NEAMAP) via the Mid-Atlantic/Southern New England Nearshore Trawl Survey (Section 3.2.1), the Massachusetts Division of Marine Fisheries Bottom Trawl Survey (ASMFC 2021) and Maine-New Hampshire Inshore Trawl Survey (ASMFC 2021). For more information on the NEFSC bottom trawl survey design, see Avarovitz (1981) and NEFSC (2015).

Bluefish are predominately caught during the fall in the inshore strata south of the Gulf of Maine, so fall inshore strata from Cape Hatteras to Cape Cod were used to build two indices for bluefish, one for the Albatross years (1985-2008) and one for the Bigelow years (2009-2021). The indices were calculated as the stratified mean catch-per-tow. The Albatross index showed high variability at the beginning of the time series followed by a generally increasing trend from the mid-1990s to the mid-2000s (Figure 16). The Albatross index declined from 2005 to the end of that time series in 2009, and the Bigelow has shown a consistent decline over its entire time series from 2009-2021 (Figure 17).

The fall stratified mean length frequencies of the Albatross and Bigelow indices were apportioned to ages by applying the annual fall age-length key (Section 3.3.1). The age-structure of the Albatross and Bigelow indices was dominated by age-0 fish (Figure 16); the Bigelow had a higher proportion of age-1 fish than the Albatross did, but was still dominated by age-0 and age-1 fish (Figure 17).

3.2.2 NEAMAP Mid-Atlantic/Southern New England Nearshore Trawl Survey

The NEAMAP Mid-Atlantic/Southern New England Nearshore Trawl Survey uses a stratified random design to sample the coastal ocean from Martha’s Vineyard, MA to Cape Hatteras, NC since the fall of 2007. NEAMAP conducts two cruises per year, one in the spring and one in the fall, and samples inshore areas that were lost from the NEFSC Bottom Trawl Survey with the vessel change in 2009. The index was calculated as the stratified mean catch-per-tow for the fall cruise where the bluefish catch and proportion positive tows were higher. The index has been variable with a somewhat declining trend, reaching a time-series low in 2019 before increasing in 2020 and 2021 (Figure 18).

The fall stratified mean length frequency of the NEAMAP index was apportioned to ages by applying the annual fall age-length key (Section 3.3.1). The age-structure of the NEAMAP index was dominated by age-0 bluefish (Figure 18).

3.2.3 Chesapeake Bay Multispecies Monitoring & Assessment Program (ChesMMAP)

The Chesapeake Bay Multispecies Monitoring & Assessment Program (ChesMMAP) uses a stratified random design to sample the mainstem of Chesapeake Bay every other month from March through November. The survey underwent a vessel change in 2019, and the calibration work has not been completed. As a result, the ChesMMAP index for bluefish includes data from 2002-2018. The index was calculated as the stratified mean catch-per-tow for the May through November cruises, where the bluefish catch and proportion positive tows was highest. The index has generally varied without trend over the time series (Figure 19).

The length frequency of the ChesMMAP index was stratified by season (May-June and July-November). The seasonal length frequencies were apportioned to ages by applying the appropriate seasonal age-length key, and the final age composition of the index was calculated by summing the seasonal index age compositions. The age-structure of the ChesMMAP index was dominated by age-0 and age-1 bluefish and had no observations greater than age-3 (Figure 19).

3.2.4 North Carolina Pamlico Sound Independent Gill Net Survey (PSIGNS)

The North Carolina Division of Marine Fisheries Pamlico Sound Independent Gill Net Survey (PSIGNS) uses a stratified random design to sample the Pamlico Sound estuary from mid-February to mid-December. Bluefish is the second most commonly caught species in the survey. The index was calculated as the stratified mean catch-per-set for all months. The index increased from 2001 through 2007 and then declined to a time-series low in 2015; subsequent years have increased slightly, and 2019 was an extremely high value (Figure 20).

The length frequency of the PSIGNS index was stratified by season (February-June and July-December). The seasonal length frequencies were apportioned to ages by applying the appropriate seasonal age-length key, and the final age composition of the index was calculated by summing the seasonal index age compositions. The age-structure of the PSIGNS index was dominated by age-1 and age-2 fish, and had the highest proportion of older fish of all fishery-independent indices used in the assessment (Figure 20).

3.2.5 MRIP Recreational CPUE

The MRIP dockside intercept program dataset was used to develop recreational total catch-per-unit-effort (i.e., harvest plus live releases in numbers) as an index of abundance for bluefish. Bluefish trips were defined using a guild approach where a trip was considered a bluefish trip if it caught either bluefish or a species that was significantly positively associated with bluefish. This was a change from the previous benchmark assessment where effort was described using “directed trips,” which describe trips where bluefish were considered a target species. The CPUE was standardized using a zero-altered negative binomial model with year, state, wave, state-wave interaction, mode (e.g., shore, private boat, charter), area fished, kind of day (i.e., weekday or weekend), and angler avidity as factors and angler-hours per trip as an effort offset. For more information on the MRIP CPUE development, see Working Paper 13 (Drew 2022c).

The MRIP CPUE peaked at the beginning of the time-series, declining through the mid-1990s after which it showed a stable to slightly increasing trend until 2016 (Figure 22). It has declined in recent years. The choice of trip definition (guild trips vs. directed trips) and standardization model (zero-altered negative binomial vs. negative binomial) resulted in significant changes in overall trend compared to the index developed during the last benchmark (directed trips standardized with a negative binomial model with no interaction terms). While the indices showed roughly similar trends – declining through the mid-1990s before stabilizing and increasing somewhat – the MRIP CPUE used in this assessment showed much more contrast than the continuity run index used in the SARC 60 assessment, starting out at a higher level, declining to lower levels, and not recovering as much after the decline (Figure 21).

The age-structure of the MRIP CPUE was developed from the recreational catch and release information, as the CPUE used both harvested and released alive fish in the calculation of the catch per unit effort. The recreational harvest numbers-at-age matrix was combined with the recreational live release numbers-at-age matrix. Unlike the recreational removals matrix, the live releases numbers-at-age were not scaled by the release mortality rate. The MRIP CPUE had a broader age-structure than the fishery independent indices (Figure 22).

3.3 Age and Length Data

3.3.1 Age-Length Keys

The WG evaluated multinomial age-length keys (ALKs) relative to traditional ALKs, and ALKs resolved at a seasonal as well as season-region level of resolution; for complete details, see Working Paper 14 (Celestino et al. 2022b). Briefly, multinomial ALKs were explored as an objective, repeatable, and efficient way to fill gaps in ALKs. The data to construct the ALKs for bluefish were sparse early in the timeseries, and throughout the time series when subset to a season (January-June and July-December) and region (Florida-North Carolina and Virginia-Maine) level of resolution. The multinomial approach to developing ALKs (Gerritsen et al. 2006) has been explored across a number of stocks assessed by ASMFC and NOAA Fisheries, and is available in modelling software [e.g., weakfish (ASMFC 2019), Stock Synthesis (Methot and Wetzel 2013)].

Age and length data collected by fishery-independent and fishery-dependent sampling by NMFS and Atlantic coast states were compiled and used to construct ALKs. When developing and comparing ALKs, the WG developed various borrowing and multinomial model configuration rules. Final multinomial ALKs were constructed using all data (all years, seasons, regions combined) input through a single model, with terms for year and season (or year, season, and region for exploration of the seasonal-regional ALKs). All spring age-0 fish were removed from the dataset prior to running multinomial models, which helped the performance of model predictions relative to biological expectations (e.g., minimized the probabilities of spring age-0 fish in ALKs). Traditional ALKs (i.e., non-model based ALKs such as those used in the 2015 assessment) were only constructed at the year-season level due to concerns related to the sparse nature of data at the year-season-region level of resolution that would require a large number of decisions related to data borrowing. Spring age-0 fish were also removed from that dataset.
The influence of multinomial and traditional ALKs, at the season and season-region level of resolution was evaluated in a statistical catch-at-age framework. Total catch and indices of abundance were apportioned into ages using seasonal traditional ALKs, seasonal multinomial

ALKs, and season-region multinomial ALKs. Statistical catch-at-age model performance was similar among model runs (Table 5). The scale of retrospective patterning was comparable between models that used multinomial ALKs, but higher for the model that used traditional ALKs. After extensive discussions, the WG did not believe data were sufficient to support an ALK model at a seasonal-regional level of resolution. All sample sizes are reduced as data are subset to finer spatial and temporal resolutions.

While Akaike's Information Criteria (AIC) among the ALK multinomial models supported seasonal-regional ALKs (vs seasonal multinomial ALKs), the WG did not have high confidence in partitioning data at this level of resolution, and so supported the use of the seasonal multinomial ALKs applied to season-region length frequencies for the catch and indices. The statistical catch at age model results suggested less retrospective patterning with the seasonal multinomial ALKs compared to the traditional seasonal keys, and so the WG supported use of the seasonal multinomial ALK for continued modelling (and ultimately the base model).

4 ASSESSMENT METHODS

> *Term of Reference #4: Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.*

4.1 History of the Bluefish Assessment

A statistical catch at age assessment model was first used to assess bluefish and provide management advice in 2005, at the Stock Assessment Workshop 41 review (NEFSC 2005). Prior to this review, several model types were explored including a modified Delury model, a surplus production model, a VPA, and catch-at-age models. At the time, the Bluefish TC concluded that age-based models such as a VPA or catch-at-age were the most appropriate for the bluefish assessment and age-based models have been used since.

At the last benchmark assessment in 2015, a number of changes were made to the data structure and assessment model. Major changes included fitting to the age composition of the surveys (as opposed to age-specific indices), separating total catch into two fleets (commercial and recreational), updated maturity-at-age information, splitting the Bigelow and Albatross survey time series into two indices, and changing MRIP index selectivity from independent estimates at-age to a logistic curve. The final model was reviewed during SAW/SARC60 (NEFSC 2015) and has been used to provide management advice since 2015.

The most recent operational assessment for bluefish took place in 2021, with data through 2019. Based on this assessment update of the 2015 benchmark model, the bluefish stock was overfished and overfishing was not occurring relative to the updated biological reference points. Spawning stock biomass was estimated to be 95,742 MT in 2019, about 47.5% of $SSB_{35\%}$

(201,729 MT) and 95% of the threshold (100,865 MT). Fishing mortality was estimated to be 0.172, which was 95% of $F_{35\%}$ (0.181). Average recruitment from 1985-2019 was 46 million age-0 fish. The terminal year estimates for fishing mortality and spawning stock biomass adjusted for retrospective error were within the 90% confidence bounds of the terminal year estimates, indicating no retrospective adjustment was needed for stock status determination.

4.2 Bluefish Research Track 2022 Model Introduction

The Research Track (RT) 2022 model building procedure for bluefish was accomplished over multiple steps. The majority of the model bridge was built using ASAP (Age Structured Assessment Program, Legault and Restrepo 1999), which was the previously approved assessment model. ASAP is an age-structured model that uses forward computations, assuming separability of fishing mortality into year and age components, to estimate population sizes given observed total catches, catch-at-age, and indices of abundance. Bluefish are modeled as age-0 through age-6+, with ages six and older pooled into a plus group. The separability assumption is partially relaxed by allowing for fleet-specific computations and by allowing the selectivity-at-age to change in blocks of years. Weights are specified for different components of the objective function, which allows for configurations ranging from relatively simple age-structured production models to fully parameterized statistical catch-at-age models. The objective function is the sum of the negative log-likelihood of the fit to various model components. Catch-at-age and survey age composition are modeled assuming a multinomial distribution, while most other model components are assumed to have lognormal error. Specifically, lognormal error is assumed for: total catch in weight by fleet, survey indices, stock recruit relationship, and annual deviations in fishing mortality. Recruitment deviations are also assumed to follow a lognormal distribution, with annual deviations estimated as a bounded vector to force them to sum to zero (this centers the predictions on the expected stock recruit relationship, or on mean recruitment when steepness is fixed at one). For more technical details, the reader is referred to the technical manual (Supporting documentation: ASAP manual, Legault 2012).

Early WG discussions led to the decision that the bluefish assessment model should be shifted into a new modeling framework, the Woods Hole Assessment Model (WHAM: Miller et al 2016, Miller and Hyun 2018, Miller et al. 2018). WHAM is a general state-space age-structured assessment model that is able to include environmental and other covariate effects on population processes. The shift from ASAP to WHAM allowed more flexibility, including the estimation of observation and process error, and the propagation of random effect parameters in stock projections. The final ASAP model was transitioned into its “ASAP-like” WHAM model counterpart, which was parameterized so that it was essentially identical to the ASAP model; after this initial WHAM model was fit, a suite of models that included random effects on the numbers-at-age were fit, and model selection via AIC was used to select a best model. Environmental indices based on a VAST analysis of forage fish availability along the east coast (Section 1.2) were also explored as covariates on the catchability of survey indices.

4.3 Bluefish Research Track 2022 Model Bridge

4.3.1 ASAP Modeling

The first step in modeling in ASAP was to conduct a continuity run, which updated the current assessment model with data through 2021. A base model was then constructed by adding new data (CAA and WAA) and indices to the continuity run, keeping the same model settings and weights. A model bridge was then built from the base model to a final ASAP model by changing model data formulation, specifications, and weighting inputs. In total, about 80 variations of ASAP models were explored during this bridge building procedure. The model steps with the most important changes that provide a linear path from the base model to the final ASAP model are presented below. See Working Paper 15 (Wood 2022b) for a detailed description of the bridge-building process, results, and diagnostics. Working Paper 16 (Wood 2022c) includes the complete diagnostic plots for the major milestone runs, including the final ASAP model; the diagnostic plots for the final ASAP model alone are linked below.

The continuity model run was carried out as update of the SAW/SARC 60 benchmark final model, which is the model currently used for management advice. Total catch, catch-at-age, weight-at-age, and indices-at-age were updated for 2020 and 2021 using previously established data protocols (Figure 23). Retrospective pattern for the continuity run was examined for F, SSB, and recruitment using 7-year peels. The analysis showed consistent and significant pattern in the estimates of F and SSB, with Mohn's rho values of -0.277 and 0.294, respectively. Recruitment estimates exhibited lower retrospective pattern that was inconsistent over the peels, with a Mohn's rho estimate of 0.170. The continuity run had poor convergence diagnostics; a jitter analysis indicated that when the parameter initial values were varied randomly, only 130 of 200 realizations of the model reached the same final objective function value as the base run and there were 18 non-converged models. Gradient values were also poor for a number of the runs that did converge, with the majority of maximum gradient values being greater than 0.0001, a value often used as threshold for an acceptable model (Carvalho et al. 2021). This diagnostic was not explored in the 2015 benchmark assessment.

The switch to multinomial age-length keys had a significant impact on estimates of SSB. The multinomial keys had the effect of spreading numbers-at-age in the older ages to younger ages, especially with the plus group. This had the result of lowering the SSB as the total biomass of mature fish was reduced. The multinomial keys substantially improved the convergence diagnostics for this model. All previous models that used the traditional keys had poor convergence diagnostics. A jitter analysis of starting parameter values showed that the previous model step with traditional keys failed to converge 52 times and only found the original model solution 129 times out of 200 jitter runs. Conversely, the model with multinomial keys was very robust to the original objective function solution and did not seem overly sensitive to the initial starting values; it failed to converge only 6 times, and found the original objective function 193 times, out of 200 jitter runs. All of the alternate objective function values were higher than the original objective function value.

The change from the directed trips method to the guild approach (Section 3.2.5) to develop the MRIP index was another significant change for the data going into the model. The MRIP index has historically been the most important index in the assessment model and effectively scales the model because of the assumed logistic selectivity. Without this flat-top selectivity, the model is able to create cryptic biomass in the older ages and can produce unrealistic results. Due to the importance of this index, small changes in trend have dramatic impacts on the scale of model results. The continuity MRIP index (i.e., using the directed trips effort and the previous standardization model) remained fairly flat throughout the time-series. Shifting to the guild approach for the index calculation resulted in a much different trend, with the guild approach index starting out at higher values and declining to lower levels compared to the continuity run directed trips index (Figure 21).

The overall effect of this new index on the model was a decrease in SSB and an increase in F. The switch to the MRIP guild index significantly reduced the retrospective pattern in SSB from 40% to 25%. The retrospective pattern in fishing mortality was also reduced from 36% to 21%. Convergence diagnostics in the ASAP framework for the model with the guild approach MRIP index were very good, with 191 of 200 jitter runs finding the original model solution and 8 non-convergences. All of the alternate objective function values were higher than the original objective function value.

The change from an age-constant natural mortality of 0.2 to the higher, age-varying estimates of M from the Lorenzen (1996) method increased SSB, decreased F, and greatly increased recruitment, as would be expected. The retrospective pattern was increased for all model results with the change in natural mortality (Table 6). Convergence diagnostics were very similar between the age-constant and age-varying M.

The previous assessment model specified a single selectivity block for each fleet for the entire time-series. To address the retrospective pattern and the patterning in the catch-at-age residuals, a selectivity block was added in each fleet beginning in the year 2000, which is the year Amendment 1 to the bluefish fishery management plan was implemented. An additional selectivity block was added in the recreational fleet from 2011-2021, to align with the increasing trend in the proportion of the recreational catch from the southern region, which tends to catch a smaller size range of bluefish (Working Paper 9 Drew 2022a). The addition of new selectivity blocks increased SSB estimates and reduced retrospective pattern.

Model fit diagnostics for the Connecticut Long Island Sound Trawl (CT LIST) survey index indicated a somewhat poor fit early on in the time-series, with two blocks of residuals from 1985-2000. This survey also caused issues with the estimation of retrospective peels, with some peels giving gradient estimates >0.001, indicating poor or no convergence. The removal of the CT LIST survey resulted in slight increases in both SSB and recruitment, and little change in fishing mortality. There was a small improvement in the retrospective pattern in SSB, and a small increase for the pattern in fishing mortality; this model did not have the retrospective peel convergence issues that occurred in previous models that included the CT LIST survey.

The WG chose model BF24 as the final bluefish ASAP model configuration, prior to migration into the state-space framework of the Woods Hole Assessment Model. A full suite of input, results, diagnostic, retrospective and MCMC plots are available for this run as part of Working Papers 15 and 16 (Wood 2022b and 2022c) and as a standalone file which can be downloaded or viewed from the following link: BF24 plots. When reviewing the ASAP plots, note that ASAP numbers the age classes starting with age-1, but the first age in the bluefish model is age-0. Therefore, all ages in the figures are increased by one relative to the biological age-class they represent (ASAP age-1 is really age-0, ASAP age-2 is really age-1, etc.). A brief summary of main model results is presented below.

The final ASAP model fleet selectivity-at-age estimates for the two fleets each show a decrease in selectivity at middle ages (ages 3-4), with selectivity increasing at older ages. The final selectivity block in the recreational fleet (2011-2021) has more of a domed shape, with older fish having much lower selectivity than previous blocks (Figure 24). Final ASAP model estimates for the index selectivities show a rapid decrease in selectivity after age-0. A few of the indices have higher selectivity towards larger/older fish, the most important being MRIP and PSIGNS, and to a lesser extent the Bigelow survey.

Abundance results from model BF24 showed a maximum of 424 million fish in 1985, declining to 166 million in 1995, and then increasing to a peak of 311 million in 2006. Total abundance declined from the peak in 2006 to a low of 147 million in 2016, a small peak to 208 million in 2018, and a terminal year estimate of 169 million fish. Spawning stock biomass started from a high of 208,791 MT in 1985 and declined over the time-series to a low of 44,931 MT in 2018, and increased since to a value of 63,320 MT in 2021. The majority of the spawning stock biomass is ages 5 and 6+ (30-60%) for the entire time-series. Fully selected fishing mortality in 2021 was 0.159, compared to an average full F from 1985 to 2021 of 0.354. Estimates of F have varied over the time-series from a peak in 1987 of 0.519 to the lowest value of 0.159 in 2021. Estimates of recruitment have remained steady over the time series, fluctuating around an average value of 127 million fish. Recruitment has been below average for the past 12 years, and was estimated at 95 million fish in 2021.

Retrospective pattern for the final model was examined for F, spawning stock biomass, and recruitment. There was a notable retrospective pattern in both SSB (Mohn's rho = 0.326) and fishing mortality (Mohn's rho = -0.277), with very little in recruitment (Mohn's rho = 0.017). Shifting this assessment model into the state-space framework of WHAM and estimating random effects helped to improve the retrospective diagnostics of this model.

The variation in the final ASAP model results for F and SSB was determined using a Markov chain Monte Carlo (MCMC) with 1000 iterations and a thinning factor of 2000 (2,000,000 iterations). Trace plots for both SSB and F show little to no patterning. There is no significant autocorrelation in the SSB or F chains. Terminal year 90% confidence intervals (CI) from the MCMC ranged from 49,856 to 71,780 MT, with a median estimate of 60,338 MT. The 2021 SSB point estimate from the final model (63,320 MT) is slightly higher than the median estimate from the MCMC distribution. The 90% CI around the terminal year F ranged from 0.112 and 0.231. The point estimate from the final model (0.159) is nearly identical to the median estimate (0.160) from the MCMC distribution.

Model BF24 had good convergence diagnostics with 192/200 jitter runs finding the original model solution, and 4 non-convergences (Figure 25).

4.3.2 Woods Hole Assessment Model (WHAM) Modeling

The Woods Hole Assessment Model (WHAM: https://github.com/timjmiller/wham) is a state-space age-structured stock assessment model developed at the Northeast Fisheries Science Center (NEFSC, Stock and Miller, 2020). WHAM is a flexible model framework that can be configured as a traditional statistical catch-at-age model, which allows for bridge building transitions from models like ASAP. In addition to the traditional catch-at-age approach, WHAM allows for the estimation of state-space effects, including annual transitions in the numbers-at-age, age and time varying random effects on natural mortality or selectivity, and the ability to incorporate environmental effects as covariates on population processes.

The final bluefish model from the ASAP model bridge (model BF24) was moved into WHAM for further model exploration. The WG made the decision to finish model exploration in WHAM because of its flexible framework, specifically allowing for the estimation of random effects on recruitment and numbers-at-age. A desirable feature of the state-space framework is that these models tend to have lower retrospective pattern in model results, and more realistic estimates of uncertainty (Stock and Miller, 2020). Model BF24 had a notable retrospective pattern in both SSB and F (Table 6) and this was a primary driver for moving the bluefish model into WHAM.

In addition to improving retrospective pattern, the final bluefish model was shifted into WHAM to explore environmental covariate links on the catchability of different surveys indices. Forage fish indices were developed using a VAST model (Section 1.2; Working Paper 4 Gaichas et al. 2022) and explored as environmental covariates on the catchability (q) of NEFSC survey indices and the MRIP catch-per-unit-effort index.

The focus of the model exploration in WHAM was to refine the final bluefish model from ASAP, and not continue building a model bridge. This refinement focused on models with random effects on recruitment and numbers-at-age. The models explored had different options for treating the yearly transitions in survival (numbers-at-age):

1. Deterministic survival: a traditional statistical catch-at-age (SCAA) model, recruitment in each year is estimated as independent fixed effect parameters.
2. Recruitment deviations (random about mean) are random effects
    a. Random effects are independent, uncorrelated: model subscript **_m2** going forward
    b. Autoregressive (AR1) by year (autocorrelated): model subscript **_m3** going forward
3. Full state-space model where survival of all ages are random effects
    a. Random effects are independent, uncorrelated: model subscript **_m4** going forward
    b. Autoregressive (AR1) deviations by year: model subscript **_m5** going forward
    c. Autoregressive (AR1) deviations by age: model subscript **_m6** going forward
    d. Autoregressive deviations by age and year (2D AR1): model subscript **_m7** going forward

To assess the fit and results of each model, a series of diagnostic criteria were applied. First, models were designated as converged if the maximum gradient was less than 1e-10 and the hessian matrix was invertible. Next, a model selection process using AIC was carried out to choose a best model among models with comparable likelihood structures. Convergence properties of the best models chosen by AIC were further explored using a jitter approach analogous to the approach used in ASAP. Parameter starting values were randomly generated using the model covariance matrix to develop random normal distributions around the MLE parameter estimates as well as a distribution scaling factor, which alters the spread of the distribution around the potential starting values by scaling the variance. Similar to the ASAP jitter approach, 200 iterations of the model were carried out to test model sensitivity to the initial parameter guesses and investigate convergence. The 200 realizations of the model objective function and gradient were examined to see how robust the model was to the starting values.

A suite of model fit diagnostic plots were also examined for each model of interest. Model fits were examined using conventional residual diagnostics, as well as one-step ahead residual diagnostics (OSA), which are more appropriate for state-space models with correlated parameters (Trijoulet et al. 2023). Finally, retrospective pattern in model results was evaluated using Mohn's rho values (Mohn 1999) calculated from 5-year model peels (Miller and Legault 2017, ICES 2020).

When reviewing the WHAM plots, note that WHAM (similar to ASAP) numbers the age classes starting with age-1, but the first age in the bluefish model is age-0. Therefore, all ages in the figures are increased by one relative to the biological age-class they represent (WHAM age-1 is really age-0, WHAM age-2 is really age-1, etc.).

4.3.2.1 Model BF24W: Run the final ASAP model as a traditional SCAA model in WHAM

The first step in WHAM modeling was to run the ASAP final model (BF24) as a traditional statistical catch-at-age model. A comparison of model results from the final ASAP model and BF24W show nearly identical results (Figure 26). The slight differences in model results can be attributed to different objective function and minimization algorithms between the two model frameworks.

One-step ahead residual diagnostics for the fleets indicate that the input CV of both fleets might be too broadly specified, with very tight blocking around 0 for the commercial fleet (fleet 1), and poor quantile distributions for both fleets (Figure 27).

4.3.2.2 Model BF26W to BF28W: Reduce CV around fleets

This series of models reduced the CV around fleet 1 by a factor of 0.5 (BF26W), the CV around fleet 2 by a factor of 0.5 (BF27W), and then both fleets' CVs by a factor of 0.5 (BF28W).

4.3.2.3 Model BF28W with different for NAA deviations specifications

Modal BF28W was used as a starting point to explore random effects models and the inclusion of environmental covariates on the catchability of selected survey indices.

The base statistical catch-at-age model (BF28W) and 6 state-space models (BF28W_m2 – BF28W_m7) with different options for treating the yearly transitions (survival) in recruitment and numbers-at-age were evaluated and compared (Table 7). Convergence diagnostics for each model run were examined and model selection via AIC was used to select a "best" model among the 6 models with comparable likelihood structures. Based on AIC selections, all of the top models were full state-space models, where survival of all ages were random effects with different correlation structures (Table 7). The model with the lowest AIC was BF28W_m7, which included correlation in the random effects by year and age (2D AR1). Model BF28W_m5 was very close in AIC but not within 2 AIC units of BF28W_m7 and was not considered equivalent based on model selection. Model BF28_m4 and BF28_m6 had similar model results but were noticeably higher in AIC.

Numbers-at-age deviations were correlated by age and year for the best model according to AIC, and were correlated by year for the next best model. The correlation by age was low and showed series of positive, negative and positive values from age-2 to age-4 in the middle of the time-series (Figure 28). The negative correlation between these ages is likely a result of the changing availability over time of this size class to the fisheries.

Results from the top 3 state-space models (BF28W_m7, BF28W_m5, and BF28W_m4) and the base statistical catch-at-age model (BF28W) showed good agreement among the model results (Table 7). The base model differed slightly in estimates of full F and SSB from 2008-2015 and in SSB again at the end of the time-series from 2016-2021, where SSB trended higher for this model (Figure 29). There were differences in the fleet selectivity block estimates, most notably with the base model in comparison to the state space models (Figure 30). In the final recreational selectivity block, the base model selectivity pattern was more domed, which likely resulted in the higher SSB estimates seen at the end of the time-series for this model. Index selectivity across the models showed differences mainly in those indices that catch older, larger bluefish. Those indices are the NEFSC Bigelow, PSIGNS, and ChesMMAP survey (Figure 31).

The final bluefish assessment model chosen by the working group was model BF28W_m7. A full presentation of parameter tables, input data, results, diagnostic, and retrospective plots are included in Working Paper 17 (Wood 2022d) and can also be downloaded or viewed separately from the following link: BF28W_m7_plots. A brief summary of results of the final model with selected plots are included below.

The final model fleet selectivity-at-age estimates for the two catch fleets showed a decrease in selectivity at middle ages (ages 3-5), with selectivity increasing at older ages. There was a decrease in the selectivity of these middle ages over time in the recreational selectivity blocks (Figure 32). Most of the index selectivities showed a domed selectivity after age-0. The MRIP CPUE index had a flat top logistic selectivity and was fully selected for the older ages. Both the NEFSC Bigelow index and the PSIGNS index had higher selectivity on the older, larger fish than the other fishery-independent indices (Figure 33).

Total abundance estimates from model BF28W_m7 peaked at a high of 599 million fish in 1985, declined to 162 million fish in 1995, and then increased to 269 million fish in 2005. Total abundance declined from 2005 to a low of 144 million in 2016, a small peak to 177 million in 2018, and a terminal year estimate of 162 million fish. Spawning stock biomass started from a high of 218,291 MT in 1985 and declined over the time-series to a low of 41,377 MT in 2018, and increased since then to 55,343 MT in 2021 (Figure 34). The majority of the spawning stock biomass is ages 5 and 6+ (30-60%) for the entire time-series. Fully selected fishing mortality in 2021 was 0.166, compared to an average full F from 1985 to 2021 of 0.309. Estimates of F have varied over the time-series from a peak in 2018 of 0.456 to the lowest value of 0.166 in 2021 (Figure 34). Estimates of recruitment remained stable over the time series, fluctuating around an average value of 128 million age-0 fish. Recruitment has been below average for the past 12 years, and was estimated at 87 million age-0 fish in 2021.

Retrospective pattern for the final model was examined for F, spawning stock biomass, and recruitment. Model BF28W_m7 exhibited a significantly improved retrospective pattern when compared to model BF24, the final ASAP model. The retrospective pattern was considered minor for SSB (Mohn's rho = 0.130), fishing mortality (Mohn's rho = -0.096), and recruitment (Mohn's rho = -0.063).

Model BF28W_m7 had excellent convergence diagnostics. Three sets of jitter analyses at increasing scale values of 1, 2, and 3 (the increase in scale broadens the distribution around the potential starting values by scaling the variance) were conducted. At a scale value of 1 (using variance estimates directly) 200/200 models converged at the original objective function. At a scale value of 2, all models converged, with 193/200 at the original objective function. Other objective function solutions were nearly identical to the original solutions (original objective function was 1468.54, other converged solutions were at 1468.69, 1468.72, and 1468.78). At a scale value of 3, all models converged with 155/200 jitter runs finding the original model solution and most of the other objective functions solutions very close to the original objective function (Figure 35). For comparison, the ASAP jitter analyses were only conducted at a scale of 1.

A historical retrospective analysis showing the model results from the 2015 benchmark assessment, 2021 operational assessment, BF01, the continuity run model, and BF28W_m7 (the final model) is presented in Figure 36.

4.3.2.4 Companion Model BF28WE: Environmental covariate on catchability of survey indices

One of the main reasons the bluefish assessment model was moved into WHAM was to explore the incorporation of environmental covariates on the catchability of different survey indices. Forage fish indices were developed using a VAST model (Section 1.2; Working Paper 4 Gaichas et al. 2022) and explored as environmental covariates on the catchability (q) of NEFSC survey indices and the MRIP CPUE index. These models are still under development and are being briefly presented as companion models for preliminary review. It is hoped that further exploration of these environmental models will lead to future improvements in the assessment.

The application of the forage fish indices as covariates on the catchability of the NEFSC science center surveys had mixed results. The forage fish index for the catchability of the NEFSC Albatross survey was explored as both a random walk and auto-regressive (AR1) process and each caused problems with the convergence of all models. Standard error around the covariate was explored using both the VAST estimated standard errors as an input standard error to the model, or allowing WHAM to estimate a single standard error of the covariate shared among time steps. All of the model runs either did not converge, or had issues with the hessian matrix calculations.

The forage fish index for the Bigelow survey did not have the same convergence issues as the Albatross index. The forage fish index was fit as a covariate on the Bigelow index catchability assuming a random walk over the time-series. All models with the forage fish covariate converged, but these models had worse fits than the base model according to AIC.

The application of the forage fish index to the MRIP CPUE index catchability was successful when implemented as an autoregressive (AR1) process over the time series with WHAM estimating a single shared standard error. The inclusion of the forage fish index improved the fit of all models (m2-m7), and model selection via AIC chose the time-varying catchability version of BF28W_m7 as the best model. This model will be referred to as model BF28W_m7ecov (where "ecov" refers to environmental covariate). Model BF28W_m7ecov had improved AIC of 2 units over BF28W_m5ecov, and by 5.6 units over BF28W_m7. The results from these top 3 models and the base model (BF28W) are presented in Figure 37.

A full presentation of parameter tables, input data, results, diagnostic, and retrospective plots are available for the best model BF28W_m7ecov are included in Working Paper 17 (Wood 2022d) and can also be downloaded or viewed separately from the following link: BF28W_m7ecov_plots.

The use of the forage fish index as a covariate on catchability led to an overall decreasing trend in catchability over time (Figure 38). The MRIP index is important in scaling the biomass results, and the lower availability at the end of the time-series led to higher recent biomass estimates from the environmental model. Spawning stock biomass started from a high of 181,804 MT in 1985 and declined over the time-series to a low of 52,697 MT in 2018, and increased since then to a value of 74,549 MT in 2021. Fully selected fishing mortality in 2021 was 0.126, compared to an average F from 1985 to 2021 of 0.271. Estimates of F have varied over the time-series from a peak in 1987 of 0.503 to the lowest value of 0.126 in 2021. Estimates of recruitment have remained stable over the time series, fluctuating around an average value of 143 million age-0 fish. Recruitment has been below average for the past 12 years, and was estimated at 106 million age-0 fish in 2021.

4.3.2.5 Model BF28W sensitivity analyses

A number of sensitivity runs of the final model (BF28W_m7) were explored. The model results and retrospective pattern results from each of these runs are presented in Table 8.

The sensitivity of the final model to the indices was explored in several ways. First, each index was removed individually, and the model was re-run to gauge the effect. Results from this series of models are in Table 8 and Figure 39. The final model was not overly sensitive to any single index, which was a shift from past bluefish assessment models. The bluefish assessment used to be heavily weighted towards the MRIP CPUE index. In many cases the model would not converge without this index included, or the model would scale the biomass to an unrealistic magnitude to find a model solution. This was no longer the case with model BF28W_m7, which converged without the MRIP CPUE index and found a solution that is in agreement with all the other index sensitivity runs (Figure 39). The model results appeared to be most sensitive to the removal of the PSIGNS index, which is an important index for tracking the abundance of older fish. Removal of this index significantly reduced SSB and increased F.

Two other index sensitivity runs were explored. Model based (GLM) versions of the indices (Working Paper 7 Celestino et al. 2022a) were substituted into the model for a sensitivity run. This change had very little impact on the model results and retrospective results.

Next, NEFSC indices that included some offshore strata were substituted into the model. The NEFSC survey encounters larger bluefish offshore in some years, and these "offshore" indices were explored to see impact of including bluefish observations from these offshore strata. The results from this sensitivity run were similar to the final model results, with both recruitment and SSB scaled upwards a small amount.

The next group of sensitivities focused on how recreational discard lengths were developed. First, a sensitivity was run that borrowed recreational discard (MRIP B2) lengths across regions, as opposed to using a cumulative length by season/region for years where the number of lengths sampled was less than 30. This sensitivity did not have good convergence properties and the hessian was not positive definite. This was due to changes in some of the fleet selectivity-at-age estimates, with some hitting the bound of 1.0. Further development could improve this model and results but were beyond the scope of a sensitivity analysis.

Next, recreational harvest lengths were borrowed for season/region years where the number of recreational discard lengths sampled was less than 30. The results of this model were nearly identical to the final model run. Finally, recreational length proportions from harvested fish (MRIP AB1) were used in place of the dead release lengths (instead of the i9, ALS, and VAS lengths). This model sensitivity also produced very similar results to the final model, with a slightly reduced recruitment and SSB, and slightly increased F (Figure 40).

Other sensitivity runs that were explored included:

1. Using the MRIP directed trip index instead of the Guild index
2. Setting the MRIP index to estimate selectivity-at-age instead of estimating a logistic curve
3. Assuming 15% recreational discard mortality instead of 9.4%
4. Assuming both the upper and lower confidence bounds for the Lorenzen M estimates

Results from each of these one-off sensitivities are presented in Table 8.

5 STATUS DETERMINATION CRITERIA

Term of Reference #5: Update or redefine status determination criteria (SDC; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

In a meeting of the Mid-Atlantic Fishery Management Council Scientific and Statistical Committee (SSC) following the 2015 benchmark assessment for bluefish, the SSC stated, "…the F_{MSY} proxy of $F_{40\%}$ might be inappropriate for Bluefish, a highly productive stock…". Citing two studies as support, the SSC used $F_{35\%}$ to set the overfishing limits for 2016-2018. The two papers the SSC cited (Rothschild et al. 2012; Thorson et al. 2012) were read and evaluated for support to use $F_{35\%}$ for bluefish in the 2022 assessment update. The WG agreed that the literature supported the use of $F_{35\%}$ for bluefish and continued the use of $F_{35\%}$ as the F_{MSY} proxy.

Many species managed in the Greater Atlantic region that use per-recruit reference points use $F_{40\%}$ (e.g., many groundfish species with analytical assessments, Atlantic herring, scup and black sea bass). However, bluefish is not the only example of a species that currently uses $F_{35\%}$ as summer flounder also uses this reference point.

5.1 Stock Status from the Continuity Run, BF01

Stock status was first determined using the continuity run model, which is the current accepted model for providing management advice, and would be used in absence of the research track assessment. Reference points were calculated using the non-parametric yield and SSB per-recruit long-term projection approach assuming 5-year averages for fishery selectivity, maturity and weights-at-age for SSB per recruit calculations. The cumulative distribution function of the 1985-2021 recruitment estimates were resampled to provide future recruitment estimates for the projections and used to estimate the SSB_{MSY} reference point associated with $F_{35\%}$ from a 100-year projection.

Existing reference points from the 2021 management track assessment (data through 2019) were F_{MSY} proxy = $F_{35\%}$ = 0.181 and SSB_{MSY} = 201,729 MT (1/2 SSB_{MSY} = $SSB_{THRESHOLD}$ = 100,865 MT). Updated reference points from the continuity run are F_{MSY} proxy = $F_{35\%}$ = 0.176 and SSB_{MSY} = 190,771 MT (1/2 SSB_{MSY} = $SSB_{THRESHOLD}$ = 93,386 MT).

A retrospective adjustment of the terminal year results for F and SSB resulted in these values being outside of their 90% MCMC confidence bounds. The retrospective pattern in F and SSB was considered major (SSB_{rho} = 0.29, F_{rho} = -0.28, based on 7-year peel) and required a retrospective adjustment to determine stock status. The 2021 retrospective adjusted F was 0.222 and falls above F_{MSY}. The 2021 retrospective adjusted value for SSB was 70,900 MT, and is lower than $SSB_{THRESHOLD}$. The results from the continuity run model indicate that the bluefish stock is overfished, and over-fishing is occurring (Figure 41). The over-fishing status has changed since the 2021 management track assessment. This change is a result of increased retrospective for F in the updated continuity run model, resulting in a retrospective adjustment that increased the terminal F value.

5.2 Stock Status from the Final Research Track Model, BF28W_m7

Both $F_{35\%}$ and $SSB_{35\%}$ were calculated internally in WHAM using average recruitment over the time series (1985-2021), and 5-year averages for fishery selectivity, maturity and weights-at-age for SSB per recruit calculations. The 5-year average was selected for those parameters to capture the most recent conditions while still smoothing some interannual variability; the full time-series of recruitment was chosen to fully capture the range of possible recruitment, given that there did not appear to be a significant regime shift in recruitment levels for bluefish over the time series. $F_{35\%}$ explicitly accounts for uncertainty from selectivity; $SSB_{35\%}$ explicitly accounts for uncertainty from selectivity and average recruitment. Uncertainty in the reference points associated with the 2D-AR1 process is implicitly accounted for through its impacts on selectivity and average recruitment. Additional sources of uncertainty in reference points not explicitly accounted for include uncertainty associated with the remaining SPR calculation inputs (e.g., natural mortality, maturity, and average weights-at-age).

Reference points from the final model (BF28W_m7) were F_{MSY} proxy = $F_{35\%}$ = 0.248 (95% CI: 0.209 – 0.299) and SSB_{MSY} proxy = $SSB_{35\%}$ = 91,897 MT (95% CI: 66,219–127,534 MT); $SSB_{THRESHOLD}$ =1/2 SSB_{MSY} proxy = 45,949 MT (95% CI: 33,110–66,768 MT). The retrospectively adjusted values of terminal year F and SSB were within the 90% confidence bounds of the unadjusted values, indicating a retrospective adjustment was not necessary to determine stock status (Figure 42). The terminal year SSB was 55,344 MT (95% CI: 35,185 – 87,052 MT) which is above the $SSB_{THRESHOLD}$ and 60% of SSB_{MSY}. Full fishing mortality was 0.166 (95% CI: 0.103 – 0.268) in 2021, which is 67% of the $F_{35\%}$ reference point (Figure 43). Accounting for uncertainty in reference points and terminal year F and SSB estimates, stock status determination based on the final model indicates that there is an 87% chance that the bluefish stock is currently not overfished and over-fishing is not occurring (Figure 44).

A comparison of stock status results from 2015 benchmark model, 2021 operational assessment, the current assessment continuity run, and the final model from this assessment is presented in Table 9.

6 PROJECTION METHODS

> *Term of Reference #6: Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions*

Short-term projections were conducted in WHAM, and incorporate model uncertainty, auto-regressive processes and uncertainty in recruitment and numbers-at-age. Removals in 2022 were assumed to be equal to the 2022 ABC (11,460 MT), and projections were carried forward for years 2023-2025 with different fishing mortality and harvest assumptions: F = 0, $F_{status\ quo}$ = 0.166, $F_{35\%}$ = 0.248, and that harvest in each year is equal to the acceptable biological catch (ABC) in each year. The annual ABC values were derived using projected OFL catch and applying the Mid-Atlantic Fishery Management Council (MAFMC) risk policy with an assumed OFL CV (MAFMC, 2020). In recent years, the ABC for bluefish has been developed using an OFL CV = 100%. Projections were carried out assuming an OFL CV of 100% and 60%.

Fishing at $F_{35\%}$ caused a decrease in biomass over the projected years, from 65,805 MT in 2022 to 61,784 MT in 2025 (Table 10). The catches associated with fishing at the reference point (OFL catch) ranged from 13,909 MT to 13,584 MT (Table 11). The probability of the stock being over the biomass threshold in 2025 was 0.84 for the $F_{35\%}$ projection.

The most realistic projections are the F status quo projection, and the MAFMC risk policy projection at an assumed CV of 100%. The risk policy approach is how management specifications are currently developed for bluefish. The probability of the stock being over the biomass threshold in 2025 was 0.93 for the F status quo projection, and 0.88 for the risk policy approach.

The projections use 5-year averages for natural mortality, maturity, fishery selectivity and weights-at-age. The 5-year average was selected for those parameters to capture the most recent conditions while still smoothing some interannual variability; the full time-series of recruitment was chosen to fully capture the range of possible recruitment, given that there did not appear to be a significant regime shift in recruitment levels for bluefish over the time series. Projections were not retrospectively adjusted, as the adjusted terminal year estimates of F and SSB fell within the 90% confidence intervals of the unadjusted values (Figure 42). The sensitivity of these projection assumptions were tested using 3-year, and 10 year averages. The projections are not overly sensitive to these assumptions. Assuming a 3-year averages leads to ~7.0% decrease in biomass, and ~6.0% decrease in catch when compared to the 5-year average. Assuming a 10-year average for the projection input results in a <1.0% difference in all results when compared to the 5-year average.

A final projection was carried out at $F_{rebuild}$. The bluefish stock is currently under a rebuilding plan, with a target date of 2028. $F_{rebuild}$ for the stock is currently set a 0.166 and a projection through 2028 was done assuming this value in each year. The 2028 SSB resulting from this projection is 79,215 MT, which is 86% of the biomass target (91,897 MT).

7 RESEARCH RECOMMENDATIONS

Term of Reference #7: Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 1 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

7.1 Status of Previous Research Recommendations

Some research recommendations were repeated in various documents (e.g., 2015 benchmark, SSC documents); for brevity, where the same, or substantially similar recommendations were made, we consolidated them under a single heading (e.g., 2015 benchmark), but noted the additional documents in which the recommendation was raised.

7.1.1 Research recommendations from SAW60 (NEFSC 2015)

7.1.1.1 High Priority

Recommendation: Determine whether NC scale data from 1985-1995 are available for age determination; if available, re-age based on protocols outlined in ASMFC (2011); if re-aging results in changes to age assignments, quantify the effects of scale data on the assessment

WG Response: The WG spoke with NC technical staff who endeavored to find the historical structures for ageing (in addition to 1985-1995 scale samples, otolith samples through 2000 were also included in the search). NC staff reached out to multiple additional agency staff at multiple offices throughout the state and the samples were not found.

Recommendation: Develop additional adult bluefish indices of abundance (e.g., broad spatial scale longline survey or gillnet survey); initiate fishery-dependent or fishery independent sampling of offshore bluefish populations to reduce reliance on MRIP sampling [also recommended in July 2015 and July 2021 SSC reviews]
WG Response: As part of the current research track assessment, ASMFC solicited data from state, federal, and academic partners as well as stakeholders via a press release and public data workshop. However, no new fishery-independent indices that capture very large fish, or offshore fish, were identified. The WG engaged with stakeholders to understand the extent of possible adult bluefish interactions with the offshore longline tilefish fishery, who indicated offshore interactions do occur with bluefish, but not consistently across states, and no clear trend was identified.

Recommendation: Expand age structure of SEAMAP index
WG Response: The WG added an age-1 index of abundance from the SEAMAP survey to the assessment; other ages classes were rarely encountered.

7.1.1.2 Moderate Priority

Recommendation: Investigate species associations with recreational angler trips targeting bluefish (on a regional and seasonal basis) to potentially modify the MRIP index used in the assessment model; Explore alternative definitions for targeting for calculating CPUE (e.g., directed trips or directed trips + incidental harvest) [also recommended in July 2015 and July 2021 SSC reviews]
WG Response: The WG developed an MRIP index using a species-association method to identify bluefish trips as well as a directed trips approach; see Section 3.2.5 and Working Paper 13 (Drew 2022d) for more details.

Recommendation: Explore age- and time-varying natural mortality from, for example, predator prey relationships; quantify effects of age- and time-varying natural mortality in the assessment model
WG Response: The WG evaluated a suite of life history and environmental data approaches to estimating age-constant and age-varying natural mortality and selected the Lorenzen (1996) age-varying approach for the base model; see Working Paper 6 (Tyrell and Truesdell 2022). The WG explored trends in large scale predator data (e.g., Shortfin mako, *Isurus oxyrinchus*) to potentially inform time-varying natural mortality, but concluded data were not sufficient to support a time-varying M at this time.

- *Next steps: If relevant predator abundance information becomes available in the future, a predator index could be used to inform time-varying natural mortality. The working group produced a condition index for bluefish of three size groups, which could be used to inform time-varying natural mortality in a WHAM model in the future; the working group prioritized using the forage fish index as a WHAM covariate in this research track assessment.*

Recommendation: Continue to evaluate the spatial, temporal, and sector-specific trends in bluefish growth and quantify their effects in the assessment model.

WG Response: The WG explored life history characteristics over various temporal and spatial scales (Working Paper 5 Truesdell et al. 2022), and constructed age-length keys and length-frequencies at the seasonal and regional level. While the age-length data were too sparse to support the season-region keys (Working Paper 14 Celestino 2022b), the catch length frequencies for both harvest and dead discards were stratified to the season and region level, an advance from the 2015 benchmark where only the harvest was stratified at that level.

Recommendation: Continue to examine alternative models that take advantage of length-based assessment frameworks. Evaluate the source of bimodal length frequency in the catch (e.g., migration, differential growth rates).
WG Response: The WG did not believe a length-based approach would improve the stock assessment for bluefish given the improvements in the age data collection across the coast and the longer time series of otolith-only data, and so did not pursue a length-based modelling approach. In addition, the WG did not have the type of data that would support a size transition matrix. The WG investigated whether the bimodal length frequency in the catch could be attributed to mid-size bluefish migrating to the Gulf of Mexico, but did not find support for this hypothesis. To some extent, with the development, expansion, and continuation of the coastwide biological collection program (Amendment I to the FMP), the bimodal pattern has become less frequent but has not disappeared; the WG suggests that this remains a research recommendation. Tagging programs (e.g., traditional, satellite) could provide additional insights.

Recommendation: Modify thermal niche model to incorporate water temperature data more appropriate for bluefish in a timelier manner [e.g., sea surface temperature data & temperature data that cover the full range of bluefish habitat (SAB and estuaries)].
WG Response: The 2015 analysis of the centers of biomass (COB) indicated that COB positions were correlated with variations in body size and abundance, but not temperature, and the annual proportion of thermal habitat suitability surveyed did not exhibit consistent, systematic trends. Therefore, the WG did not update the thermal niche model for this assessment, but included temperature as a covariate in the VAST forage fish index to serve a similar function as a covariate to inform catchability of the indices; see Working Paper 4 (Gaichas et al. 2022) for more details.

7.1.2 Research recommendations from SSC (July 2015)

Recommendation: Develop Bluefish-specific MSY reference points or proxies.
WG Response: Bluefish-specific MSY proxy reference points were developed for this assessment (see Section 5 above).

Recommendation: Low frequency environmental variability may have caused changes in the timing of the movement of juvenile Bluefish through the region that, in turn, may have affected availability. Changes in the selectivity of age-0 Bluefish in the survey relative to water column or surface temperature and date should be examined.
WG Response: The WG investigated the influence of temperature effects on bluefish as part of the ESP. See Section 1.1 and Working Papers 1 and 3 (Tyrell et al. 2022, Tyrell 2022) for more detail. However, more work, including additional bluefish data collection, needs to be done to incorporate this information into the assessment model framework in a quantitative way (see also

Section 7.2 below). The WHAM framework will allow for continued exploration and testing of covariates influencing time-varying catchability and selectivity.

- *Next steps: Additional survey data in the late fall would be needed to determine whether bluefish spawning is extending later in the year, which may be possible due to warmer temperatures extending later in the fall. Environmental covariates on recruitment could be incorporated into WHAM to test for improvements to model fit.*
- *Next steps: Further VAST models could be developed that incorporate additional scientific surveys, e.g., ChesMMAP and NEAMAP. The effect of environmental variability and timing of sampling could also be further investigated with VAST models, which can account for the day of sampling using a catchability covariate and can account for environmental variability using density covariates.*

Recommendation: Evaluate methods for integrating disparate indices produced at multiple spatial and temporal resolutions into a stock-wide assessment model, especially for a migratory species like Bluefish [also a July 2021 SSC review recommendation]
WG Response: The WG continued the use of the Conn (2010) approach to develop a single recruitment index from multiple state seine surveys as a means to addressing this research recommendation; see Section 3.1.2 and Working Paper 12 (Drew 2022c). The WG also explored using VAST to develop a forage fish index from multiple surveys (Section 1.2 and Working Paper 4 Gaichas et al. 2022) and to develop a standardized index with a single time series from the NEFSC Albatross and Bigelow vessels (Section 1.1.2 and Working Paper 3 Tyrell 2022); both approaches need more development before they can be incorporated into the base model of the assessment (see also Section 7.2 below).

- *Next steps: The bluefish Albatross-Bigelow VAST index could be further developed with environmental covariates (such as temperature). Additionally, multiple surveys could be combined in the VAST index.*

7.1.3 Research recommendations from SSC (July 2021)

Recommendation: A primary source of uncertainty is the recreational catch time series. The MRIP trend does not seem consistent with hypothesized reasons for differences between the mail and phone surveys. This historical correction to the MRIP estimates for bluefish should be explored further to evaluate the causes of differences from other species and to consider their plausibility.
WG Response: The WG examined differences in the calibrated and uncalibrated MRIP estimates of bluefish catch and found that while the magnitude of the calibration effect differed by mode and state, overall, we do generally see differences over time consistent with the hypothesized reasons for differences between the mail and phone surveys, and similar to trends in other mid-Atlantic species like summer flounder (*Paralichthys dentatus*), tautog (*Tautoga onitis*), and striped bass (*Morone saxatilis*). More detail is available in Working Paper 9 (Drew 2022a).

Recommendation: Investigate whether and how the selectivity pattern in discards has changed over time; the SSC questioned the methods for estimating the weight of recreational discards and the disparity between the use of volunteer angler data and the assumptions used by GARFO.
WG Response: For this assessment, the WG stratified released length frequency by region when calculating the weight of dead recreational releases to account for differences in the size

structure of removals and the release length samples between the regions. In addition, during the course of the present research track assessment, the WG communicated with the Greater Atlantic Regional Fisheries Office (GARFO) staff to ensure there is no longer a discrepancy between how the assessment estimates the weight of dead recreational releases and how that component is estimated for management; the agreed upon methods are consistent with other managed species (e.g., black sea bass, summer flounder).

Recommendation: Investigate patterns and trends in recent recruitments; the SSC noted low recruitment estimates in 2019 and asked whether it was possible to detect shifts between spring vs late summer recruiting cohorts.
WG Response: The WG's review of recruitment data largely suggested that data were not clear or sufficient to resolve whether there has been a shift between spring versus late summer recruiting cohorts; see Working Paper 1 (Tyrell et al. 2022) for a more detailed review of recruitment information available for bluefish.

- *Next steps: In order to quantitatively distinguish between spring-spawned and summer-spawned bluefish cohorts, regular seasonal sampling targeting small (<10cm) bluefish would need to be conducted over the broader Mid-Atlantic region and would have to extend later into the fall than current surveys.*

Recommendation: Long term environmental variability may have caused changes in the timing of the movement of juvenile Bluefish and the distribution of adults throughout the region that, in turn, may have affected availability.
WG Response: The WG explored development of VAST index of small pelagic fish aggregate abundance via predator diet data as a covariate for bluefish availability (Section 1.2 and Working Paper 4 Gaichas et al. 2022) and incorporating environmental covariates into index development via VAST (Section 1.1.2 and Working Paper 3 Tyrell 2022) and GLM-based standardization (Working Paper 7 Celestino et al. 2022a); both approaches need more development before they can be incorporated into the base model of the assessment (see also Section 7.2 below).

- *Next steps: More formal examination of time series changepoints and relationships of the forage indices with other ecosystem indicators will be explored during the NEFSC's 2023 State of the Ecosystem report development cycle, and can be included in future bluefish assessments.*
- *Next steps for Albatross-Bigelow VAST: there are additional VAST model changes that can be explored to better understand the influence of environmental covariates on bluefish distribution. The VAST model presented in this report could be further developed to successfully incorporate environmental covariates such as temperature.*

7.2 New Research Recommendations

7.2.1 High Priority

Expand collection of recreational release length frequency data
Recreational release mortality accounts for approximately 15% of total removals in weight in recent years, but information on the size structure of released fish is limited, particularly in the South Atlantic. The assessment now stratifies length frequency of released fish by region, but requires borrowing across years with low sample sizes (n<30), and this borrowing should be minimized or avoided where possible to better capture year class effects. Expansion and

promotion of volunteer angler survey programs would be one option to reduce this source of uncertainty in the assessment.

Continue development and refinement of the forage fish / availability index as well as incorporation of this index into a base model for bluefish management advice
Preliminary modelling that incorporated the forage fish index suggested an improved model fit relative to a model without the index. The forage fish index could provide information on availability of bluefish to different surveys and fisheries, and could potentially help the model resolve conflicts between indices that occur more offshore and indices that occur more inshore. Additional work could include:

- Investigate sources of piscivore diet data for "inland waters" (Chesapeake Bay, Delaware Bay, Long Island Sound) to integrate into the model, potentially providing more insight into availability to the MRIP index. (ChesMMAP has diet data; other surveys or studies should also be investigated)
- Investigate sources of piscivore diet data south of Cape Hatteras to expand to full bluefish range
- Investigate other potential environmental covariates (e.g., higher resolution SST)
- Continue modelling within the WHAM framework to resolve issues identified in TOR4
- Continue to explore environmental linkages to catchability, selectivity, recruitment, and natural mortality using WHAM

Initiate additional fisheries independent surveys and/or fishery-dependent sampling programs to provide information on larger, older bluefish
This remains a high priority given the limited information on older (e.g., age 2+) bluefish collected by existing fishery independent surveys. This item addresses the need to adequately characterize dynamics of older fish that are currently not well sampled by fishery independent trawl surveys, as well as to understand the extent of summer and fall spawning and the contribution of these fish to year classes. This item also would help address unresolved issues identified above (e.g., relative cohort strengths, offshore movements, environmental effects). Further engagement with stakeholders can help identify areas of incidental bluefish catch in offshore fisheries and inform potential development of voluntary or required reporting programs and data sources.

Continue coastwide collection of length and age samples from fishery-dependent and fishery-independent sources.
The availability of bluefish to different fisheries varies throughout the year along the coast; in order to accurately characterize the age-structure of the removals, adequate samples, stratified spatially and temporally, need to be collected. The increased sampling at the state level as a result of Amendment 1 to the Bluefish FMP improved the available data and reduced gaps in the ALK. Current sampling levels should be maintained at a minimum.

7.2.2 Medium Priority

Further index of abundance development and refinement

The large number of indices input into the model sometimes provide conflicting signals and add additional parameters that need to be estimated. Exploring environmental drivers of bluefish distribution and exploring index consolidation using VAST or other modeling approaches could provide more coherent indices and/or provide information on catchability covariates to resolve conflicting signals and improve model fits.

Develop a recreational economic demand model

Recreational demand models can inform managers of the likely economic and biological implications of alternative regulatory and stock conditions. Given the large role recreational effort plays in the bluefish fishery, efforts in developing recreational demand models should be prioritized to develop measures that will meet both biological and socioeconomic goals for the bluefish fishery (Appendix 1 of Working Paper 1 Tyrell et al. 2022).

7.2.3 Low Priority

Development of an updated recreational release mortality study

Given the importance of recreational releases in the bluefish fishery for both accurately estimating total catch and therefore population scale and in the correct allocation of dead catch to the commercial and recreational sectors, reducing uncertainty on the release mortality estimate is important, especially if it has changed over time with changing angler behavior. The WG discussed: (1) examination of release mortality based on study factors (hook type, fish length, etc.), and (2) a comparison of release mortality estimates generated from a variety of methods.

A coordinated tagging program to help understand migration patterns potentially contributing to patterns in length frequency distributions.

To the extent that spatiotemporal variation in availability is contributing to the bimodal length frequency distribution, this could help resolve a source of uncertainty in the assessment. The WG was not able to resolve the source of bimodal length frequency distributions and has hypothesized offshore migration or summer/fall residency in southern waters makes those size classes of bluefish unavailable to fisheries and surveys. A coordinated fishery-independent tagging program could also help to more definitively resolve migrations between the Gulf of Mexico and Atlantic coast. A coordinated program could also address the release mortality recommendation above and provide a different source of growth information to compare to age-based methods.

Commercial discard length frequency data

There are currently no length data to characterize the length frequency of commercial discards. This source of mortality is small relative to other sources of fishing mortality, but does represent a source of uncertainty.

8 BACKUP ASSESSMENT APPROACH

> *Term of Reference #8: Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.*

A backup assessment approach is required to be in place as a hedge against a scenario where the primary catch-at-age model fails peer review. Such alternative models could include biomass dynamic-type models (e.g., as used for red crab), swept area approaches (e.g., witch flounder), catch curves, index-based methods (e.g., Georges Bank Atlantic cod) or other approaches. In one case, a statistical catch-at-age approach was put forward as the primary model and modifications that were still in the statistical catch-at-age framework were suggested as a backup approach (e.g., American Plaice Research Track).

The Working Group chose the index-based method Ismooth (previously known as PlanBSmooth; see Chris Legault's GitHub repository for more information) as the backup model due to its performance in the analyses performed by the Index Based Model Working Group (NEFSC 2020) and because it has a history of application at the NEFSC as an approach that has been used to develop ABCs (e.g., Georges Bank cod, Gulf of Maine / Northern Georges Bank and Southern Georges Bank / Mid-Atlantic monkfish).

In general, this approach applies recent trends in an index or indices to recent dead catch to generate ABC advice. There are two steps in the process. The model calculates an average of normalized indices that are selected for inclusion, applies a loess smooth to those values, fits a linear model to the final three years of log-transformed smoothed data, and extracts the slope of the fit. The results are then applied to recent dead catch levels. In this case, the previous three years of dead catch are averaged and the Ismooth exponentiated slope is multiplied by that average to generate the advice.

Ismooth was one of a number of data poor approaches examined by the NEFSC Index Based Methods Working Group (NEFSC 2020, Legault et al., 2023). The primary focus of this Working Group was to quantify the performance of data poor approaches in circumstances that led to severe retrospective errors in statistical catch-at-age models. That group found that none of the data poor methods outperformed a retrospectively adjusted catch-at-age model over the long-term, but also concluded that the Ismooth approach performed reasonably well relative to other methods, especially with respect to maintaining an acceptable level of SSB and constraining F. Thus, the Ismooth approach represents a reasonable choice if the statistical catch-at-age model were to fail.

The WG simulated the performance of Ismooth relative to historical bluefish ABCs that were based on results of the ASAP model; see Working Paper 18 (Truesdell 2022) for additional information. In general, the retrospective advice calculated by the Ismooth model was correlated with the actual ASAP-derived ABCs that were recommended for management use by the SSC, especially when the MRIP index was included when developing the Ismooth advice (Figure 45). Accordingly, as a one-off ABC tool (i.e., when differences between approaches do not compound over time), Ismooth offers similar advice to the previously accepted statistical catch-at-age model (ASAP) given the historical indices that were used to compile the Ismooth estimate.

The WG explored other data-limited approaches for estimating sustainable yield including Depletion-Corrected Average Catch (DCAC; MacCall 2009) and Depletion-Based Stock Reduction Analysis (DBSRA; Dick and MacCall 2011) as was done in the previous benchmark

(NEFSC 2015). However, McCall (2009) noted that the DCAC method is not recommended for species where natural mortality is greater than approximately 0.20. Because of that, DCAC was dropped from further consideration as an alternative model for bluefish given the updated natural mortality rate. The DBSRA model produced significantly higher estimates of biomass and sustainable yield than the age-structured model and had a low rate of accepted runs when parameterized with updated bluefish life history information; because of this and concerns about the underlying surplus production model framework of the DBSRA, the WG did not recommend this approach for providing alternate catch advice.

Swept area approaches were also investigated but given the importance of the recreational sector a method that could incorporate the MRIP index was preferable. In addition, bluefish catchability and selectivity in trawl nets is not well understood which would decrease confidence in a trawl survey-only swept area approach. Catch curves were considered but were not recommended by the WG as a backup approach, as the WG did not know of other assessments that used catch curves to produce catch advice.

The WG does not anticipate the need for the backup approach to be applied. The scenarios in which the selected catch-at-age model would be abandoned are limited to new data that caused complete convergence failure, the discontinuation of critical data streams or logistical issues that precluded the model fitting process altogether. In the case of severe retrospective errors, the Index Based Model Working Group found that a retrospectively adjusted statistical catch-at-age model did not perform worse than index or data poor approaches, so this potential issue is not expected to cause a transition to the backup assessment. If new data causes major issues with model fitting, modifications within the catch-at-age framework (e.g., data weighting, random effect structure, etc.) would be exhausted before moving to the backup assessment approach. Such changes could be implemented through an Expedited or Enhanced management track peer review.

9 REFERENCES

ASMFC. 2011. Proceedings of the Atlantic States Marine Fisheries Commission Bluefish Ageing Workshop. ASMFC, Alexandria, VA. 26p. Available online at: http://www.asmfc.org/uploads/file/bluefishAgeingWorkshopReport_May2011.pdf

ASMFC. 2019. Atlantic States Marine Fisheries Commission, Weakfish Stock Assessment Update report. Arlington, VA. 86 pp. Available online at: http://www.asmfc.org/uploads/file/5de7fc7c2019WeakfishAssessmentUpdate.pdf

ASMFC. 2020. Memorandum from the bluefish Technical Committee to the bluefish Management Board, regarding recommendations on bluefish age and length frequency data sampling. 5 p. Available online at: http://www.asmfc.org/uploads/file/6022ea5dBluefishTC_Report_Sept2020.pdf

ASMFC. 2021. Report from the 2021 Bluefish State Data Review Workshop. Arlington, VA. 115p. Available online at: https://www.asmfc.org/uploads/file/63487cfe2021BluefishStateDataReviewReport.pdf

Avarovitz, T.R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. Pages 62–67 In: W.G. Doubleday and D. Rivard (editors), Bottom trawl surveys. Canadian Special Publication of Fisheries and Aquatic Sciences 58. 273 p.

Carvalho, F., H. Winker, D. Courtney, M. Kapur, L. Kell, M Cardinale, M. Schirripa, T. Kitakado, D. Yemane, K. R. Piner, and M. N. Maunder. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research 240, p.105959.

Conn, PB. 2010. Hierarchical analysis of multiple noisy abundance indices. Canadian Journal of Fisheries and Aquatic Sciences 67:108-120.

Dick, E. J. and A. D. MacCall. 2011. Depletion-based stock reduction analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110(2): 331-341.

Gerritsen, H.D., McGrath, D. and Lordan, C. 2006. A simple method for comparing age-length keys reveals significant regional differences within a single stock of haddock (*Melanogrammus aeglefinus*). ICES Journal of Marine Science, 63: 1096-1100.

Hewitt, D. A., and J. M. Hoenig. 2005. Comparison of two approaches for estimating natural mortality based on longevity. Fishery Bulletin 103 (2): 433.

Klein-MacPhee, G. 2002. Bluefish: family Pomatomidae. In Bigelow and Schroeder's fishes of the Gulf of Maine (B. B. Collette, and G. Klein-MacPhee, eds.), p. 400–406. Smithsonian Institution Press, Washington, DC.

Legault, C. M. 2012. User manual for ASAP 3. 22 p.

Legault, C. M. and V. Restrepo. 1999. A flexible forward Age-Structured Assessment Program. ICCAT Coll. Vol. Sci. Pap. 49.

Legault, C. M. J Wiedenmann, J. J. Deroba, G. Fay, T. J. Miller, E. N. Brooks, R. J. Bell, J. A. Langan, J. M. Cournane, A. W. Jones, and B Muffley. 2023. Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 80 (1): 27-42.

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): 627–42.

MAFMC. 2020. OMNIBUS Acceptable Biological Catch and Risk Policy Framework Adjustment. Environmental Assessment. 116 pp. Document available at this link.

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

Methot, R.D. and Wetzel, C.R. (2013). Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86-99. https://doi.org/10.1016/j.fishres.2012.10.012

Miller, T. J., Hare, J. A., and Alade, L. A. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to Southern New England yellowtail flounder. Canadian Journal of Fisheries and Aquatic Sciences 73(8): 1261-1270. Doi: 10.1139/cjfas-2015-0339

Miller, T. J. and Hyun, S-Y. 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. Canadian Journal of Fisheries and Aquatic Sciences 75(5): 691-703. Doi: 10.1139/cjfas-2017-0035

Miller, T. J., O'Brien, L., and Fratantoni, P. S. 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of Georges Bank

Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences 75(12): 2159-2171. Doi: 10.1139/cjfas-2017-0124
Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science 56(4): 473-488.
Munch, S. 1997. Recruitment dynamics of bluefish, *Pomatomus saltatrix*, on the continental shelf from Cape Fear to Cape Cod, 1973-1995. M.S. thesis, State University of New York at Stony Brook, NY, 127 p.
National Research Council (NRC). 2006. Review of Recreational Fisheries Survey Methods. Washington, DC: The National Academies Press.
NEFSC. 2015. 60th Northeast Regional Stock Assessment Workshop (60th SAW) assessment report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-08; Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026 http://doi.org/10.7289/V5W37T9T
NEFSC. 2020. Draft report of the Index Based Methods Working Group, submitted for peer review November 20th 2020. 128 pp. Report available at this link.
Ng, E. L., J. J. Deroba, T. E. Essington, A. Grüss, B. E. Smith, and J. T. Thorson. 2021. Predator stomach contents can provide accurate indices of prey biomass. ICES Journal of Marine Science 78(3):1146–1159.
Norcross, J. J., S. L. Richardson, W. H. Massmann, and E. B. Joseph. 1974. Development of young bluefish (Pomatomus saltatrix) and distribution of eggs and young in Virginian coastal waters. Transactions of the American Fisheries Society 103(3): 477-497.
Olla, B.L. and A.L. Studholme. 1971. The effect of temperature on the activity of bluefish, *Pomatomus saltatrix* L. Biological Bulletin 141:337-349.
Perretti, C. T., and J. T. Thorson. 2019. Spatio-temporal dynamics of summer flounder (Paralichthys dentatus) on the Northeast US shelf. Fisheries Research 215:62–68.
Papacostas, K. J., and J. Foster. 2020. A replication approach to controlled selection for catch sampling intercept surveys. Fisheries Research 229:1-9.
Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax. 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. Journal of Climate 20(22):5473–5496.
Ross, M.R. 1991. Recreational Fisheries of Coastal New England. University of Massachusetts Press, Amherst, MA.
Rothschild, B. J., Y. Jiao, and S. Y. Hyun. 2012. Simulation study of biological reference points for Summer Flounder. Transactions of the American Fisheries Society 141(2):426-436
Schoener, T. W. 1970. Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. Ecology 51(3):408–418.
Shepherd, G.R. and D.B. Packer. 2006. Essential Fish Habitat Source Document: Bluefish, *Pomatomus saltatrix*, Life History and Habitat Characteristics 2nd edition. NOAA Technical Memorandum, NMFS-NE-198:100.
Stock, B. C., and Miller, T. J. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research, 240: 105967. Doi: 10.1016/j.fishres.2021.105967
Taylor, D.L., Rowe, P.M., and K.W. Able. 2006. Habitat use of the inner continental shelf off southern New Jersey by summer spawned bluefish (*Pomatomus saltatrix*). Fishery Bulletin 104:593-604.

Taylor, D.L., Nichols, R.S., and K.W. Able. 2007. Habitat selection and quality for multiple cohorts of young-of-the-year bluefish (*Pomatomus saltatrix*): Comparisons between estuarine and ocean beaches in southern New Jersey. Estuarine, Coastal and Shelf Science 73:667-679.

Then, A. Y., J. M. Hoenig, N. G. Hall, D. A. Hewitt, and Handling editor: E. Jardim. 2015. 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.

Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210:143–161.

Thorson, J. T., and L. A. K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Science 74(5):1311–1321.

Thorson, J. T., J. M. Cope, T. A. Branch, and O. P. Jensen. 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, J. T., and K. Kristensen. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fisheries Research 175:66–74.

Trijoulet, V., C. M. Albertsen, K. Kristensen, C. M. Legault, T. J. Miller, and A. Nielsen. 2023. Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research 257, p.106487.

Wigley S. E., Rago P. J., Sosebee K. A., Palka D. L. 2007. The analytic component to the Standardized Bycatch Reporting Methodology Omnibus Amendment: sampling design end estimation of precision and accuracy (2nd edition). U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc . 07-09; 156 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

Wuenschel, M.J., Able, K.W., Buckel, J.A., Morley, J.W., Lankford, T., Branson, A.C., Conover, D.O., Drisco, D., Jordaan, A., Dunton, K., and D.H. Secor. 2012. Recruitment patterns and habitat use of young-of-the-year bluefish along the United States east coast: insights from coordinated coastwide sampling. Reviews in Fisheries Science 20:80-102.

10 TABLES

Table 1. Maturity-at-age through age-6 as calculated using various approaches. "Benchmark 2015" refers to the ogive used in the previous assessment, "NMFS 2022" refers to analyses performed during the 2022 research track assessment using data through 2021 but from federal sources only, and "Midyear model" refers to the GLM that was fit using federal and state data together.

Age	Benchmark 2015	NMFS 2022	Midyear model
0	0.00	0.000	0.000
1	0.40	0.417	0.456
2	0.97	0.965	0.926
3	1.00	0.999	0.995
4	1.00	1.000	1.000
5	1.00	1.000	1.000
6+	1.00	1.000	1.000

Table 2. Total removals of bluefish in metric tons by sector, 1985-2021.

Year	Commercial Landings	Commercial Discards	Recreational Landings	Recreational Dead Releases	Total Catch
1985	6,124		47,754	1,045	54,923
1986	6,657		75,470	1,611	83,738
1987	6,579		64,160	2,012	72,750
1988	7,162		36,475	905	44,542
1989	4,740	29	36,464	1,279	42,511
1990	6,250	32	31,553	1,976	39,811
1991	6,138	116	26,766	2,486	35,506
1992	5,208	38	22,533	1,769	29,548
1993	4,819	32	16,396	2,369	23,617
1994	4,306	162	14,176	3,140	21,783
1995	3,629	81	13,381	2,516	19,607
1996	4,213	166	10,760	2,756	17,895
1997	4,113	53	12,638	3,640	20,444
1998	3,741	74	15,414	2,995	22,224
1999	3,335	79	10,695	6,863	20,972
2000	3,660	83	11,141	6,289	21,174
2001	3,956	23	15,121	7,271	26,370
2002	3,116	37	13,904	4,581	21,638
2003	3,361	22	15,053	2,120	20,556
2004	3,673	62	17,570	4,744	26,050
2005	3,213	26	17,945	4,055	25,239
2006	3,354	34	16,912	5,708	26,009
2007	3,390	27	18,382	5,815	27,614
2008	2,731	22	17,410	5,428	25,591
2009	3,119	33	18,339	4,767	26,258
2010	3,304	87	21,269	6,384	31,044
2011	2,454	95	15,706	3,815	22,070
2012	2,212	14	15,291	2,833	20,350
2013	1,977	12	15,732	2,472	20,194
2014	2,251	18	12,324	2,880	17,473
2015	1,917	14	13,725	3,689	19,345
2016	1,946	14	10,634	1,837	14,431
2017	1,876	7	15,620	1,793	19,297
2018	1,105	8	5,857	1,579	8,548
2019	1,359	10	6,800	1,702	9,871
2020	1,112	9	5,923	1,253	8,296
2021	1,090	12	5,471	1,391	7,963

Table 3. Fishery-independent indices accepted by the Bluefish Working Group. "In Conn" indicates the index is part of the composite YOY index.

State	Index	Used in 2015?	Use in 2022?
NH	NH Seine Survey	Yes (in Conn)	Yes (in Conn)
RI	Beach seine (Narragansett Bay)	Yes (in Conn)	Yes (in Conn)
-	NEFSC Fall Trawl Survey	Yes	Yes
NY	WLIS Seine Survey	Yes (in Conn)	Yes (in Conn)
NJ	DE R. Seine Survey	Yes (in Conn)	Yes (in Conn)
MD	Striped Bass Seine Survey	Yes (in Conn)	Yes (in Conn)
VA	NEAMAP	Yes	Yes
VA	ChesMMAP	No	Yes
VA	Juv. Striped Bass Seine	Yes (in Conn)	Yes (in Conn)
NC	PSIGNS	Yes	Yes
SC	SEAMAP	Yes	Yes

Table 4. Fishery-independent surveys analyzed and excluded by the Bluefish Working Group.

State	Index	Used in 2015?	Use in 2022?	TC Comments	WG Comments
MA	MA Inshore Trawl Survey	No	No	Explore additional standardization; consider as a YOY index if trawl Conn dataset is expanded	Trawl Conn not used
RI	Trawl – Seasonal	No	No	Explore additional standardization; consider as a YOY index if Conn dataset is expanded	Trawl Conn not used
CT	Long Island Sound Trawl Survey	Yes	No	Use again in 2022	Remove; the index is dominated by age-0 fish, covers a limited spatial area, and was poorly fit by the model
NY	Peconic Bay Trawl	No	No	Explore additional standardization; consider as a YOY index if Conn dataset is expanded	Trawl Conn not used
NJ	NJ Ocean Trawl	Yes	No	Revise strata choice, standardization; consider as YOY index	Trawl Conn not used
DE	30' Trawl	No	No	Explore additional standardization	Limited spatial coverage for recruitment index; trawl Conn not used
MD	Coastal Bays Juvenile Trawl Survey	No	No	Explore as part of trawl composite YOY survey	Trawl Conn not used
NC	IGNS	No	No	Consider River Regions data to expand spatial extent of PSIGNS	Trends in other regions the same as PSIGNS; not worth shortening the time series
NC	P195	No	No	Explore additional standardization; consider as a YOY index if Conn dataset is expanded	Trawl Conn not used

Table 5. Model outputs and diagnostics from ASAP runs using various temporal and spatial levels of ALK and data resolution. ρ = Mohn's rho

ALK	2021 SSB (mt)	Recruitment (millions of fish)	F	SSB ρ	R ρ	F ρ	# at initial objective function (out of 200)	# unique objective function solutions	# Not converged (out of 200)
Traditional-Seasonal	59,540	28.7	0.19	0.341	0.080	-0.23	190	2	4
Multinomial-Seasonal	51,562	27.4	0.19	0.215	0.024	-0.18	193	2	3
Multinomial-Season-Region	43,916	27.1	0.21	0.222	0.033	-0.19	192	2	4

Table 6. Model table showing linear steps in the ASAP and WHAM model bridge building process. R is recruitment (in millions of age-0 fish). "W" in the model names indicates WHAM model runs. ~ indicates jitter analysis was not performed for that run. Ρ is Mohn's rho measure of retrospective patterning.

Model	Description	2021 SSB (MT)	2021 R (mil)	2021 F	SSB ρ	R ρ	F ρ	# at OG OBFunc	Jitter Sol	Not conv
BF00	BLF 2021 MT model	95,742	27.9	0.172	0.226	0.192	-0.221	~	~	~
BF01	BLF RT Continuity Run	91,745	39.4	0.160	0.294	0.170	-0.277	132	51	18
BF03	Update all new data	85,975	39.2	0.172	0.364	0.174	-0.323	142	38	21
BF04	New LW parameters	86,581	39.1	0.172	0.359	0.174	-0.320	~	~	~
BF05	New Rec discard mortality	82,103	35.9	0.159	0.380	0.186	-0.334	~	~	~
BF07	Add commercial discards	82,018	36.2	0.160	0.378	0.185	-0.332	140	28	31
BF08	New Indices: MRIP Continuity	88,424	35.3	0.158	0.319	0.123	-0.313	129	17	52
BF09	New Indices: MRIP Continuity, multinomial ALKs	70,336	26.3	0.158	0.352	0.042	-0.321	193	2	6
BF10	New Indices: MRIP Continuity, multinomial ALKs, Rec discard length by season/region	67,029	26.7	0.138	0.405	0.051	-0.361	197	2	2
BF11	New Indices: MRIP Guild, multinomial ALKs, Rec discard length by season/region	47,734	25.8	0.172	0.253	0.033	-0.214	191	2	8
BF12	New M: Lorenzen based on empirical WAA	65,946	97.3	0.110	0.346	0.113	-0.266	188	3	9
BF18	5 Sel blocks	79,849	97.9	0.116	0.293	0.035	-0.220	183	2	9
BF19	Fix bounded selectivities F2to3	82,858	98.6	0.113	0.288	0.014	-0.221	194	2	5
BF20	MRIP PSE for fleet 2	91,149	101.7	0.107	0.257	0.023	-0.205	194	4	2
BF21	MRIP index input CV (from 0.3)	84,212	85.3	0.116	0.193	0.000	-0.223	191	3	6
BF22	No CT survey	88,051	93.1	0.111	0.187	-0.001	-0.229	193	2	6
BF23	Adjust MRIP CV to reduce RMSE (x1.6)	94,886	102.0	0.102	0.225	-0.014	-0.209	199	1	1
BF24	Adjust fixed selectivity at age 2 for some blocks	63,320	94.6	0.159	0.326	0.017	-0.277	192	4	4
BF26W	Reduce fleet 1 CV	63,606	95.7	0.160	0.270	-0.066	-0.215	~	~	~
BF27W	Reduce Fleet 2 CV	68,546	96.4	0.152	0.249	-0.062	-0.198	~	~	~
BF28W	Reduce both fleets CV	68,631	96.4	0.152	0.248	-0.063	-0.197	~	~	~

Table 7. Results and diagnostics for different state-space model variations of the WHAM model BF28W examining different options for treating the yearly transitions (survival) in recruitment and number-at-age. R is recruitment (in millions of age-0 fish). Ρ is Mohn's rho measure of retrospective patterning.

Model	Description	dAIC	AIC	2021 SSB (MT)	2021 R (mil)	2021 F	R ρ	SSB ρ	F ρ	Con-verged?	Positive definite Hessian?
BF28W	Base model: traditional statistical catch-at-age	~	~	68,631	96.4	0.152	-0.063	0.248	-0.197	TRUE	TRUE
m7	All NAA transitions are random effects correlated by year and age	0	3229	55,344	86.5	0.166	0.010	0.130	-0.096	TRUE	TRUE
m5	All NAA transitions are random effects correlated by year	3	3232	55,070	82.3	0.167	0.019	0.126	-0.097	TRUE	TRUE
m4	All NAA transitions are random effects independent, identically distributed	46.2	3275	58,114	98.6	0.160	-0.008	0.172	-0.144	TRUE	TRUE
m6	All NAA transitions are random effects correlated by age	46.9	3276	58,786	99.9	0.159	-0.004	0.177	-0.148	TRUE	TRUE
m2	Recruitment transitions are random effects independent, identically distributed	111	3340	73,843	104.1	0.144	-0.022	0.236	-0.195	TRUE	TRUE
m3	Recruitment transitions are random effects correlated by year	111	3340	72,329	101.3	0.146	-0.020	0.245	-0.198	TRUE	TRUE

Table 8. Results, retrospective, and convergence properties of the final model sensitivity runs for the WHAM final model (BF28W_m7, bolded row). R is recruitment (in millions of age-0 fish). ρ is Mohn's rho measure of retrospective patterning.

Model	2021 R (mil)	2021 SSB (MT)	2021 F	R ρ	SSB ρ	F ρ	Con-verged?	Positive Definite Hessian?
BF28W_m7	**86.5**	**55,344**	**0.166**	**0.010**	**0.130**	**-0.096**	**TRUE**	**TRUE**
rmALB	83.1	53,880	0.171	0.0075	0.127	-0.0924	TRUE	TRUE
rmBIG	86.6	56,327	0.163	0.0281	0.1222	-0.0885	TRUE	TRUE
rmMRIP	101.4	64,964	0.142	-0.0233	0.176	-0.0993	TRUE	TRUE
rmNEA	81.3	57,488	0.162	0.0071	0.1326	-0.0995	TRUE	TRUE
rmSEA0	90.3	55,826	0.165	0.0045	0.1266	-0.0932	TRUE	TRUE
rmPSIGN	75.1	38,725	0.236	0.0635	0.25	-0.1689	TRUE	TRUE
rmYOY	77.4	53,209	0.175	0.0278	0.1473	-0.1118	TRUE	TRUE
rmCHES	86.0	54,749	0.168	0.0048	0.1256	-0.0908	TRUE	TRUE
rmSEA1	86.9	55,633	0.165	0.0116	0.1316	-0.0983	TRUE	TRUE
NEFSC offshore	97.6	59,020	0.169	0.046	0.128	-0.094	TRUE	TRUE
GLM indices	90.5	57,758	0.158	0.0535	0.1513	-0.1155	TRUE	TRUE
MRIP direct	101.4	71,334	0.131	0.004	0.130	-0.096	TRUE	TRUE
MRIP SAA	86.8	60,378	0.165	0.007	0.121	-0.093	TRUE	**FALSE**
Borrow Region	83.4	95,775	0.130	0.090	0.204	-0.132	TRUE	**FALSE**
Borrow AB1	83.6	55,473	0.172	0.0066	0.141	-0.1023	TRUE	TRUE
Use AB1 for B2	80.9	53,674	0.186	0.0071	0.1326	-0.0921	TRUE	TRUE
B2 15% DM	93.9	58,842	0.172	0.0103	0.1244	-0.0935	TRUE	TRUE
M Lorenzen Low	31.4	42,296	0.226	0.0007	0.1316	-0.1003	TRUE	**FALSE**
M Lorenzen High	480.3	205,189	0.045	0.2398	0.4017	-0.2348	TRUE	TRUE

Table 9. Biological reference points from the 2015 benchmark, 2021 operational assessment, the continuity run (BF01), and the final model (BF28W_m7).

Reference Point	SAW60	OA2019	BF01: Cont Run	BF28W_m7: Final Model
$F_{35\%}$	0.190	0.181	0.176	0.248
SSB_{TARGET}	101,343 MT	201,729 MT	190,771 MT	91,897 MT
$SSB_{THRESHOLD}$	50,672 MT	100,865 MT	93,386 MT	45,949 MT

Table 10. Short-term (2022-2025) projections of SSB and the probability of being above $B_{THRESHOLD}$ in 2025 for bluefish under 3 different F scenarios.

Projection scenario	2022	2023	2024	2025	P (2025) > $B_{threshold}$
F_{MSY} = 0.248	65,805 (39,305-110,170)	66,340 (37,604-117,034)	64,083 (35,017-117,275)	61,784 (32,086-118,971)	0.84
F_0 = 0	65,805 (39,305-110,170)	72,637 (41,394-127,462)	83,806 (46,270-151,792)	94,956 (49,788-181,098)	0.99
F_{status_quo} = 0.166	65,805 (39,305-110,170)	68,357 (38,820-120,367)	70,009 (38,411-127,601)	71,150 (37,110-136,412)	0.93
MAFMC risk policy (60% CV)	65,805 (39,305-110,170)	67,891 (37,217-123,847)	68,583 (33,654-139,765)	68,804 (29,551-160,198)	0.85
MAFMC risk policy (100% CV)	65,805 (39,305-110,170)	68,514 (37,767-124,295)	70,385 (35,116-141,078)	71,553 (31,586-162,089)	0.88

Table 11. Short term (2022-2025) projections of total catch for bluefish under 3 different F scenarios.

Projection scenario	2022	2023	2024	2025
F_{MSY} = 0.248	11,460	13,909 (8,098-23,889)	13,957 (7,784-25,022)	13,584 (7,157-25,784)
F_0 = 0	11,460	0	0	0
F_{status_quo} = 0.166	11,460	9,569 (5,564-16,458)	10,127 (5,628-18,223)	10,292 (5,399-19,623)
MAFMC risk policy (60% CV)	11,460	10,581 (P* = 0.311)	11,118 (P* = 0.314)	11,202 (P* = 0.316)
MAFMC risk policy (100% CV)	11,460	9,225 (P* = 0.311)	10,027 (P* = 0.321)	10,357 (P* = 0.327)

11 FIGURES

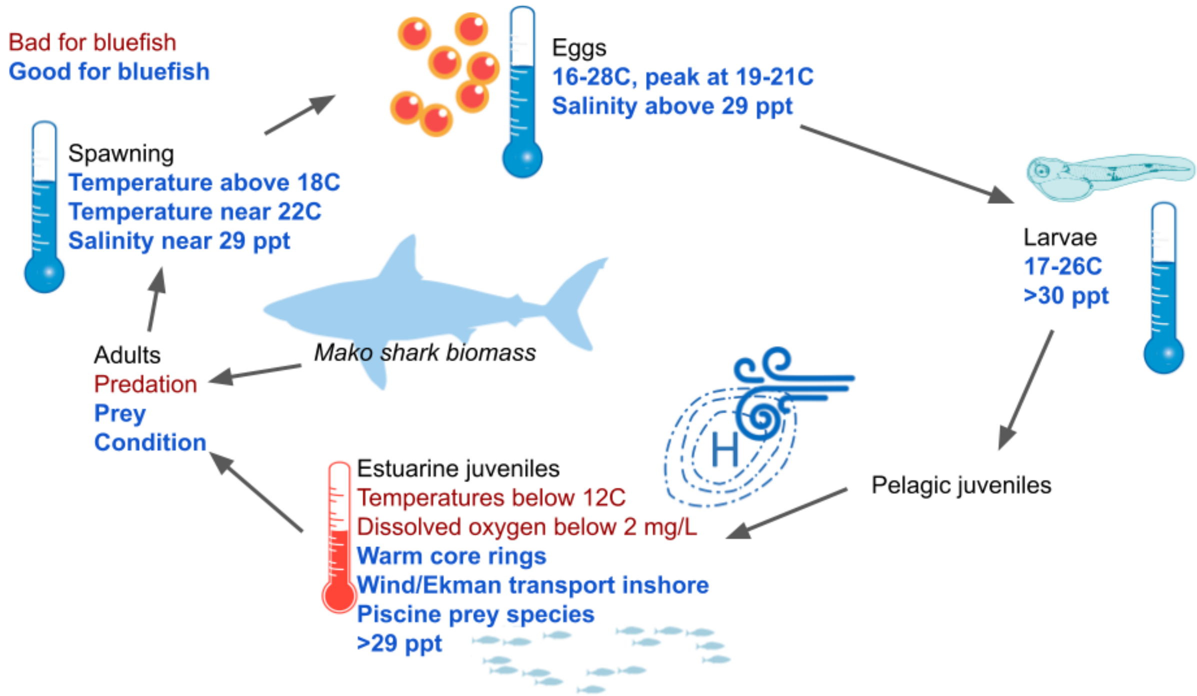


Figure 1. Life history conceptual model of bluefish identifying environmental factors with positive (blue text) or negative (red text) effects on different life stages of bluefish.

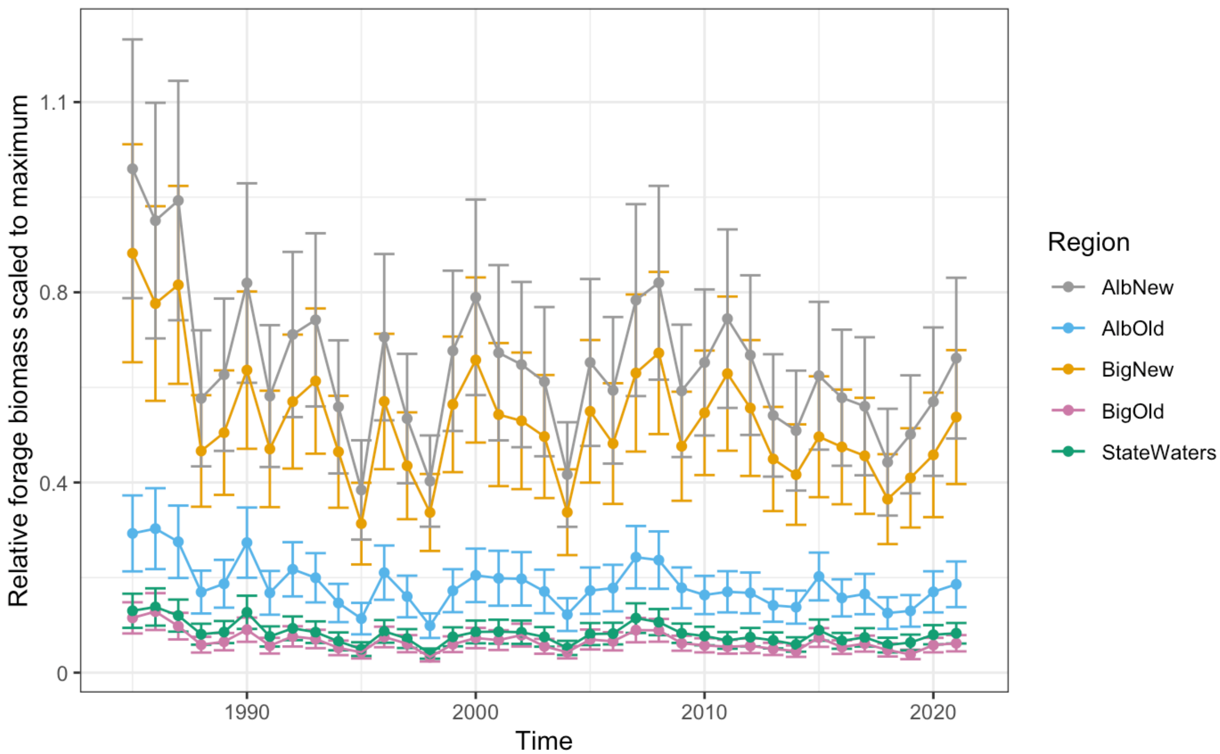


Figure 2. Forage fish indices for Fall 1985-2021 in Mid-Atlantic and Georges Bank nearshore and offshore areas. AlbNew= Albatross New, all inshore and new offshore survey strata (largest area); AlbOld= Albatross Old, includes all inshore survey strata; BigNew= Bigelow New, includes the subset of inshore survey strata that can be sampled by the R/V Henry Bigelow plus new offshore strata; BigOld = Bigelow Old, includes the subset of inshore survey strata that can be sampled by the R/V Henry Bigelow; StateWaters includes the coastline to 3 nautical miles offshore (smallest area)

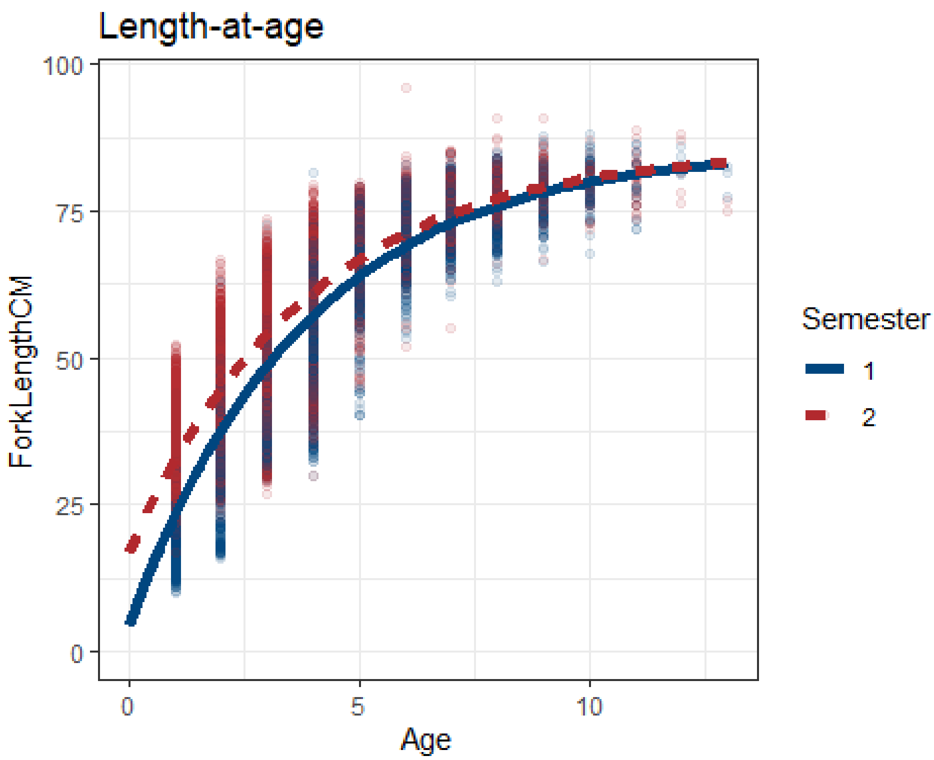


Figure 3. Fitted von Bertalanffy relationship by season using age-length data from state and federal fishery-dependent and fishery-independent sources. Jan-Jun=Semester 1, July-Dec=Semester 2.

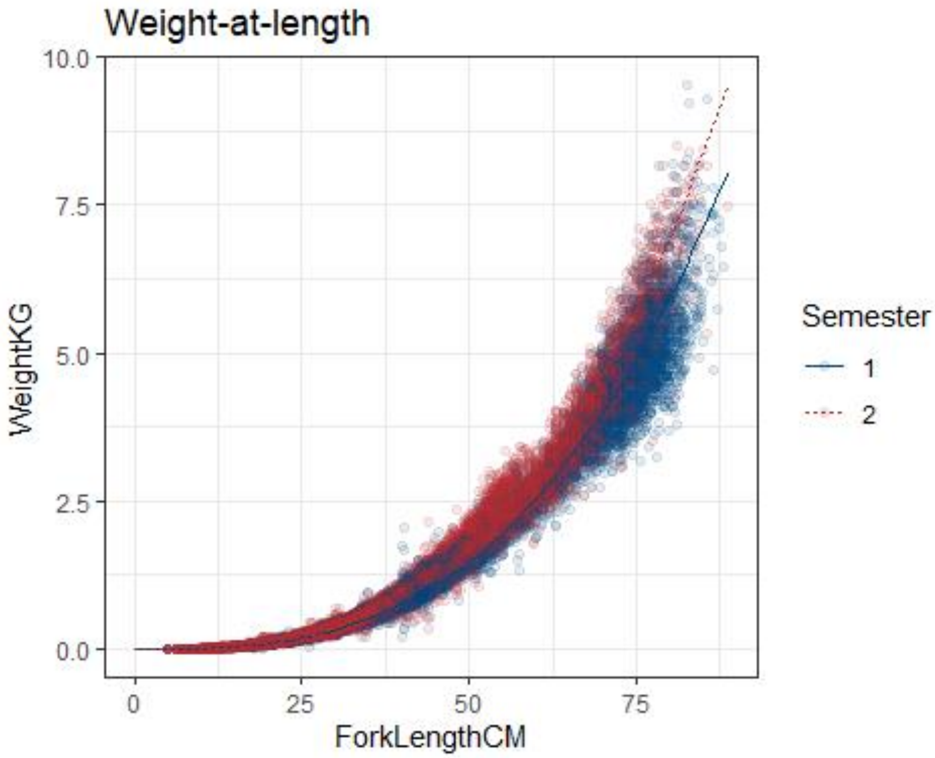


Figure 4. Fitted length-weight relationship by season using age-length data from state and federal fishery-dependent and fishery-independent sources. Jan-Jun=Semester 1, July-Dec=Semester 2.

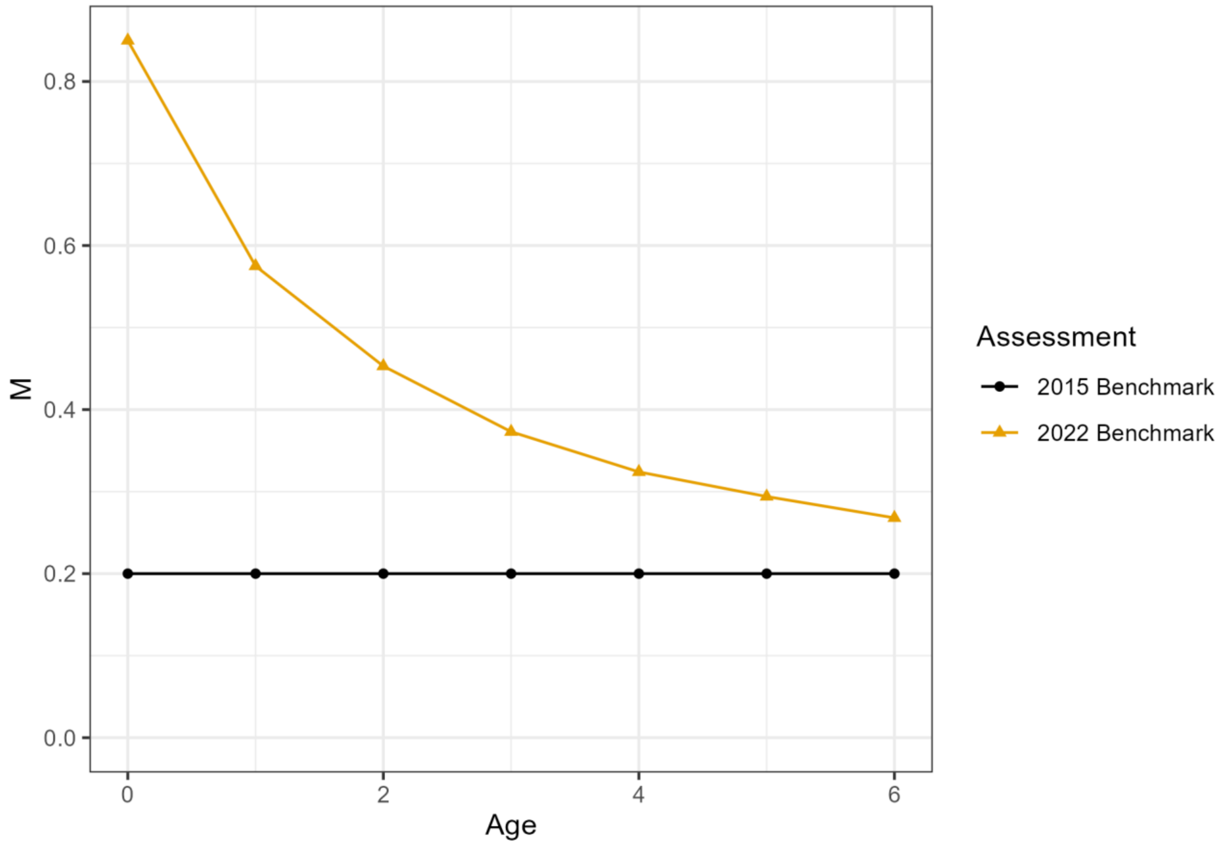


Figure 5. Comparison of the estimates of M-at-age used in the 2015 benchmark assessment and this assessment.

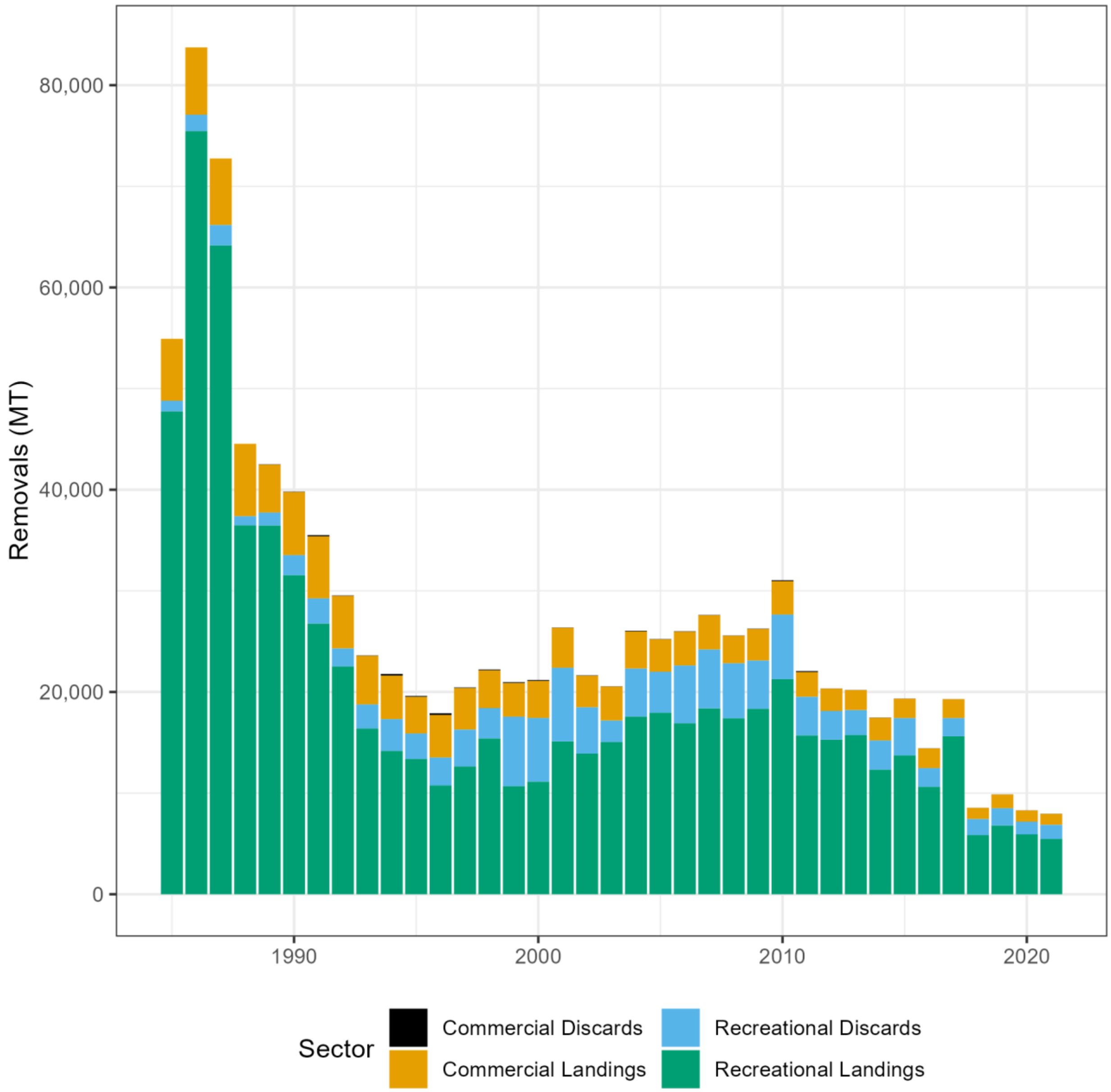


Figure 6. Total removals of bluefish on the Atlantic coast by sector, 1985-2021.

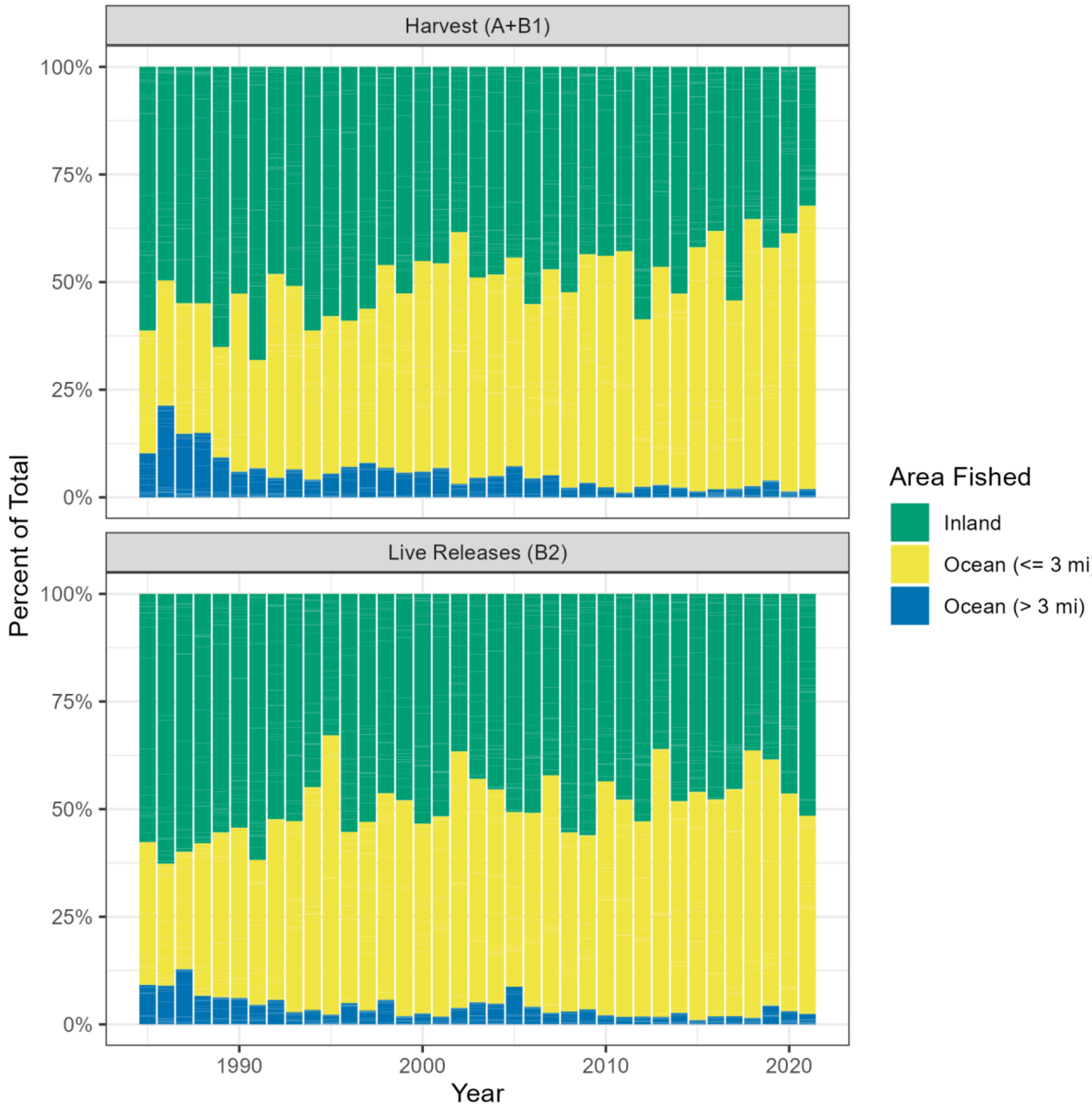


Figure 7. Percent of recreational harvest (top) and live releases (bottom) in numbers by area fished over time.

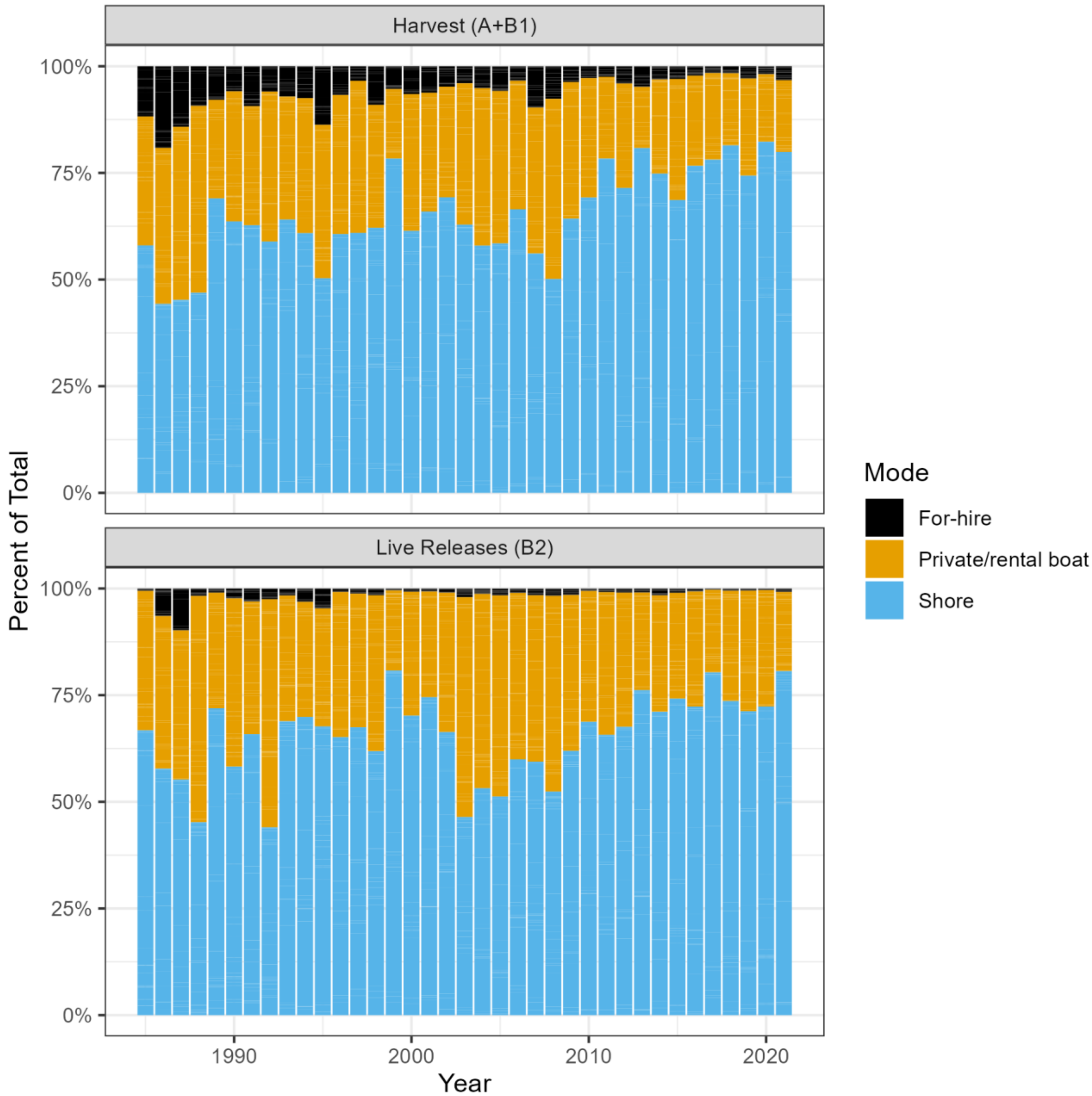


Figure 8. Percent of recreational harvest (top) and live releases (bottom) in numbers by mode of fishing over time.

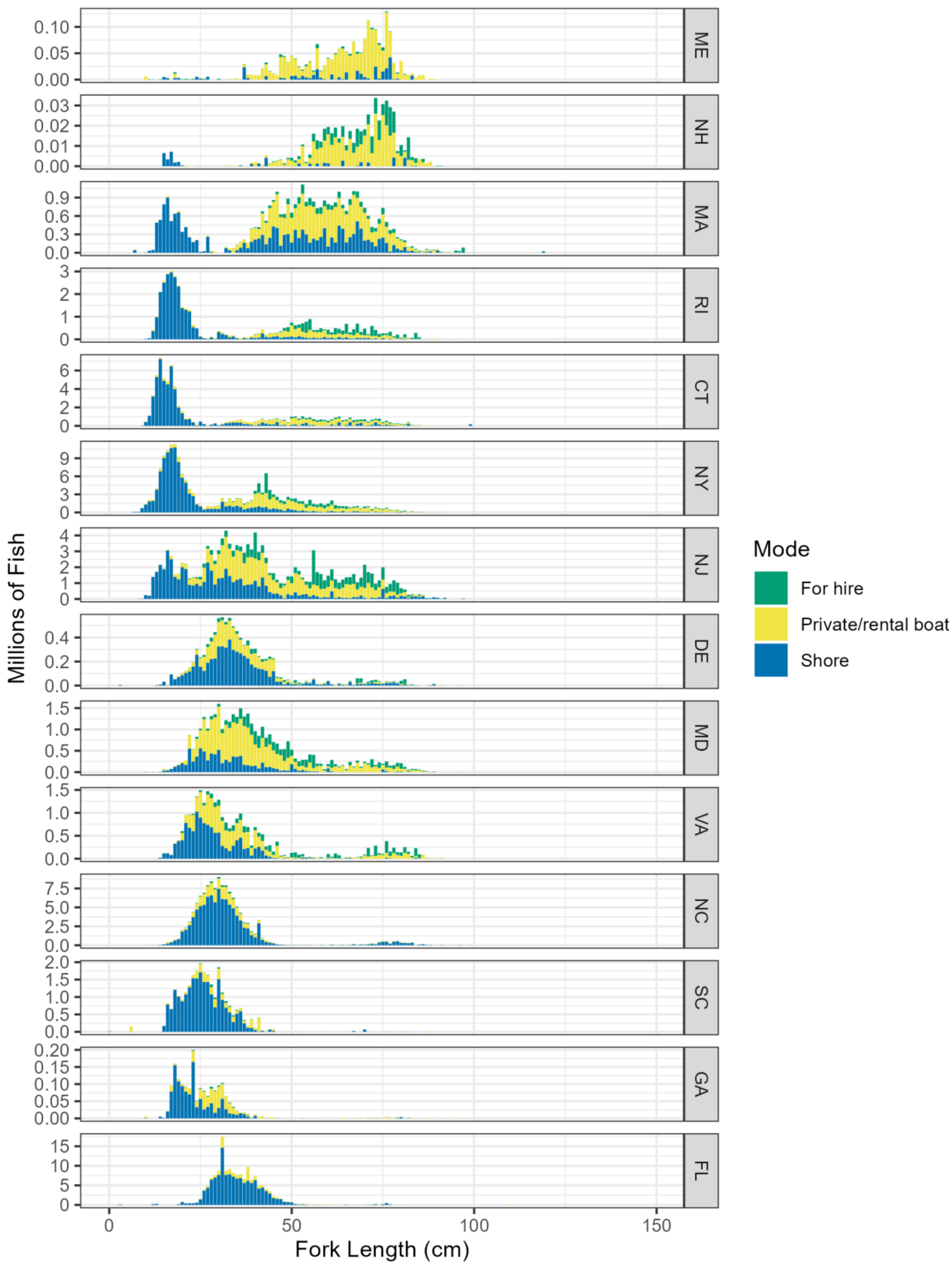


Figure 9. Length frequency of recreationally harvested bluefish by state and mode of fishing, 1982-2020 pooled.

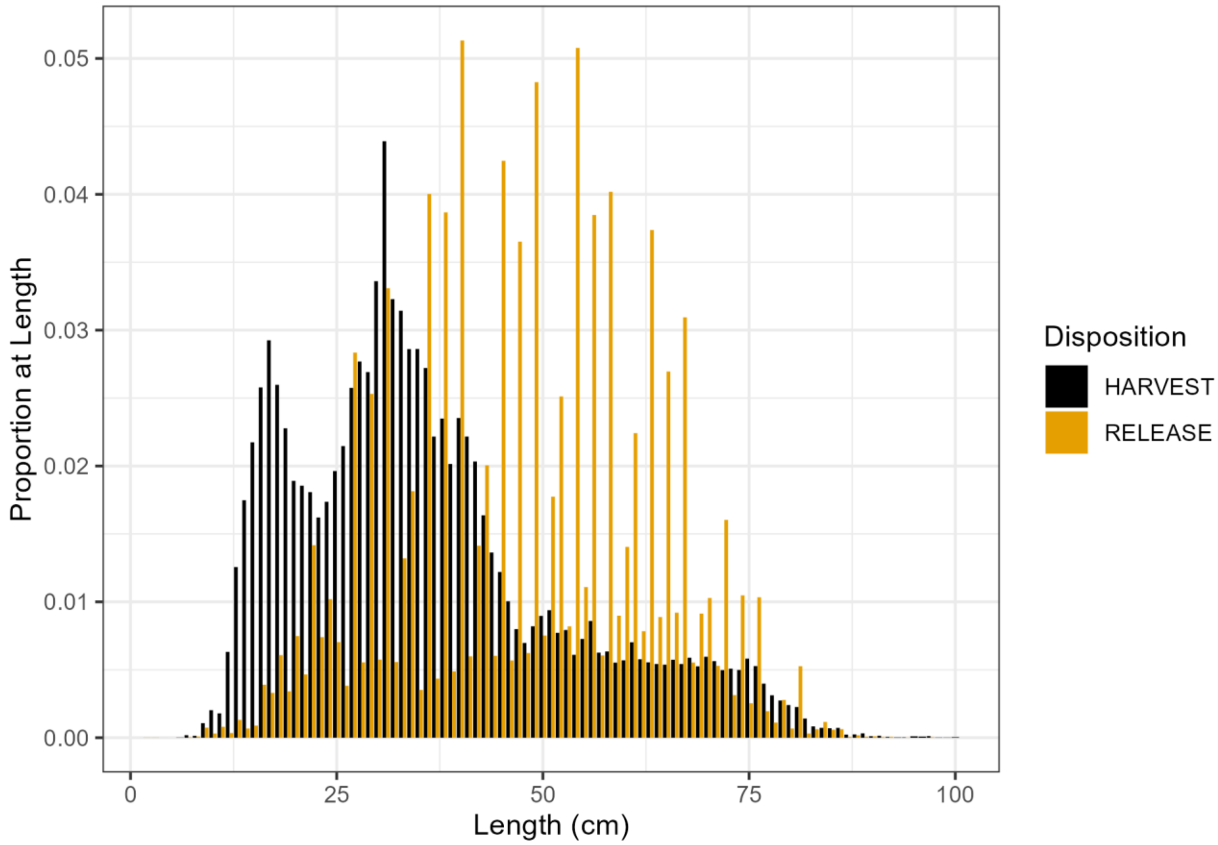


Figure 10. Proportion at length for harvested and released bluefish sampled, pooled over region, season, and years.

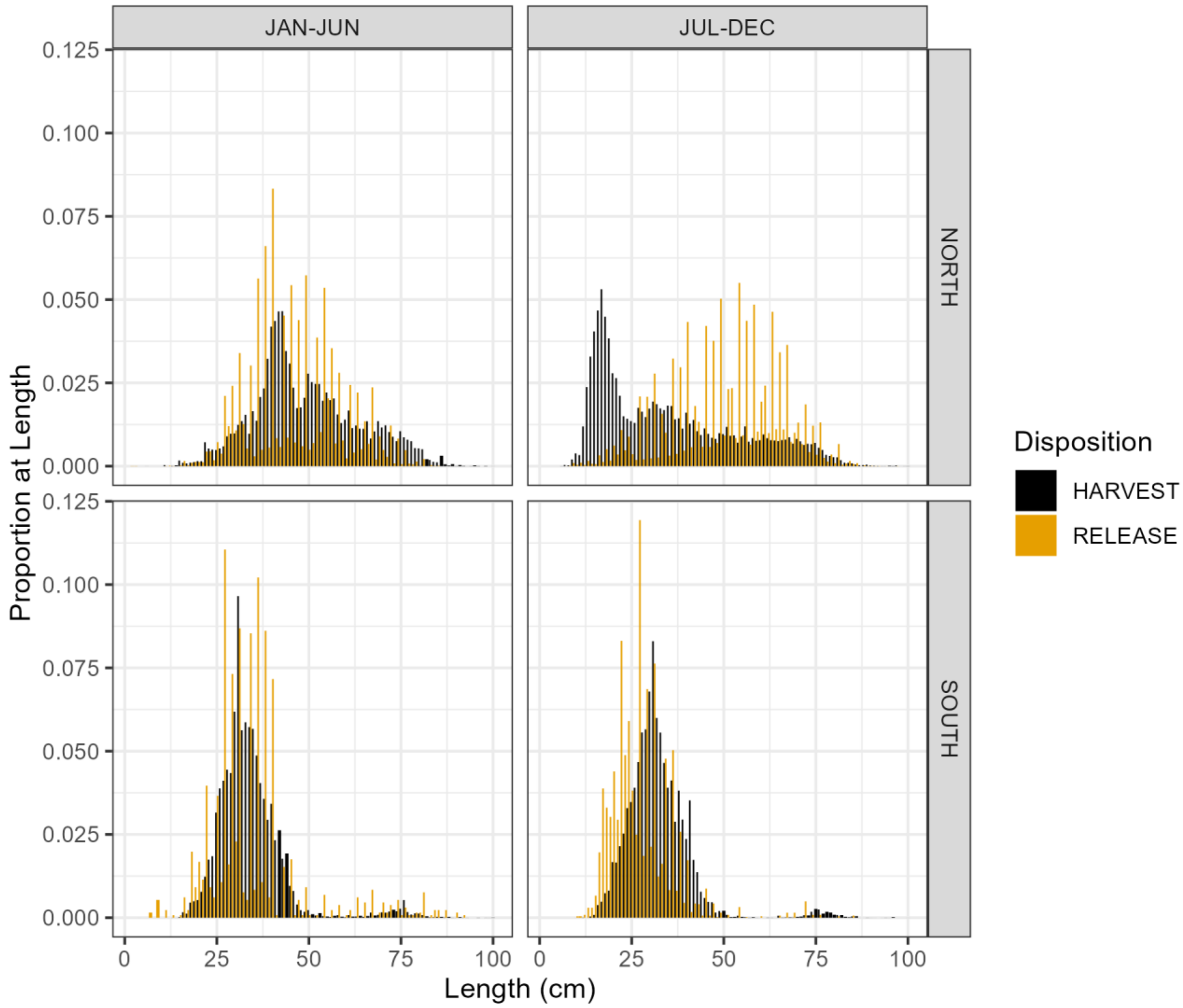


Figure 11. Proportions of length of harvested and released bluefish by region and season. Data pooled over 1985-2021.

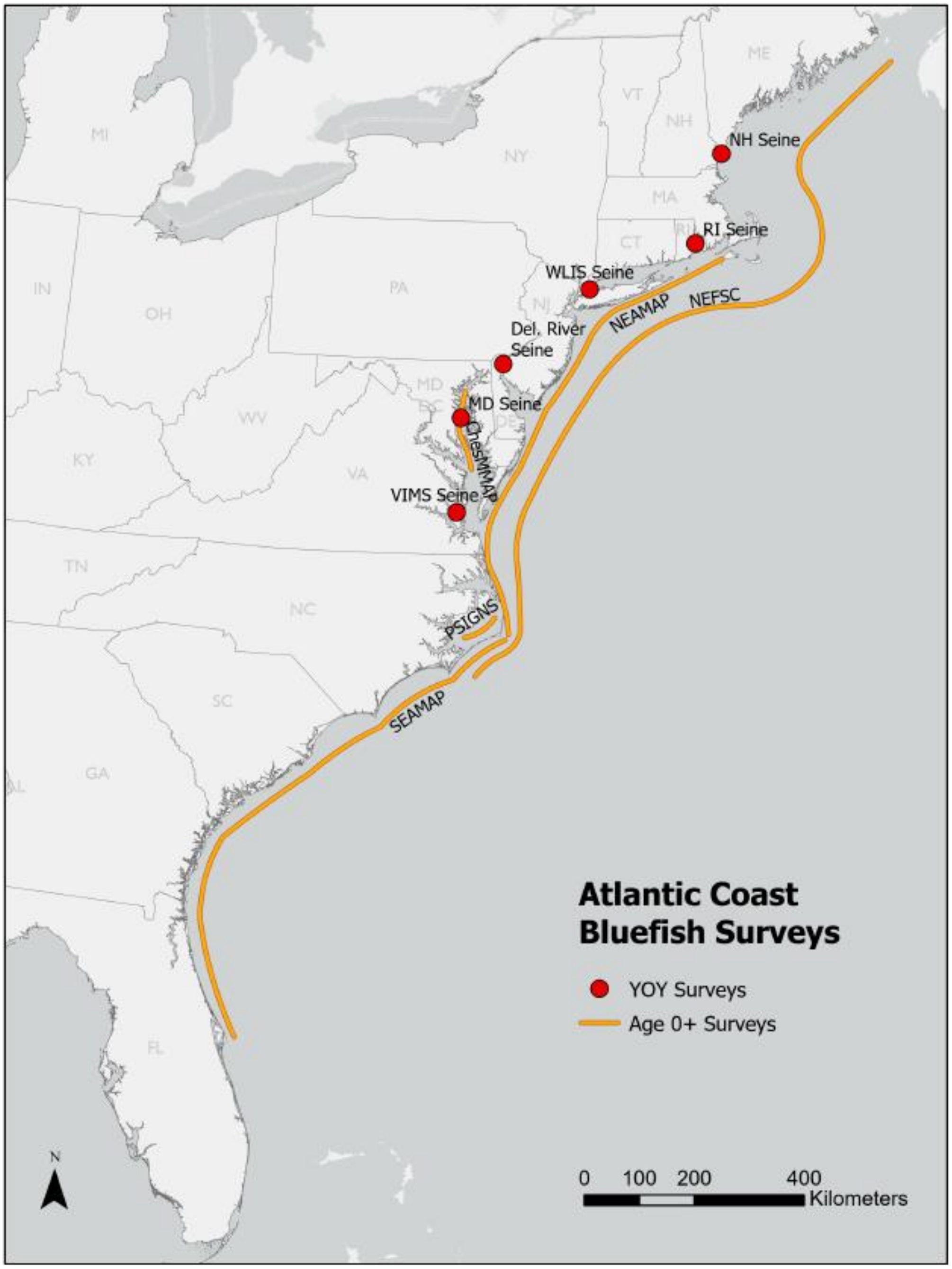


Figure 12. Map of the east coast of the United States showing the approximate locations of each of the surveys used in the final assessment model.

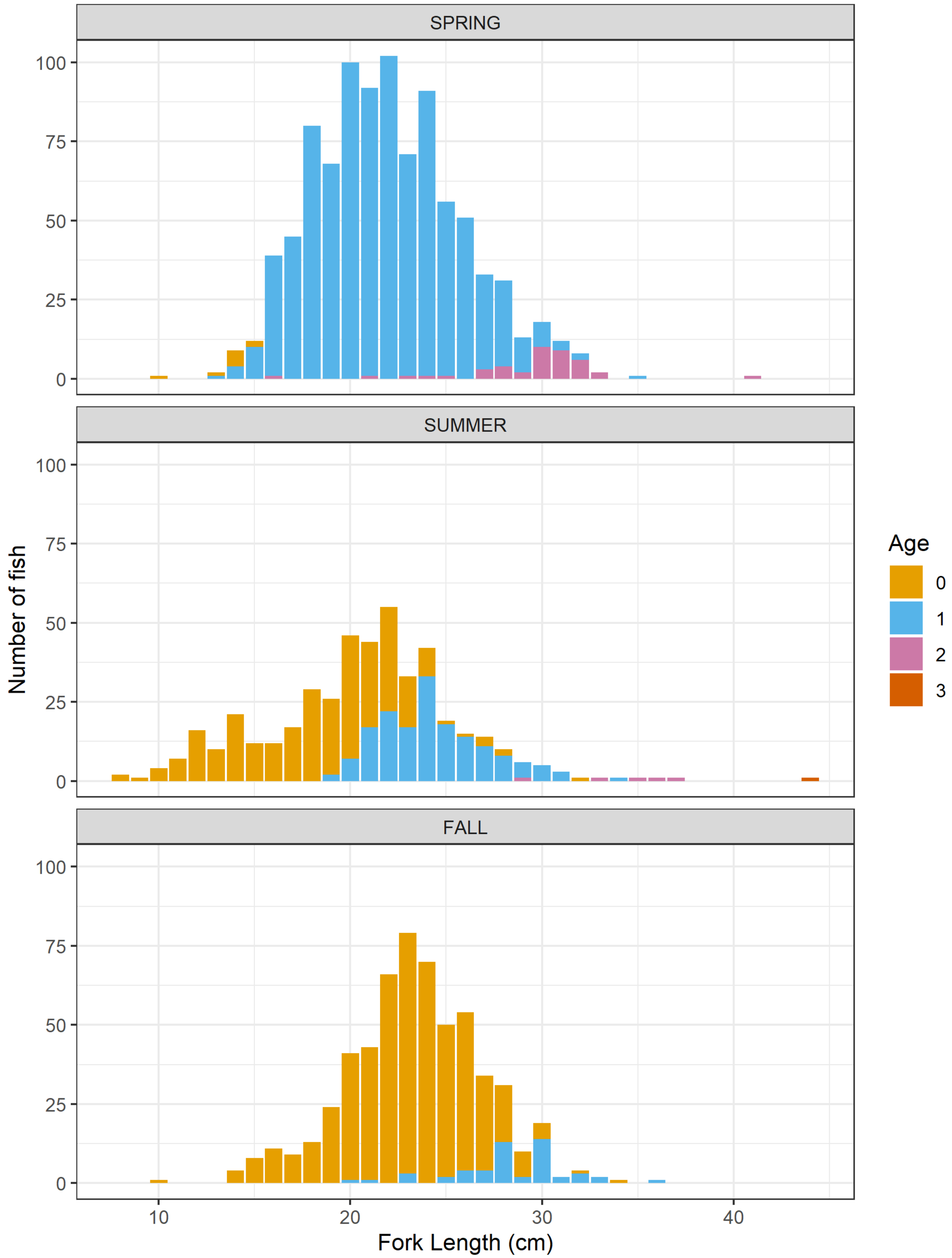


Figure 13. Length and age frequency of the SEAMAP survey by season, pooled over all years.

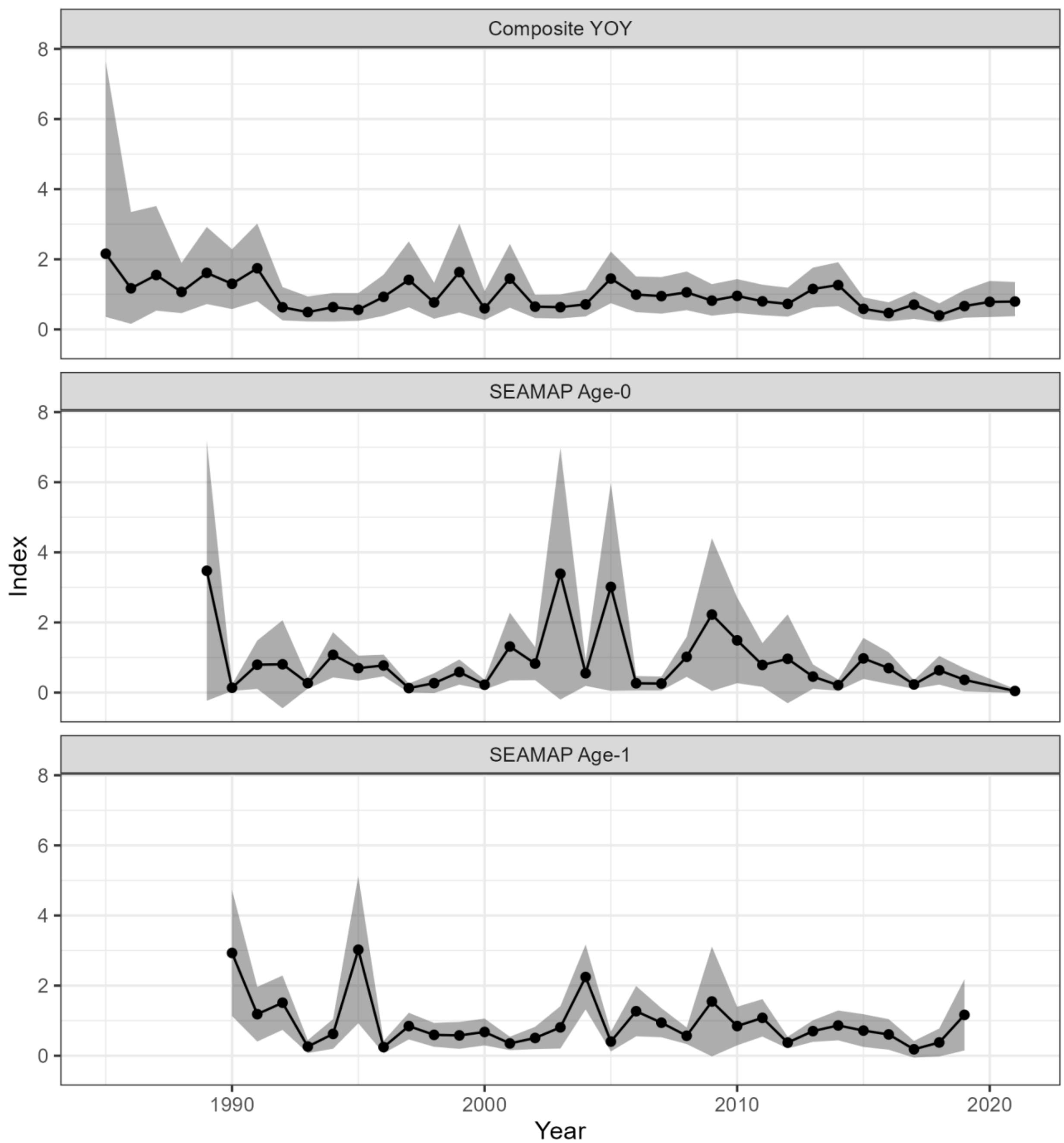


Figure 14. Indices of bluefish recruitment (i.e, age-0 and age-1 only) used in the ASAP and WHAM models.

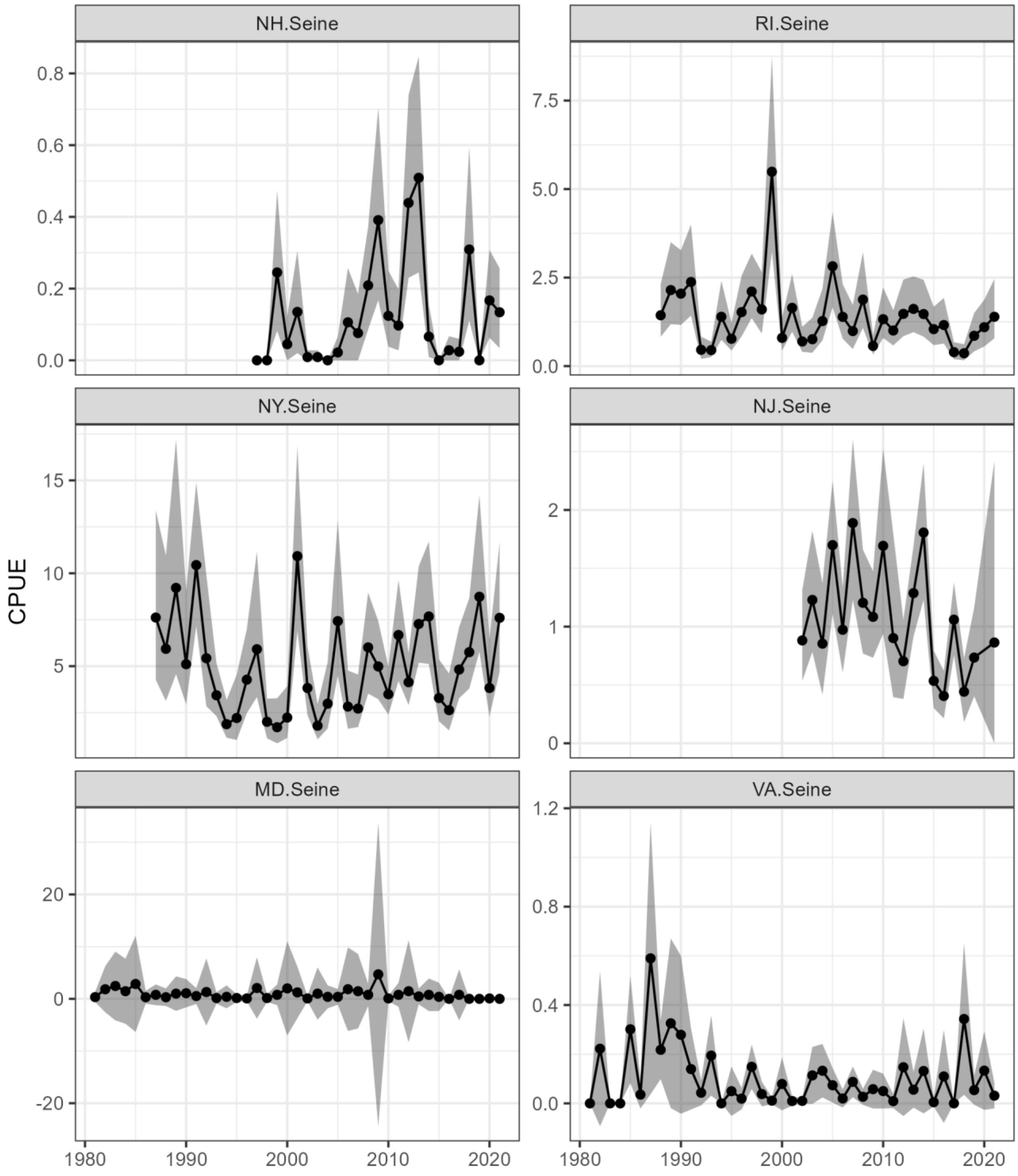


Figure 15. Indices of young-of-year abundance from state seine surveys used in the composite YOY index. Shaded area indicates 95% confidence interval.

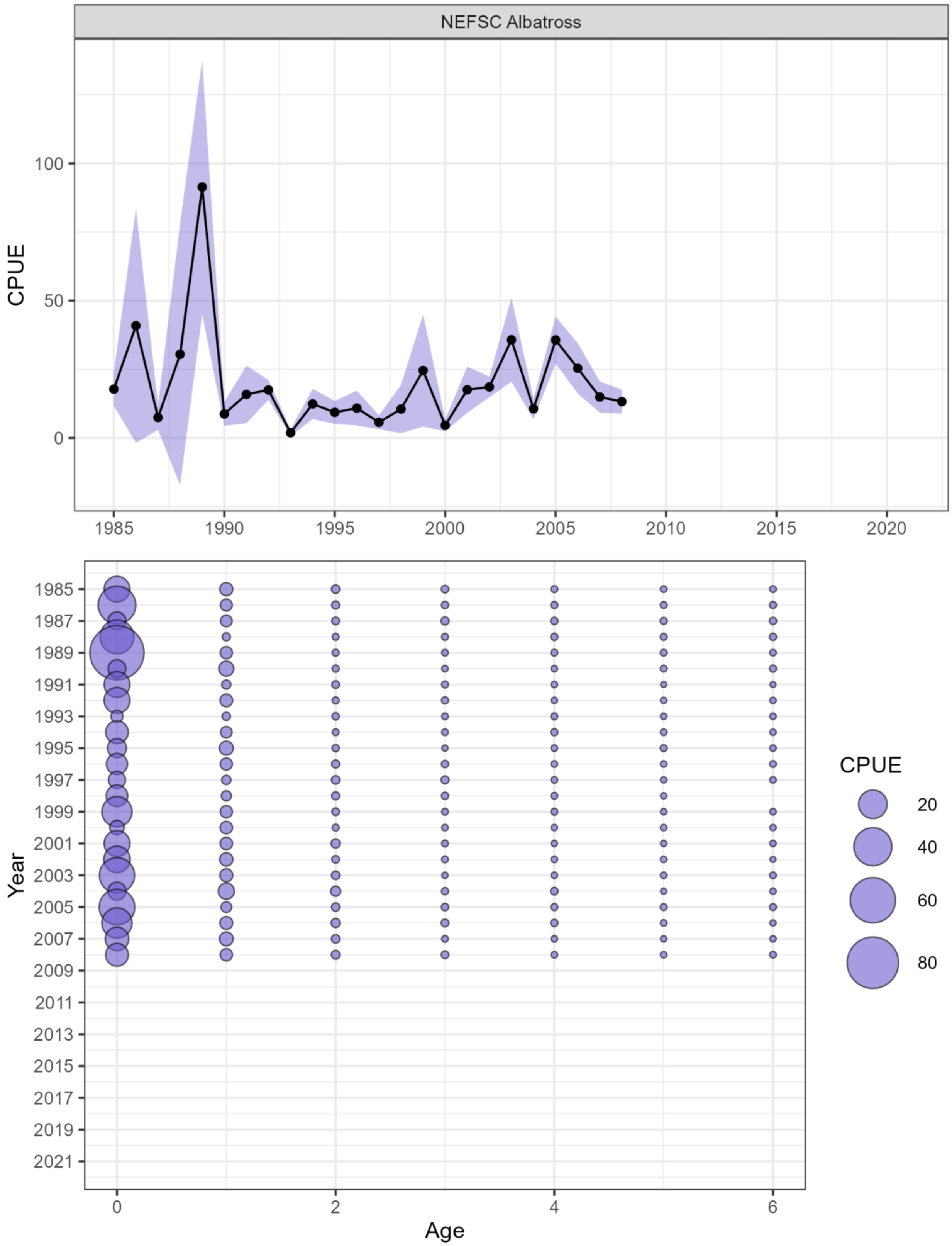


Figure 16. Index of abundance (top) and age composition (bottom) from the NEFSC fall trawl survey (Albatross years). Shaded area indicates 95% confidence interval.

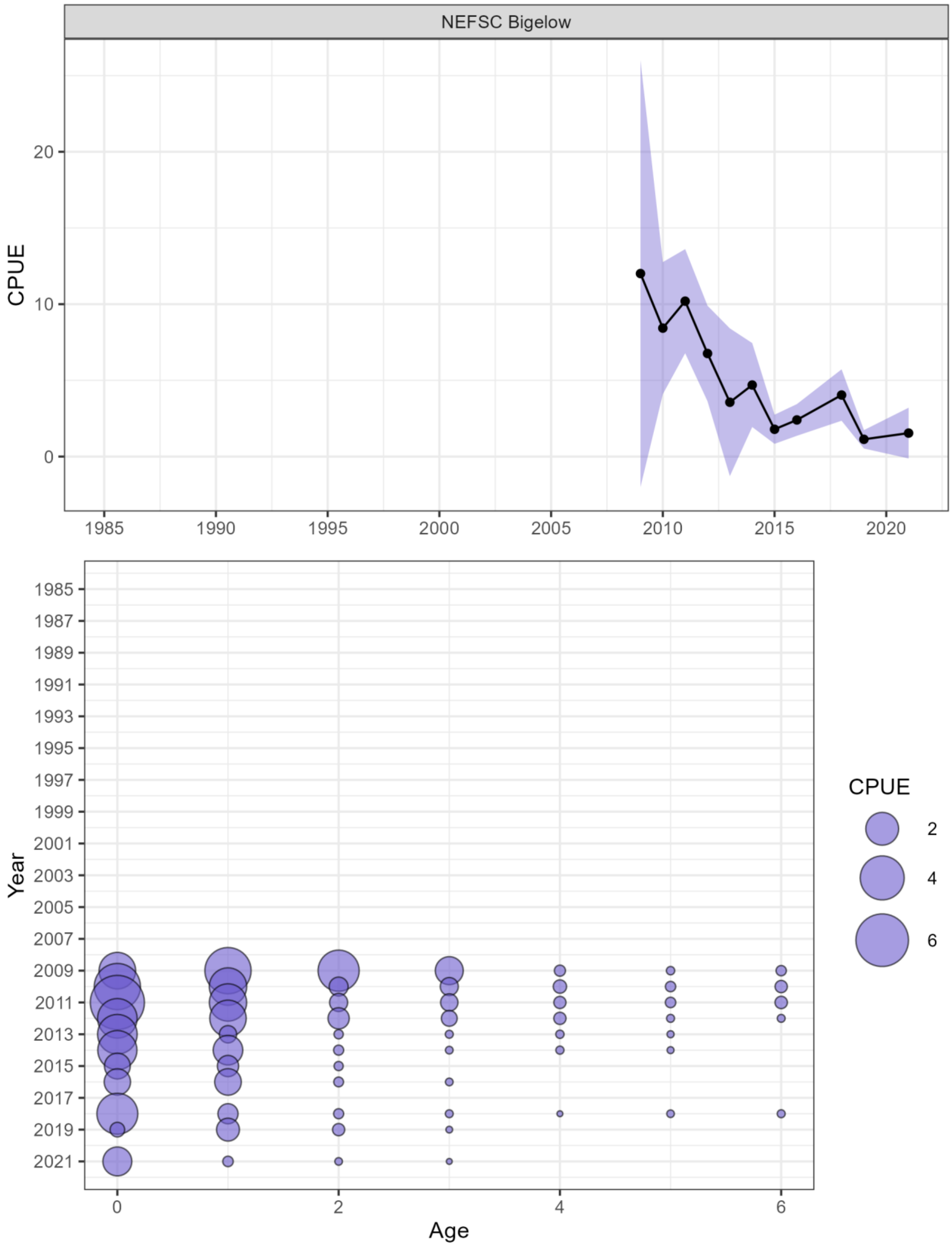


Figure 17. Index of abundance (top) and age composition (bottom) of the NEFSC fall trawl survey (Bigelow years). Shaded area indicates 95% confidence interval.

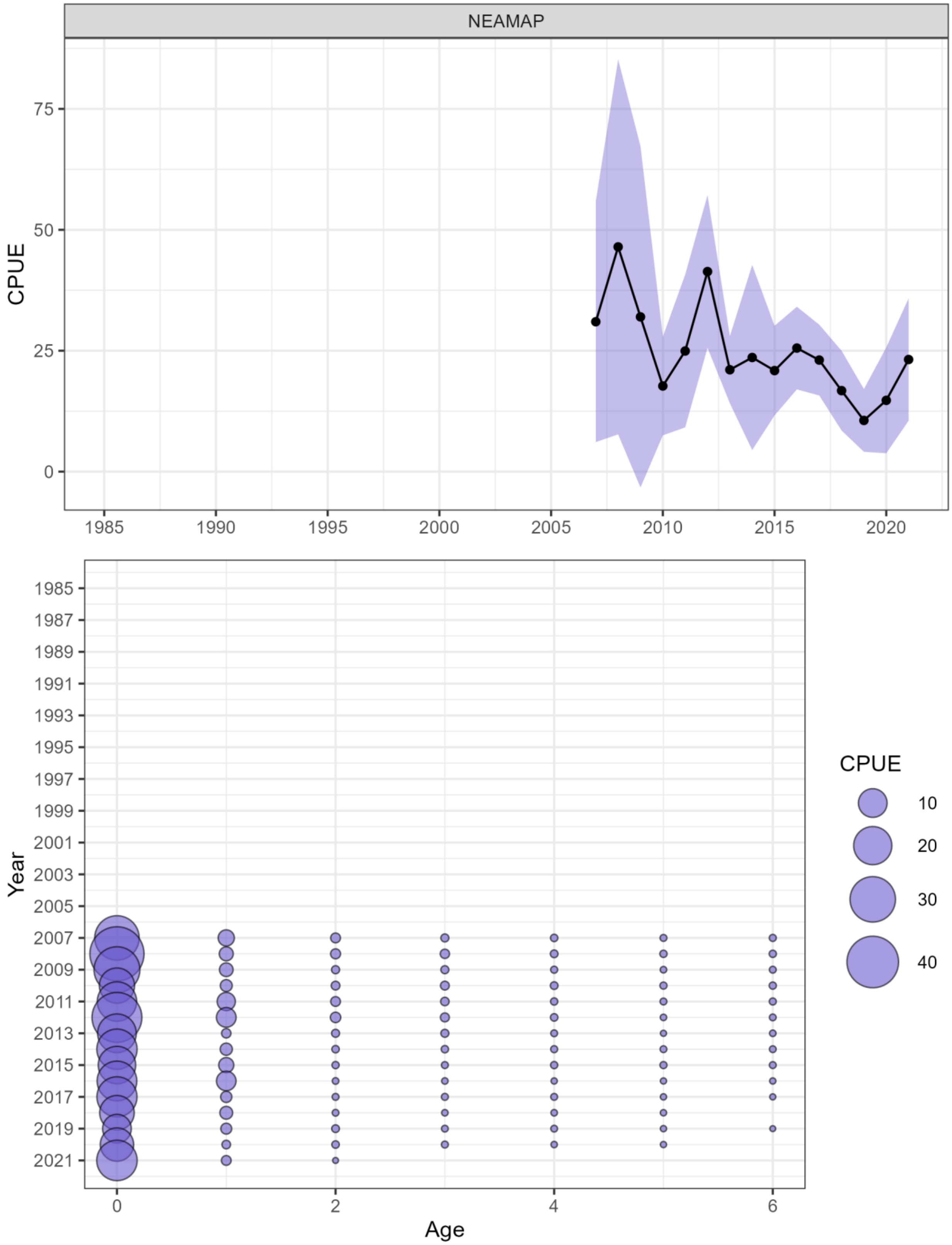
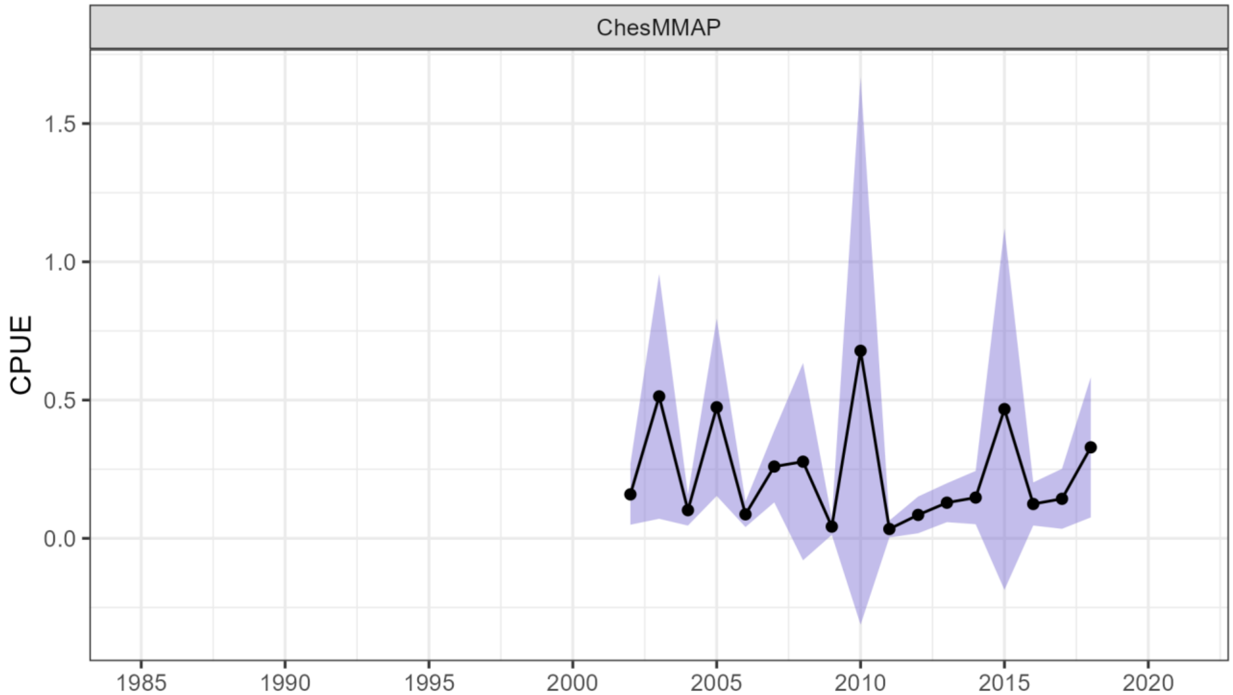


Figure 18. Index of abundance (top) and age composition (bottom) of the NEAMAP survey. Shaded area indicates 95% confidence interval.



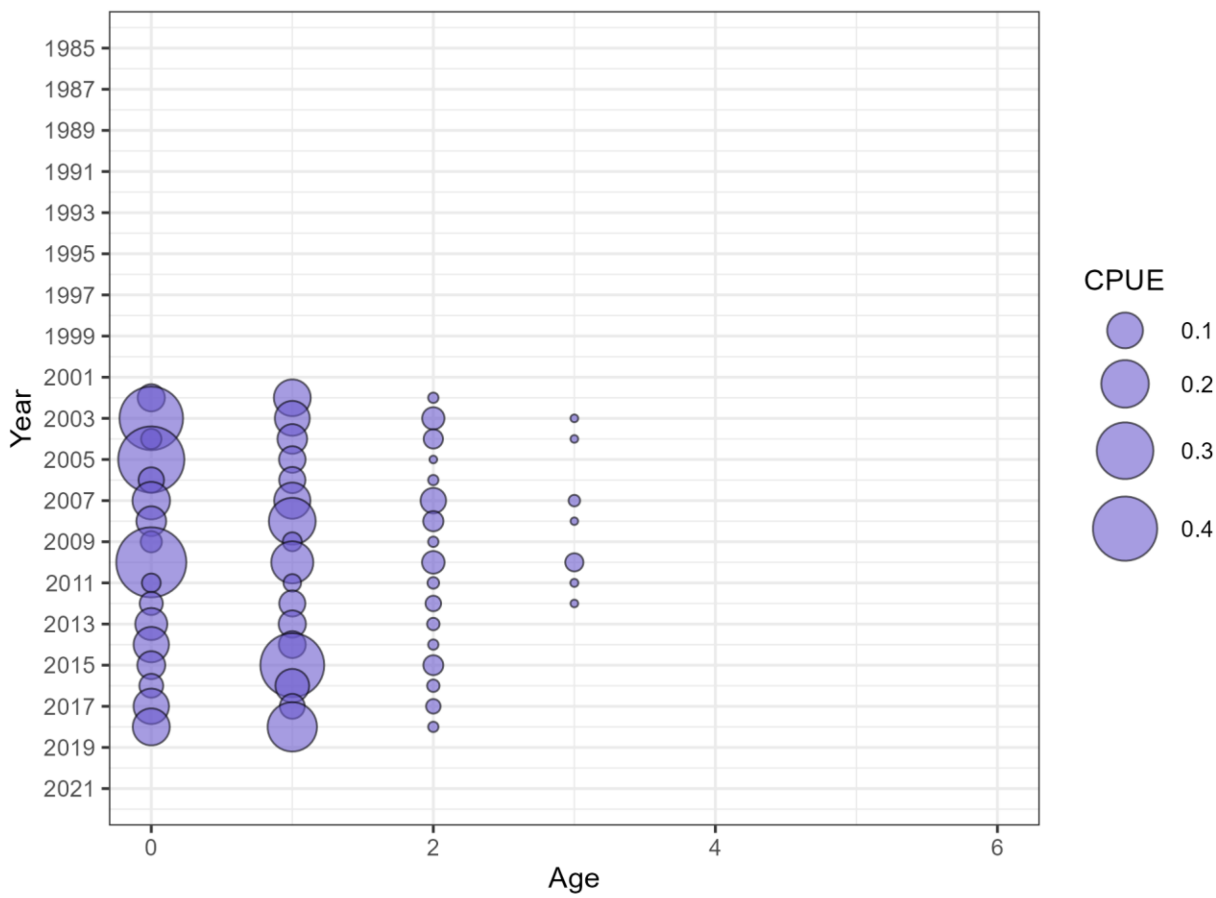


Figure 19. Index of abundance (top) and age composition (bottom) of the ChesMMAP survey. Shaded area indicates 95% confidence interval.

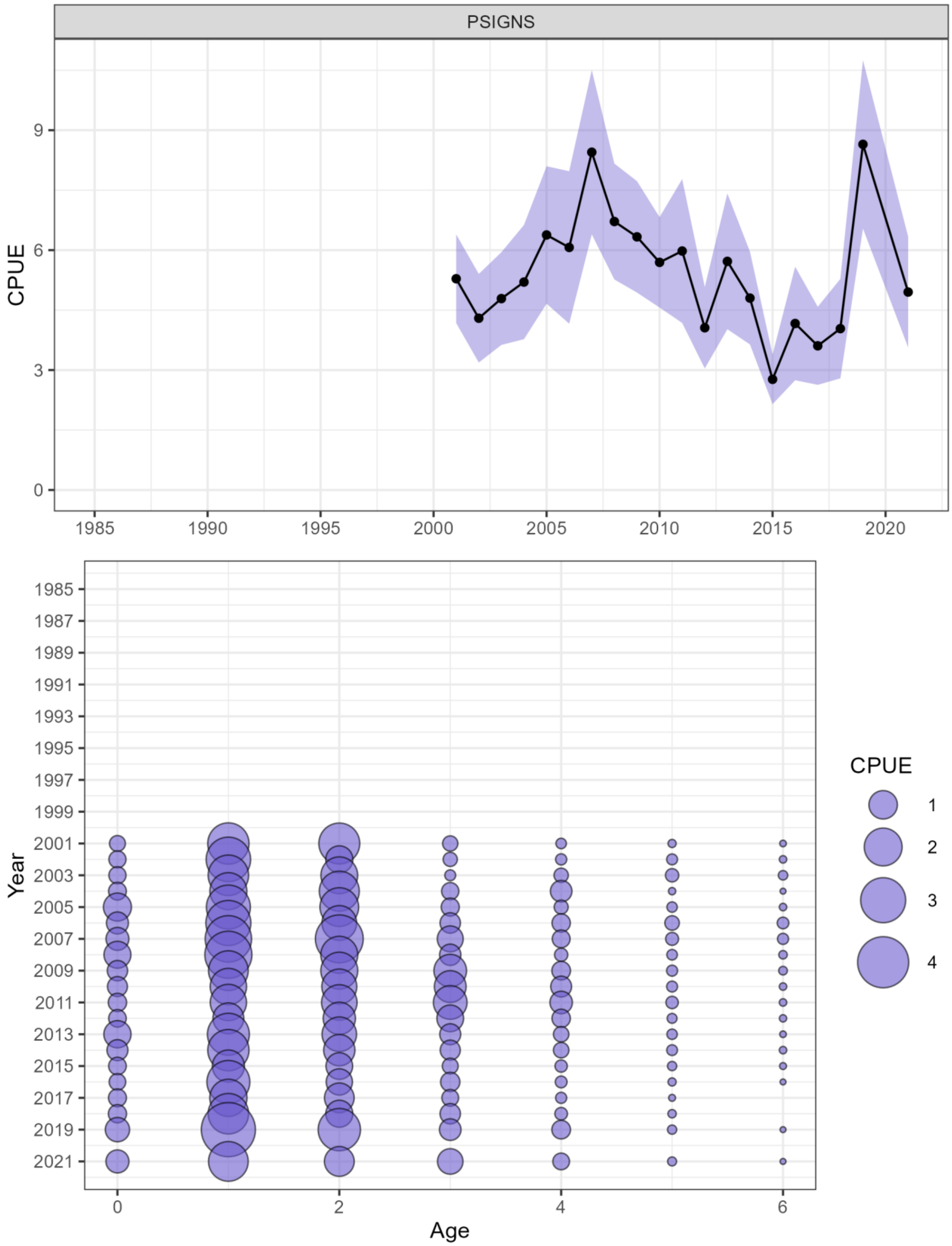


Figure 20. Index of abundance (top) and age composition (bottom) of the NC PSIGN survey. Shaded area indicates 95% confidence interval.

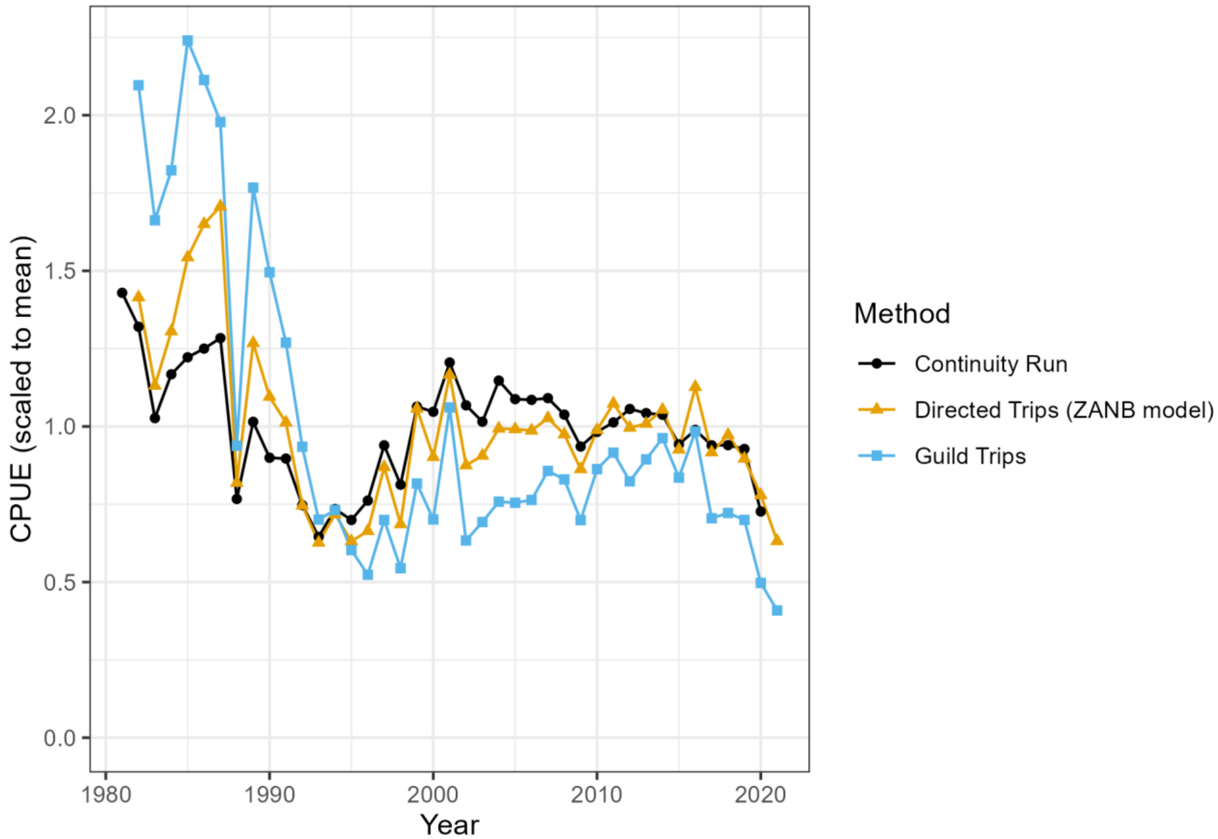


Figure 21. Comparison of trends in the MRIP CPUE index developed using different trip selection and standardization methods.

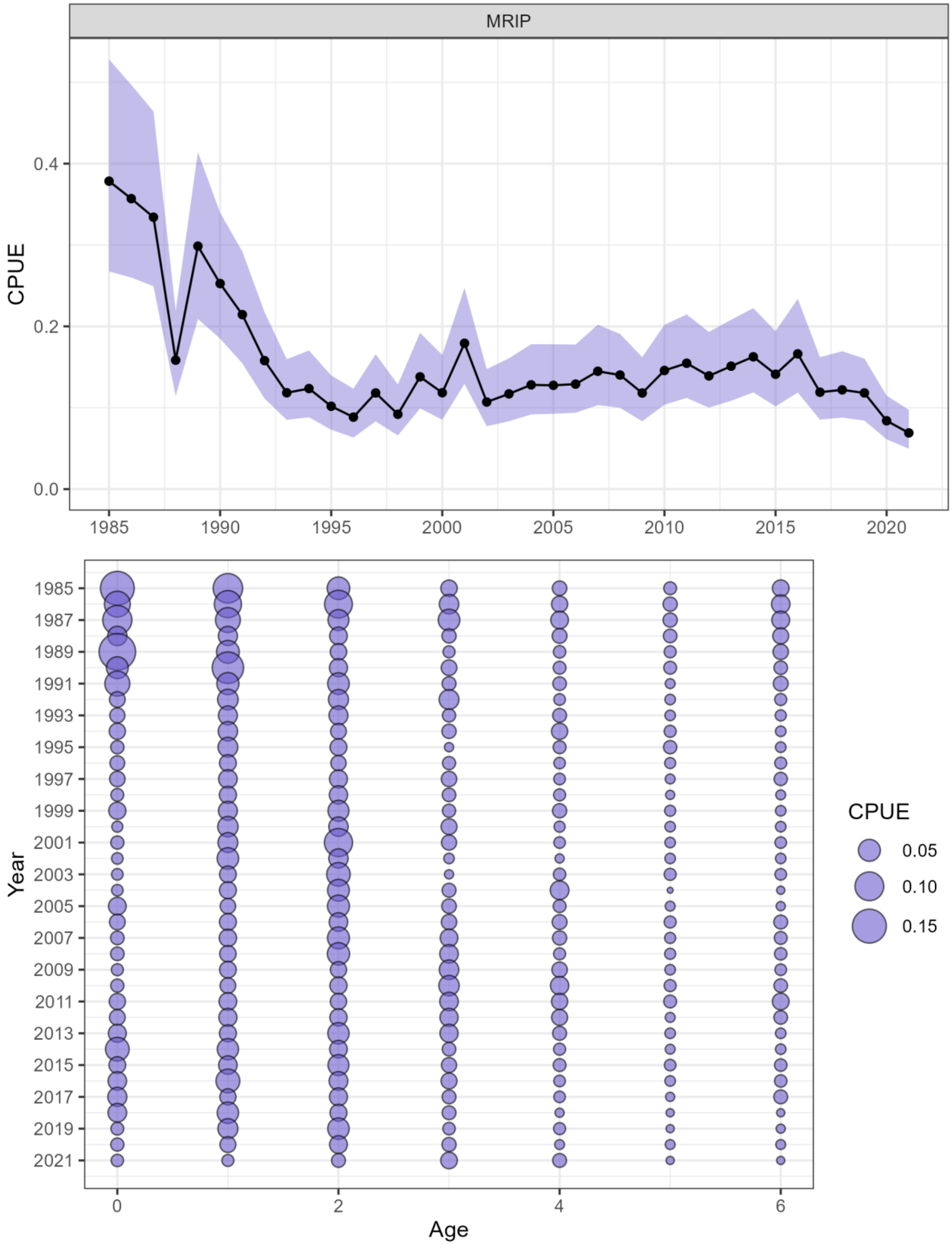


Figure 22. MRIP CPUE (top) and age composition (bottom).Shaded area indicates 95% confidence interval.

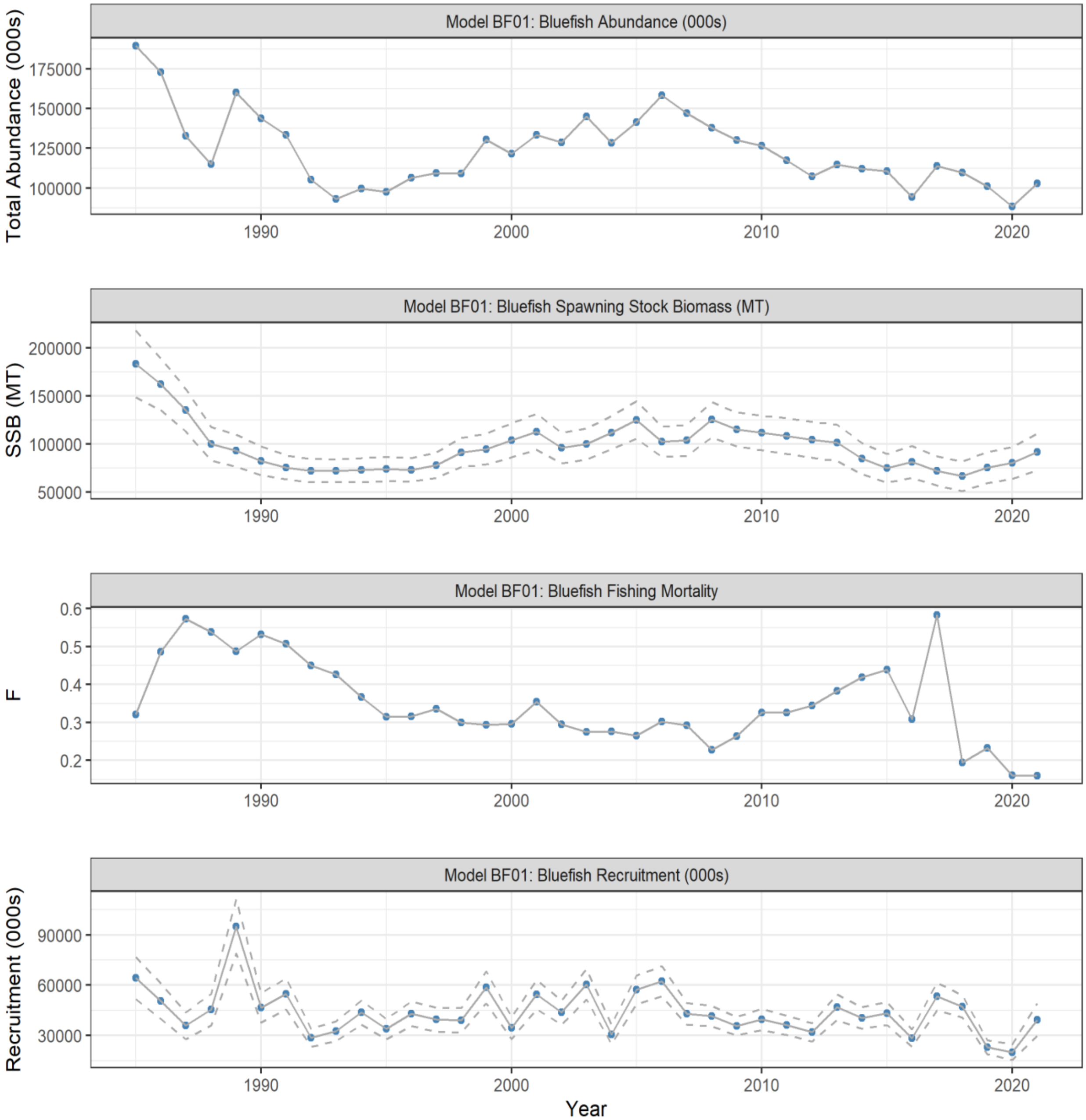


Figure 23. Model results of abundance, SSB, fishing mortality, and recruitment for the continuity run (Model BF01).

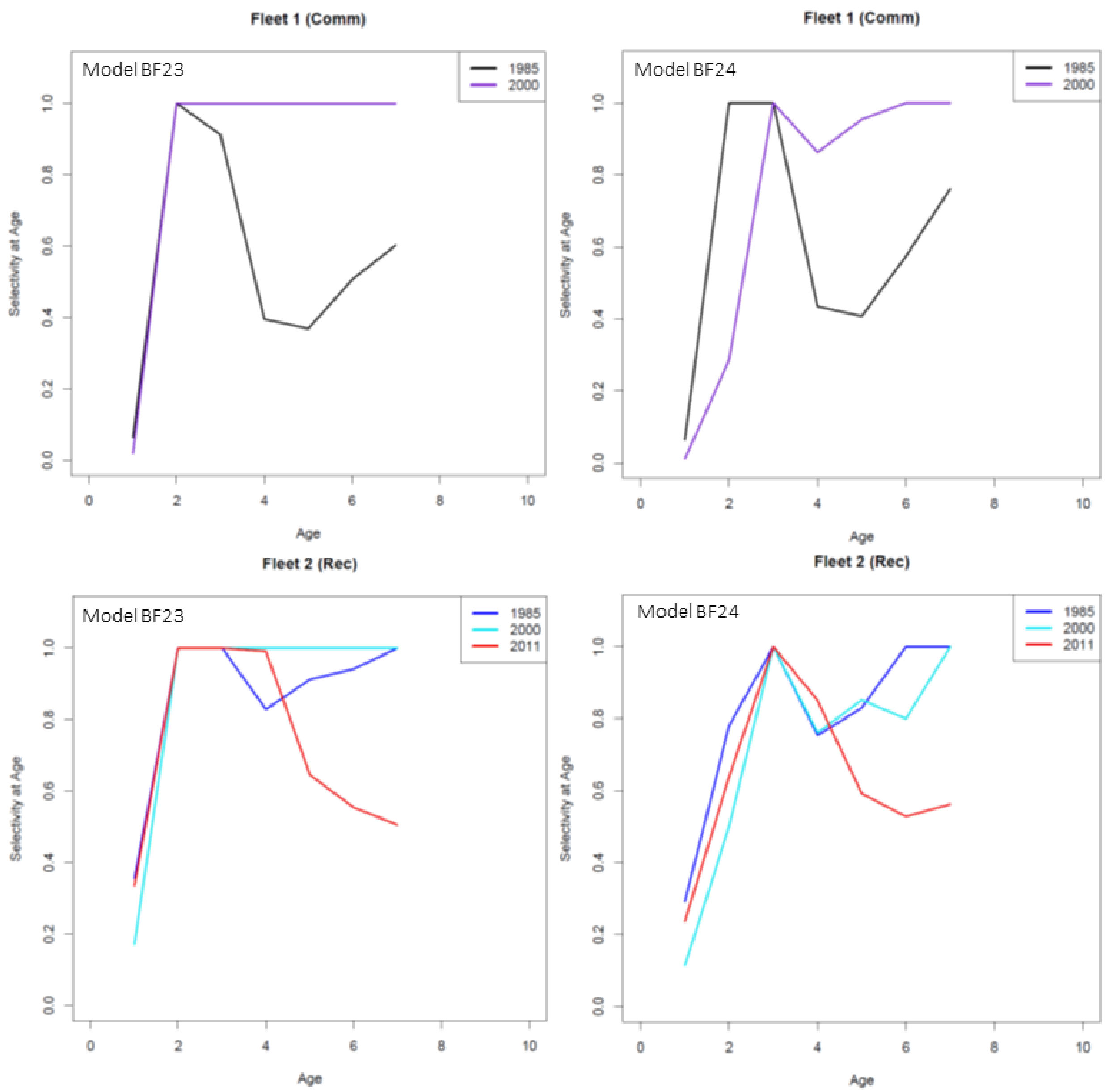


Figure 24. Fleet selectivity block comparison between model BF23 (left) and final model BF24 (right) after addressing poor age composition residual blocks in BF23.

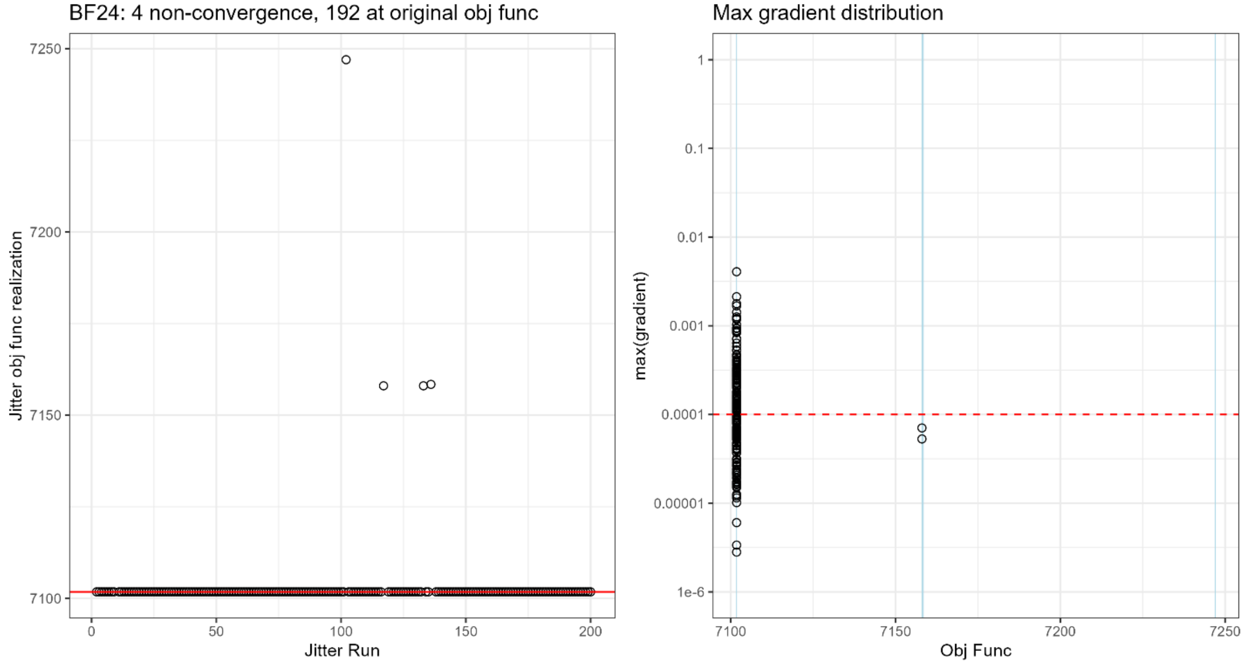


Figure 25. Convergence diagnostics for the final ASAP model, model BF24. The left plot shows objective function results for 200 random sets of starting values, with the original objective function designated by the horizontal red line. The right plot shows gradient values from each model distributed around a 0.0001 criterion for 'good' convergence (blue vertical lines represent other objective function solutions, some where the gradient result is above or below the y-axis range of the plot).

Figure 26. Comparison between model results from ASAP model BF24 (blue lines and points) and the same model run as a traditional statistical CAA model in WHAM (orange lines).

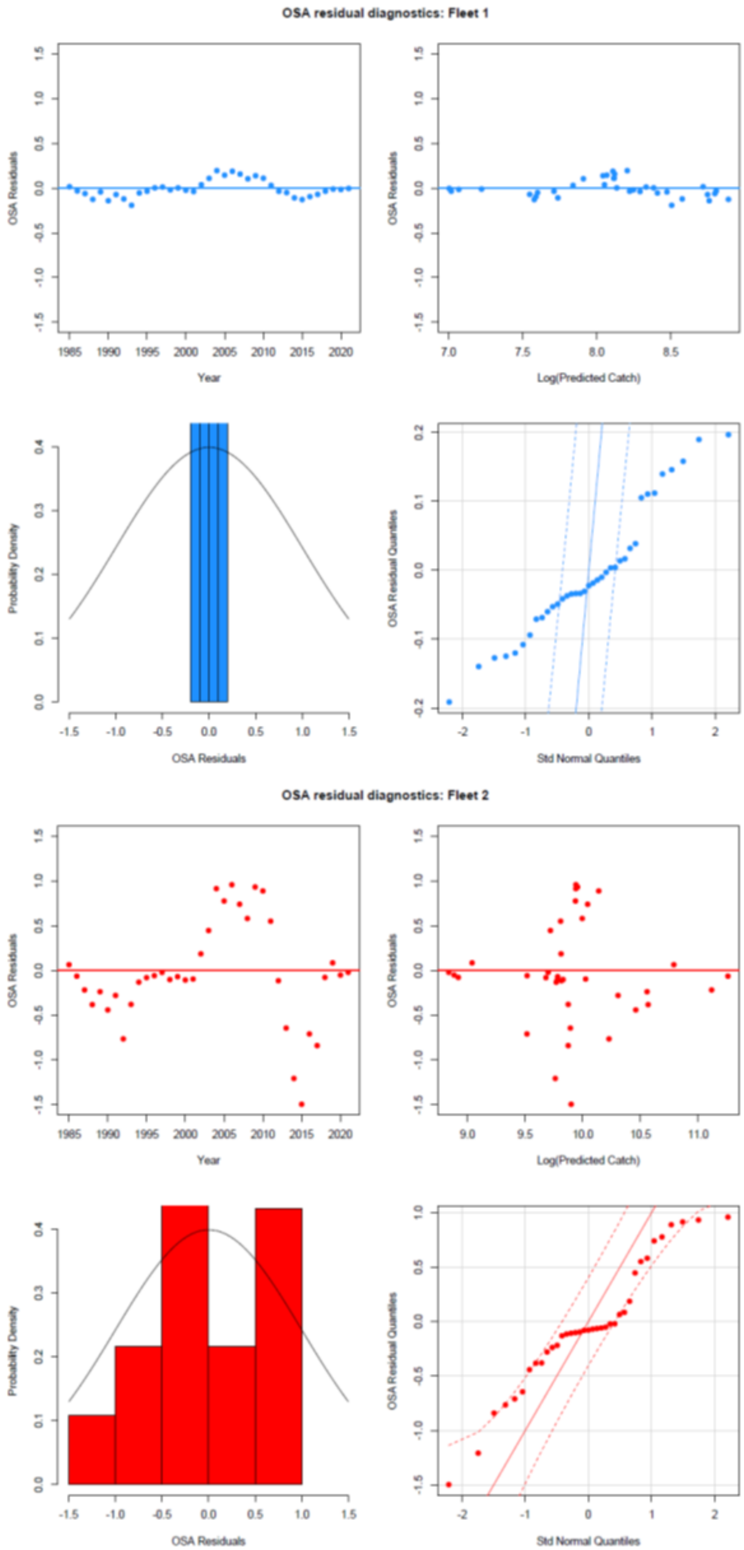


Figure 27. One-step ahead residual diagnostics for the 2 fleets. Patterns in the diagnostics for both fleets led to a reduction in the fleet input CVs.

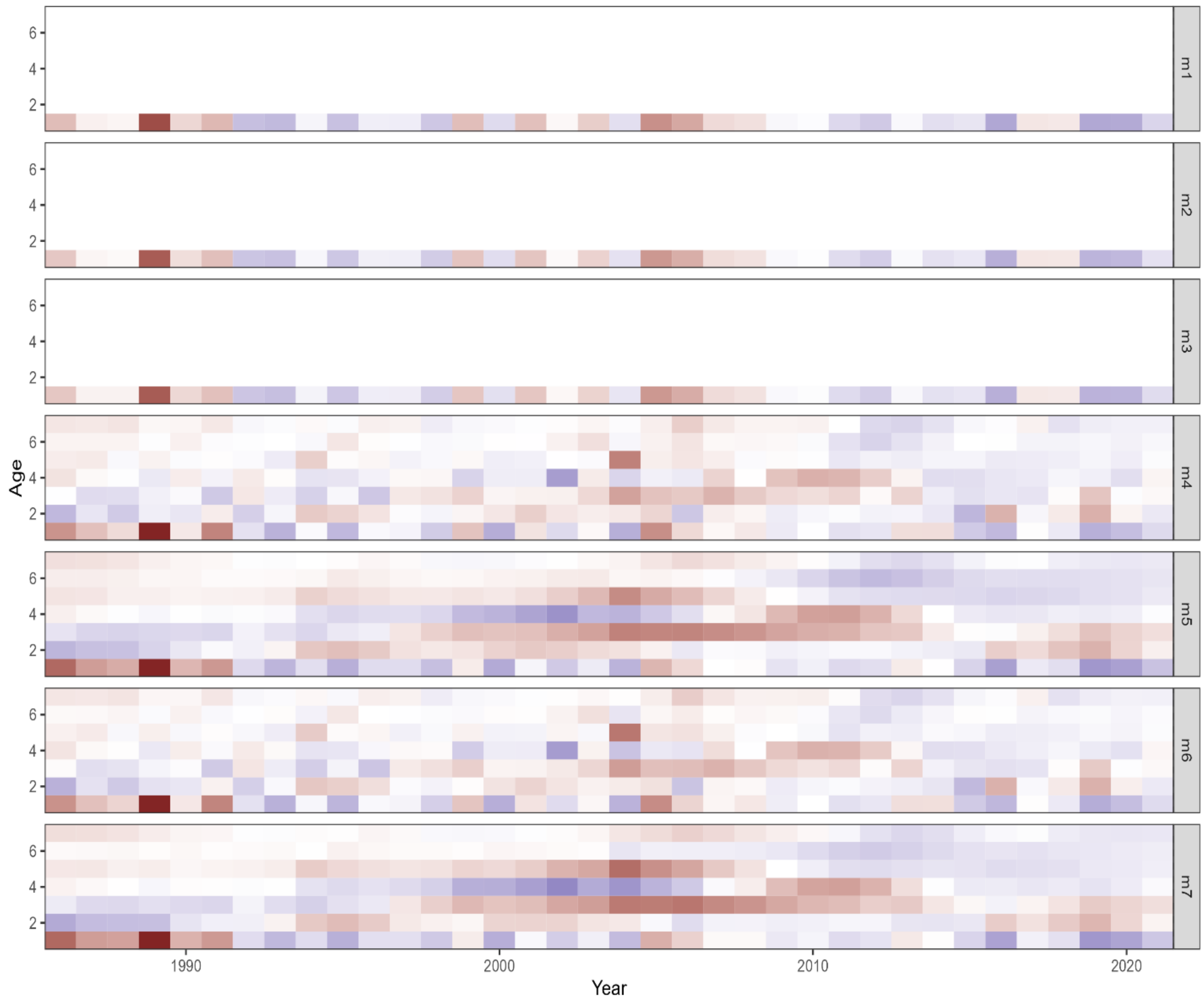


Figure 28. Number-at-age deviations for the models BF28W_m1 through BF28W_m7. Red indicates positive deviations and blue indicates negative deviations.

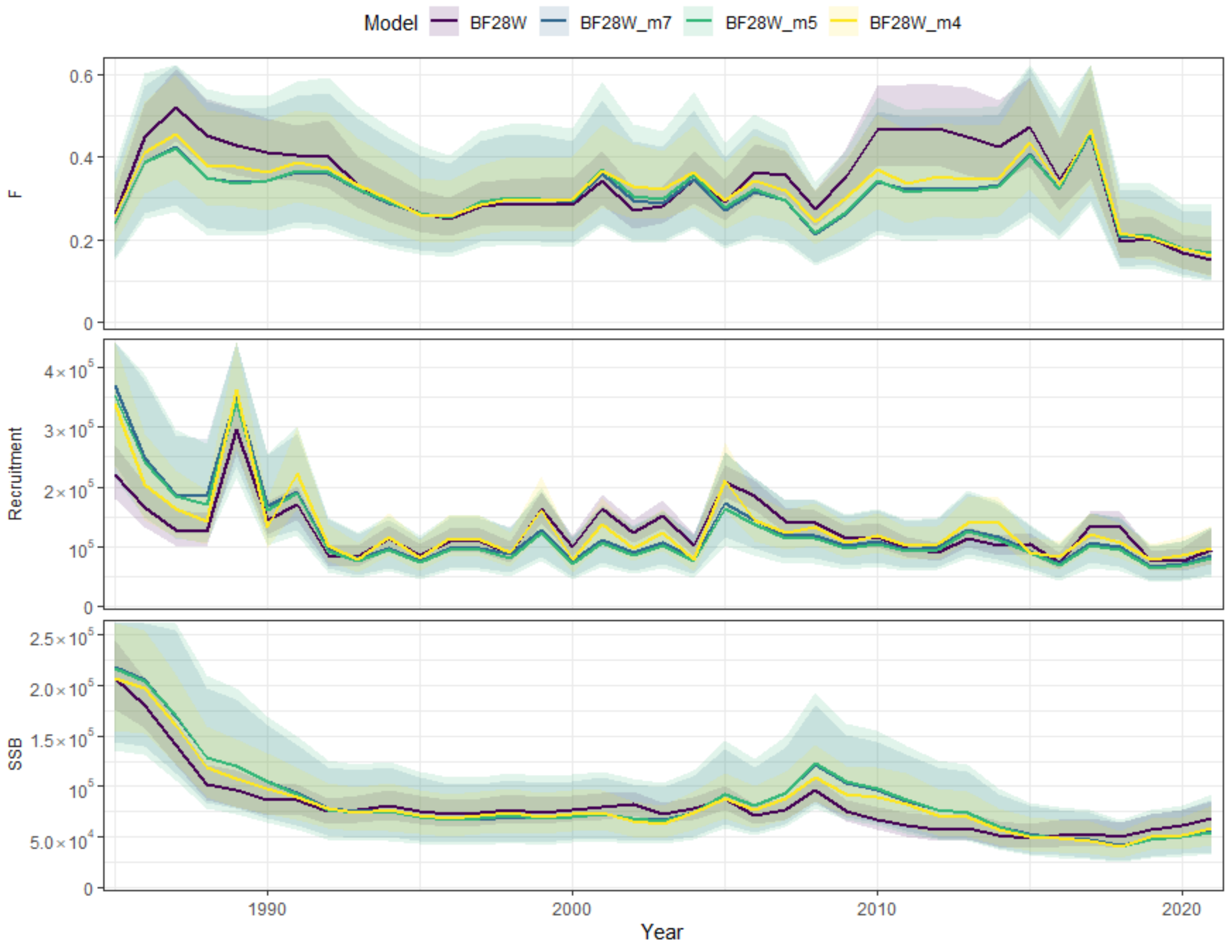


Figure 29. A comparison of SSB, F, and recruitment between the final bluefish model (BF28W_m7), the base statistical catch-at-age model (BF28W), and the top two closest models chosen by AIC.

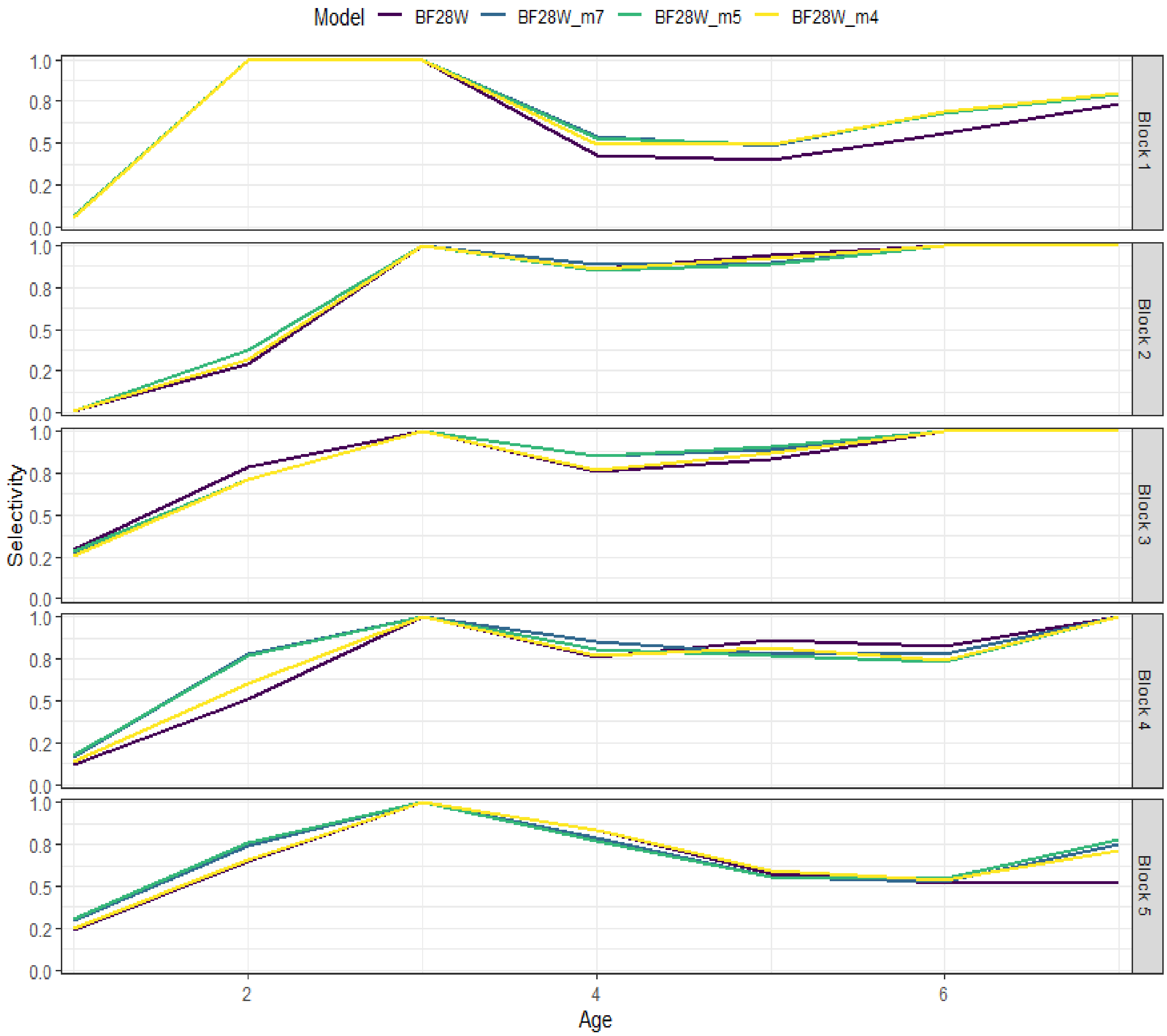


Figure 30. A comparison of the fleet selectivity block estimates between the final bluefish model (BF28W_m7), the base statistical catch-at-age model (BF28W), and the top two closest models chosen by AIC. Block 1: commercial fleet 1985-1999, Block 2: commercial fleet 2000-2021, Block 3: recreational fleet 1985-1999, Block 4: recreational fleet 2000-2010, Block 5: recreational fleet 2011-2021.

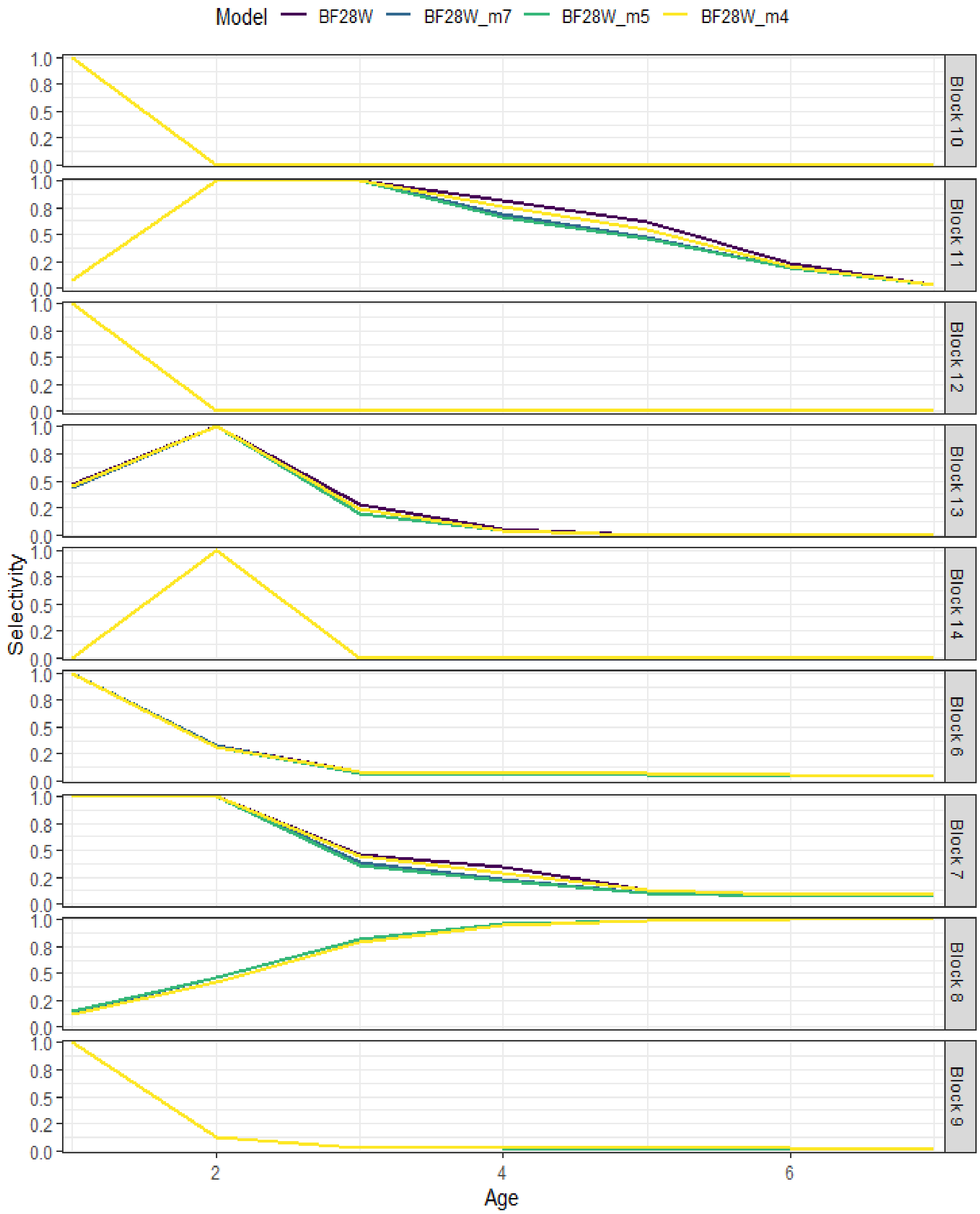


Figure 31. A comparison of the Index selectivity estimates between the final bluefish model (BF28W_m7), the base statistical catch-at-age model (BF28W), and the top two closest models chosen by AIC. Block 6: NEFSC Albatross, Block 7: NEFSC Bigelow, Block 8: MRIP CPA, Block 9: NEAMAP, Block 10: SEAMAP Age 0, Block 11: PSIGNS, Block 12: Conn YoY, Block 13: ChesMMAP, Block 14: SEAMAP Age1.

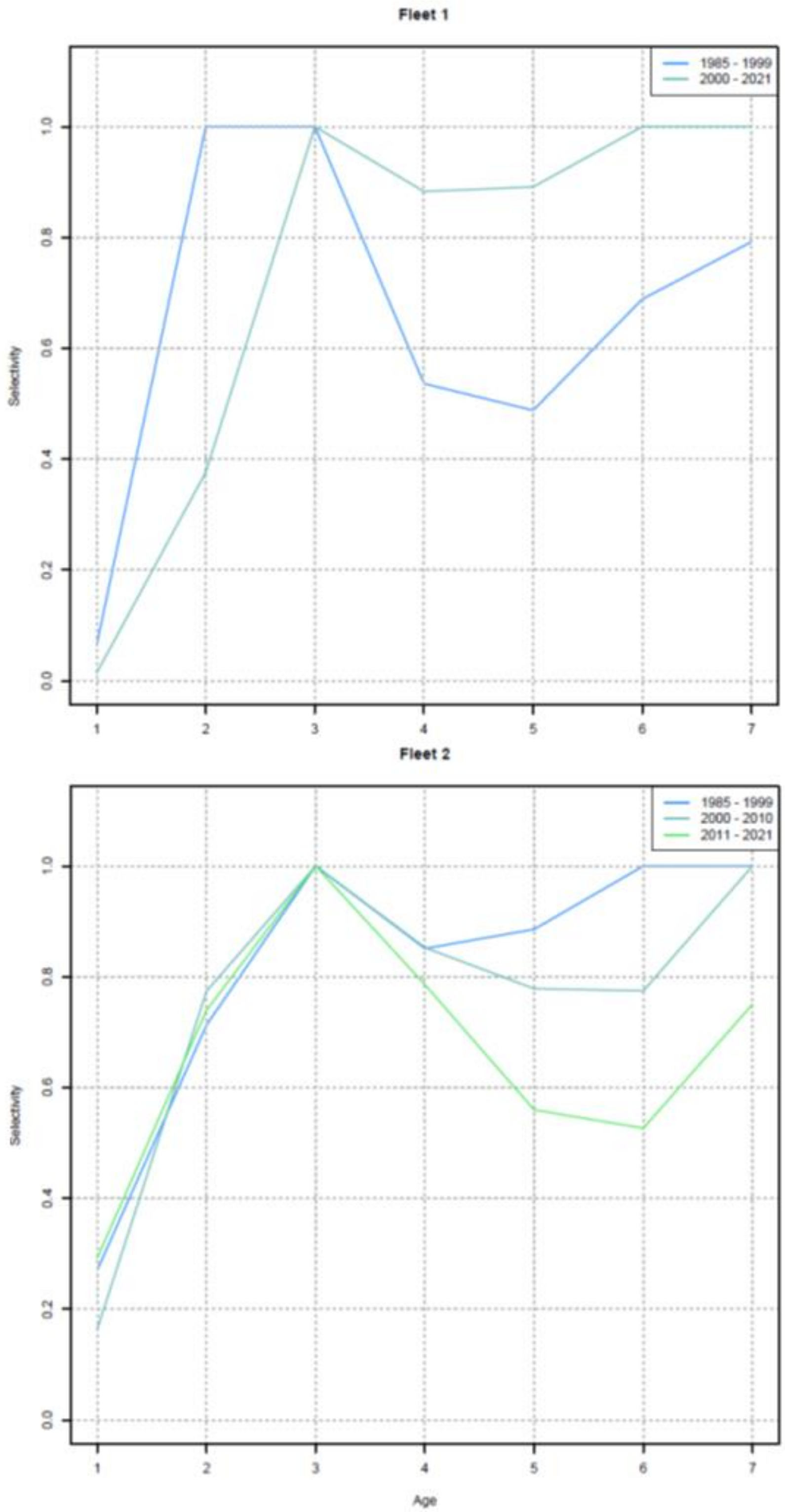


Figure 32. Selectivity estimates for the commercial (top) and recreational (bottom) fleets from the final model BF28W_m7.

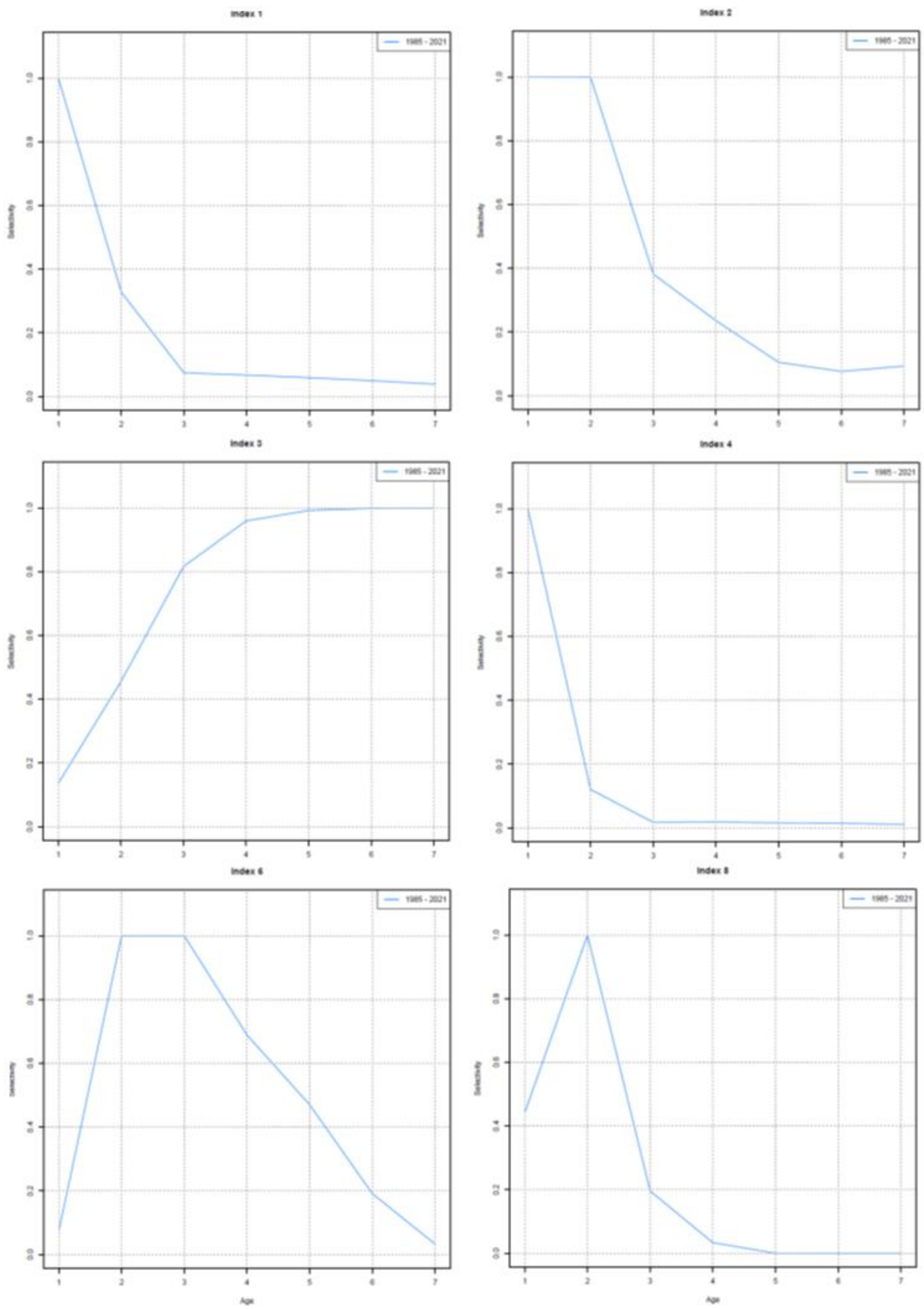


Figure 33. Final index selectivity estimates for all indices with age comps in the final model BF28W_m7. Index 1: NEFSC Alb, Index 2: NEFSC Big, Index 3: MRIP CPA, Index 4: NEAMAP, Index 6: PSIGNS, Index 8: ChesMMAP.

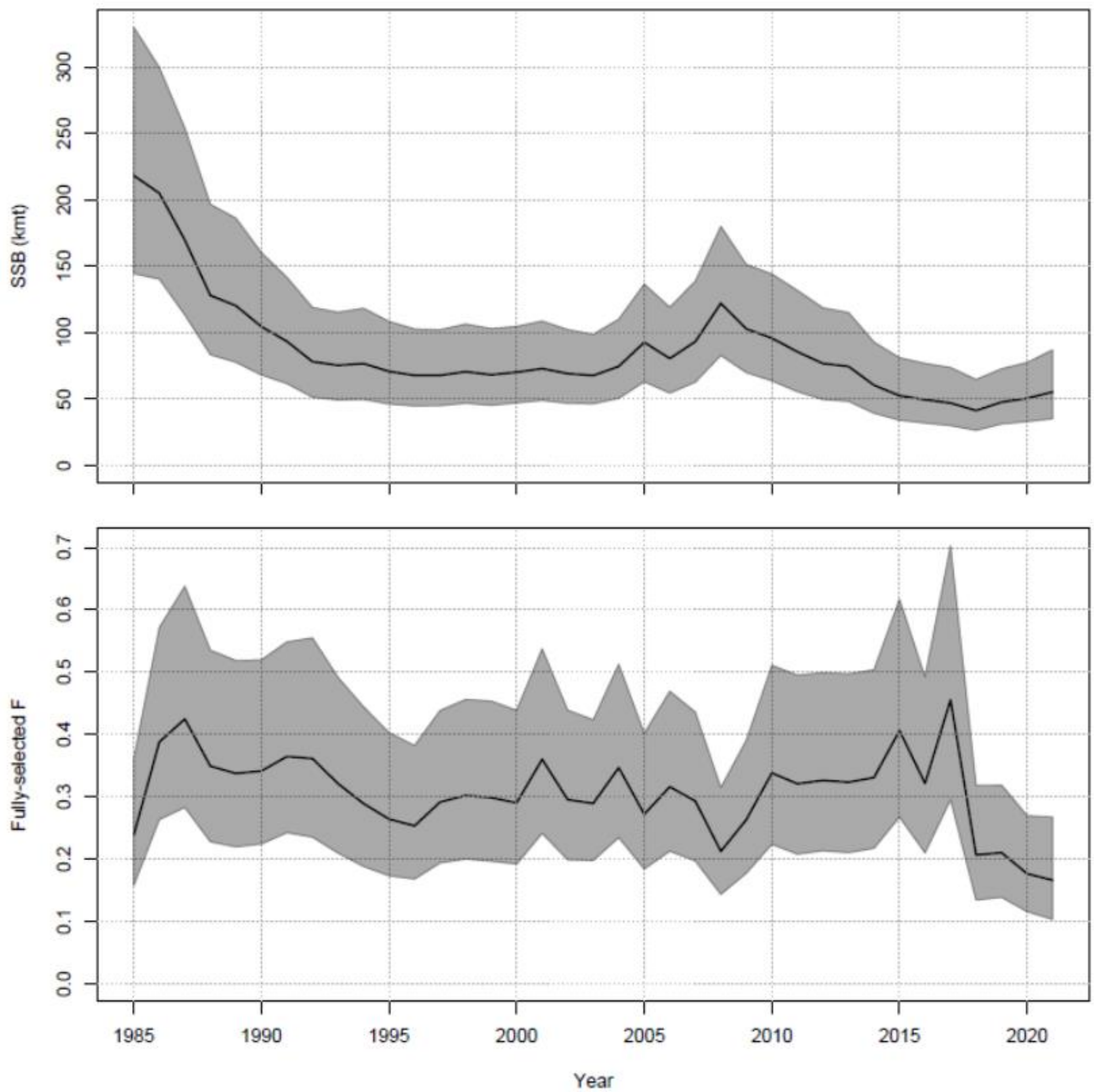


Figure 34. Spawning stock biomass (top) and fully selected fishing mortality (bottom) results from the final model BF28W_m7 from 1985-2021.

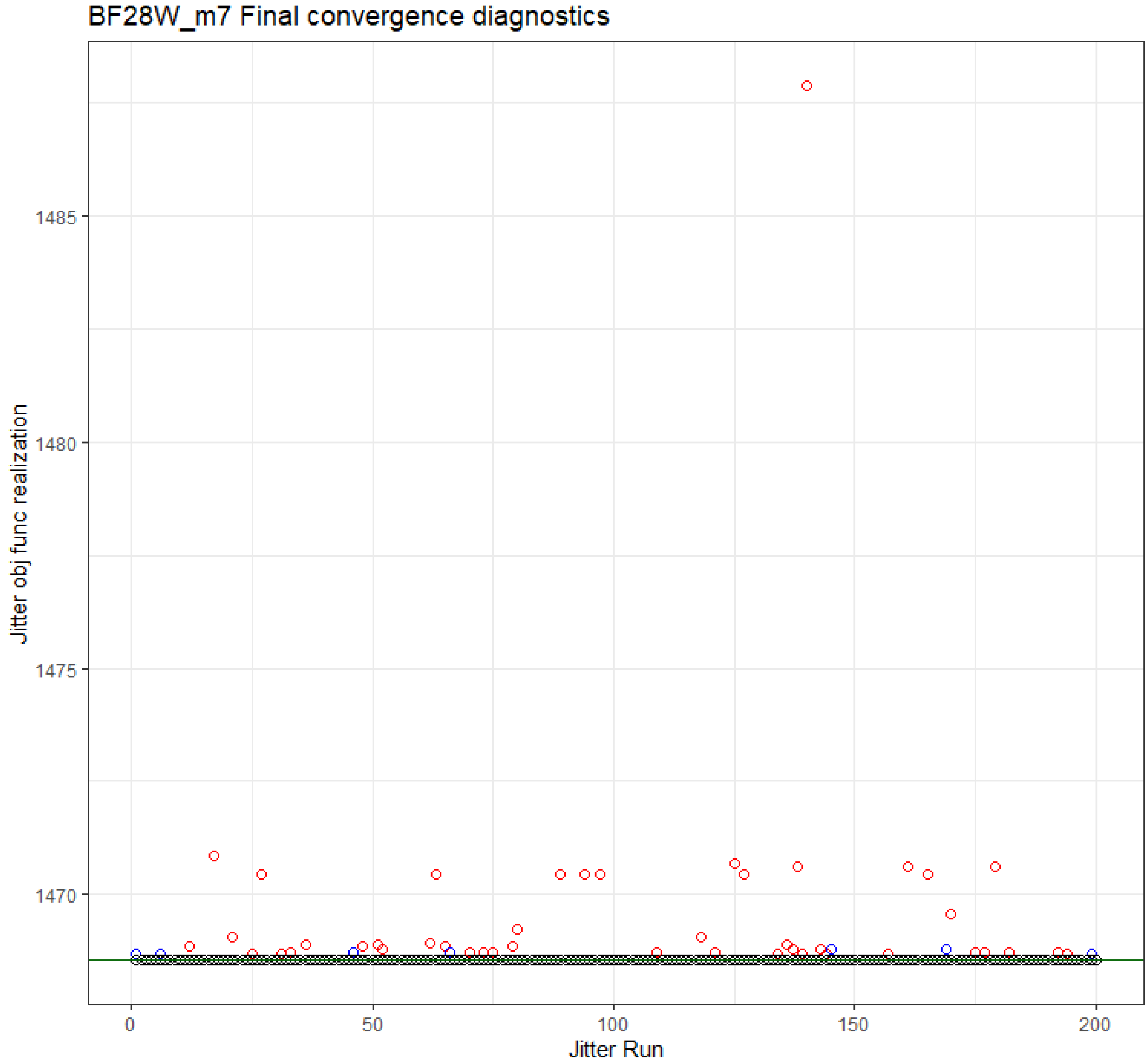


Figure 35. Jitter analysis to investigate convergence properties of final model BF28W_m7. 200 jitter runs at 3 different variance scales were run with convergence results shown (Scale 1 = black, Scale 2 = blue, Scale 3 = red). The green line indicates the original model objective function.

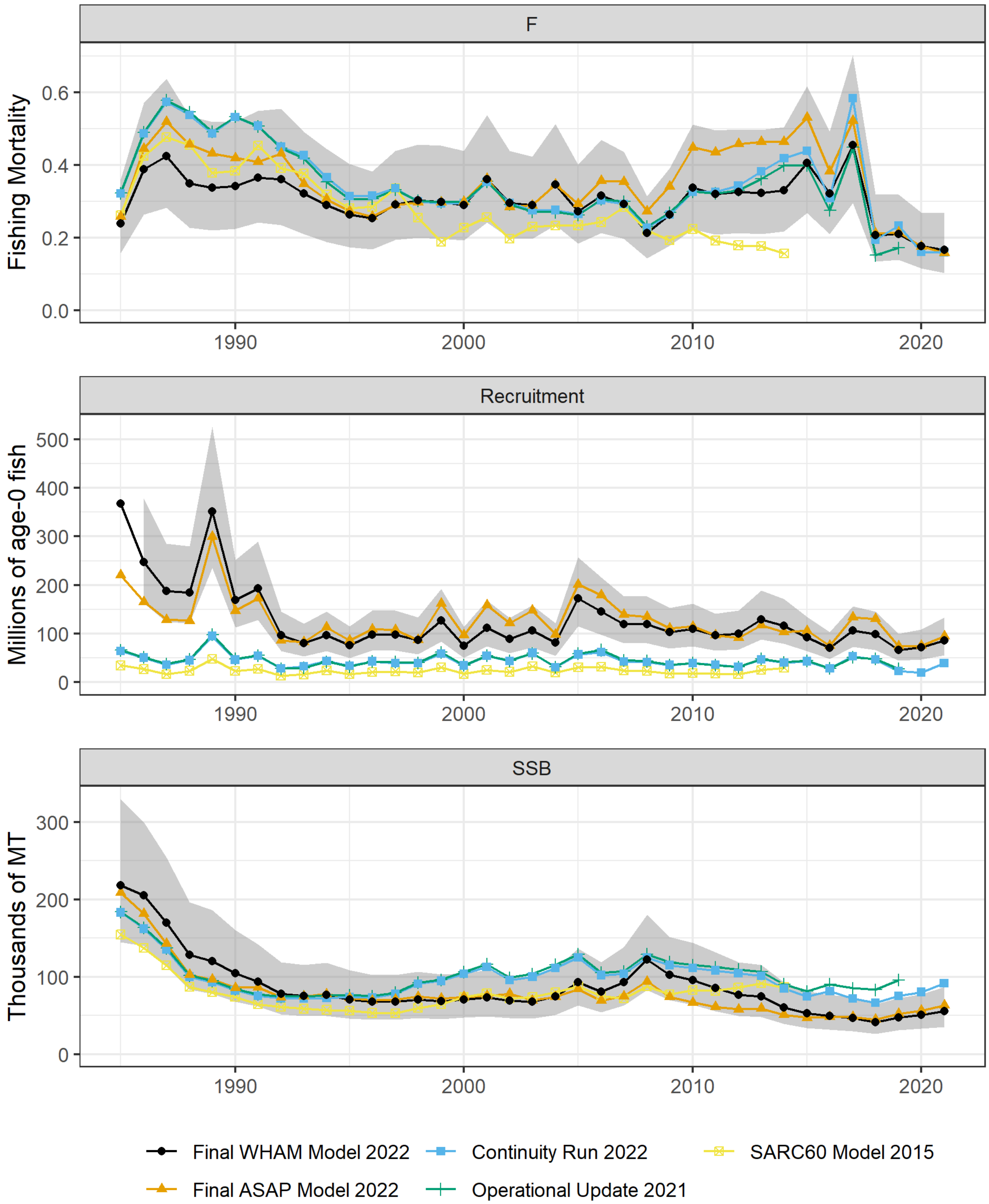


Figure 36. Historical retrospective of model results from the final WHAM model, the final ASAP final model, the continuity run update of the SAW60 model, the operational assessment in 2021, and the SARC60 benchmark model. The shaded area indicates the 95% confidence intervals for the final WHAM model estimates.

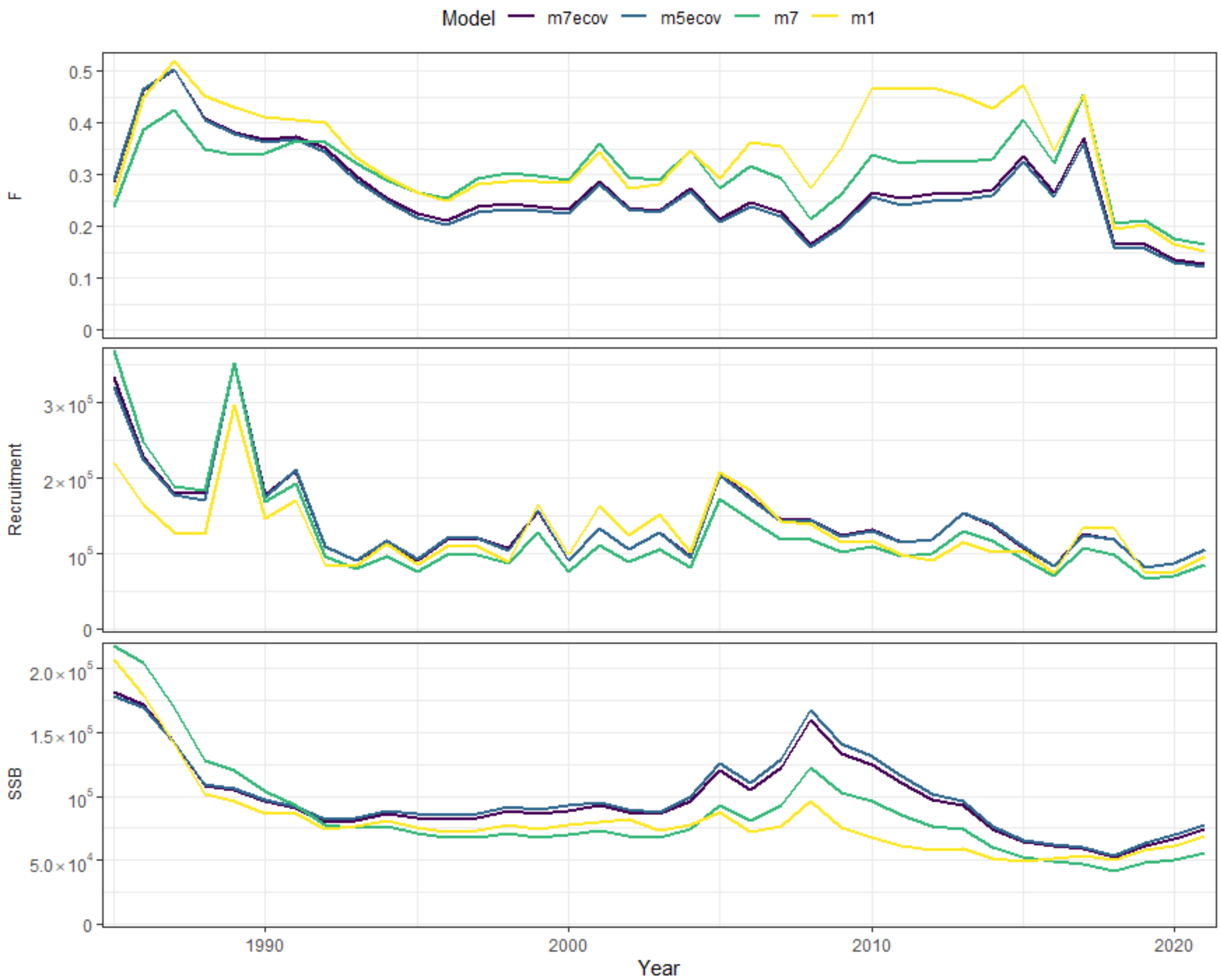
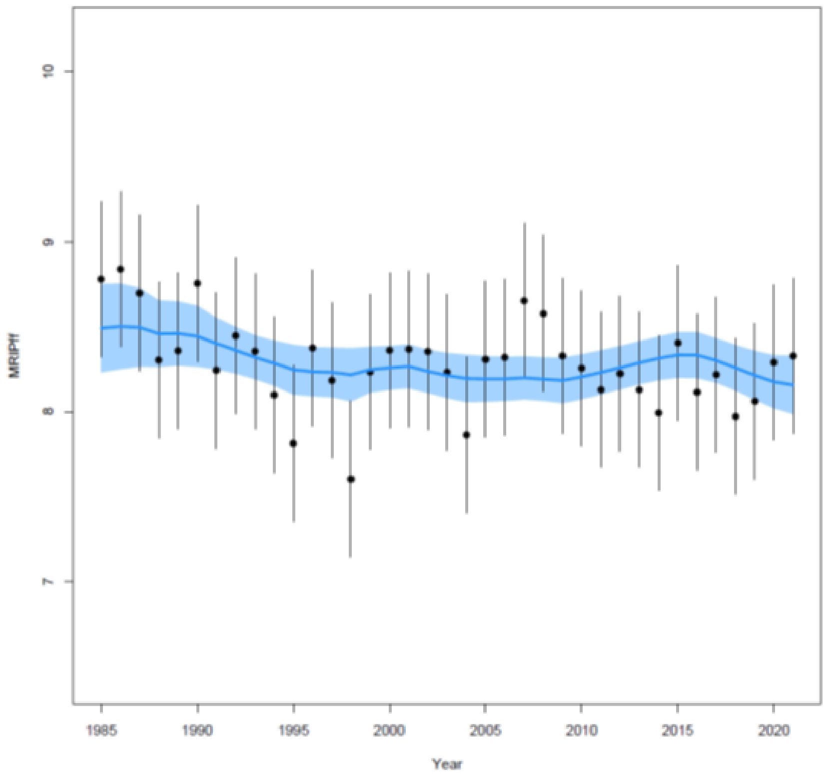


Figure 37. A comparison of the results from the base model (BF28W) and the top 3 models that include the environmental covariate on the MRIP index catchability.



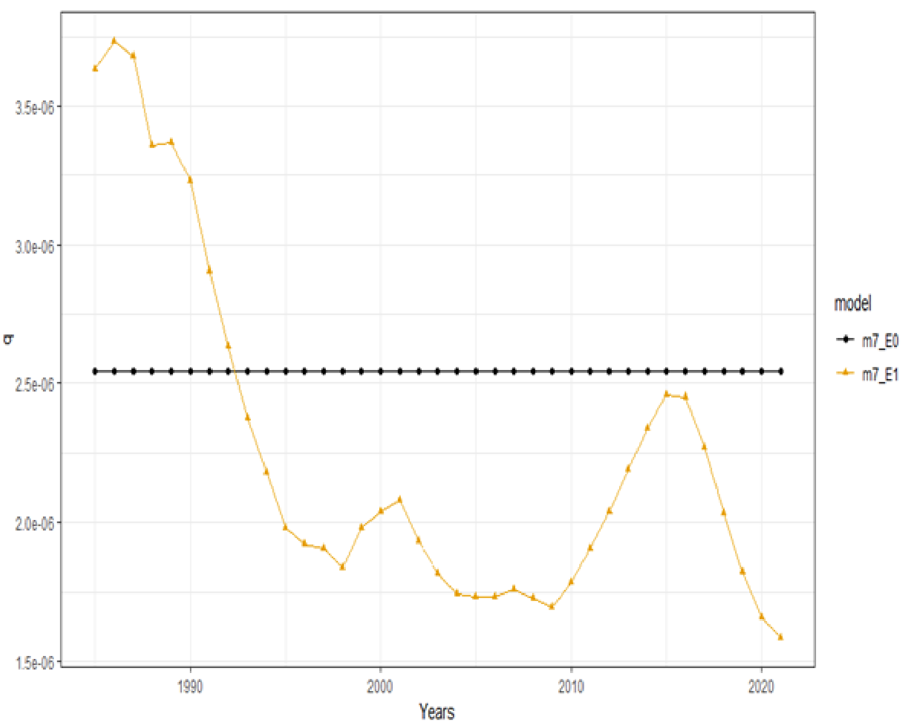


Figure 38. Fit to the forage fish index used as a covariate on the catchability (availability) of the MRIP index (top) and resulting trend in estimated catchability over time for the MRIP index (bottom).

Figure 39. Results from the sensitivity analyses testing the impact of removing each index on the model results.

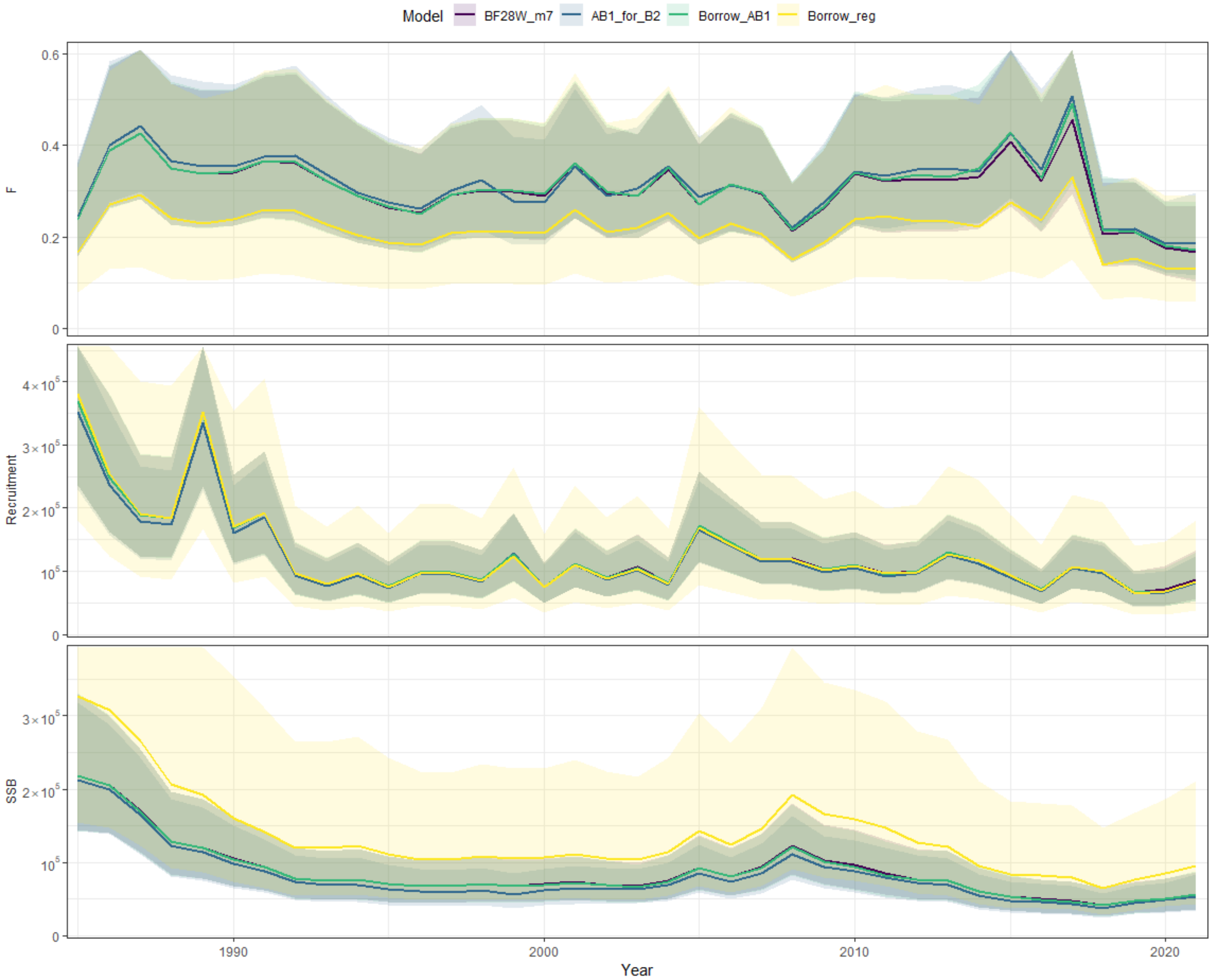


Figure 40. Results from the sensitivity analyses testing different borrowing rules for the recreational discard lengths. The borrowing from region model ("Borrow_reg") had poor convergence properties with bounded parameter estimates.

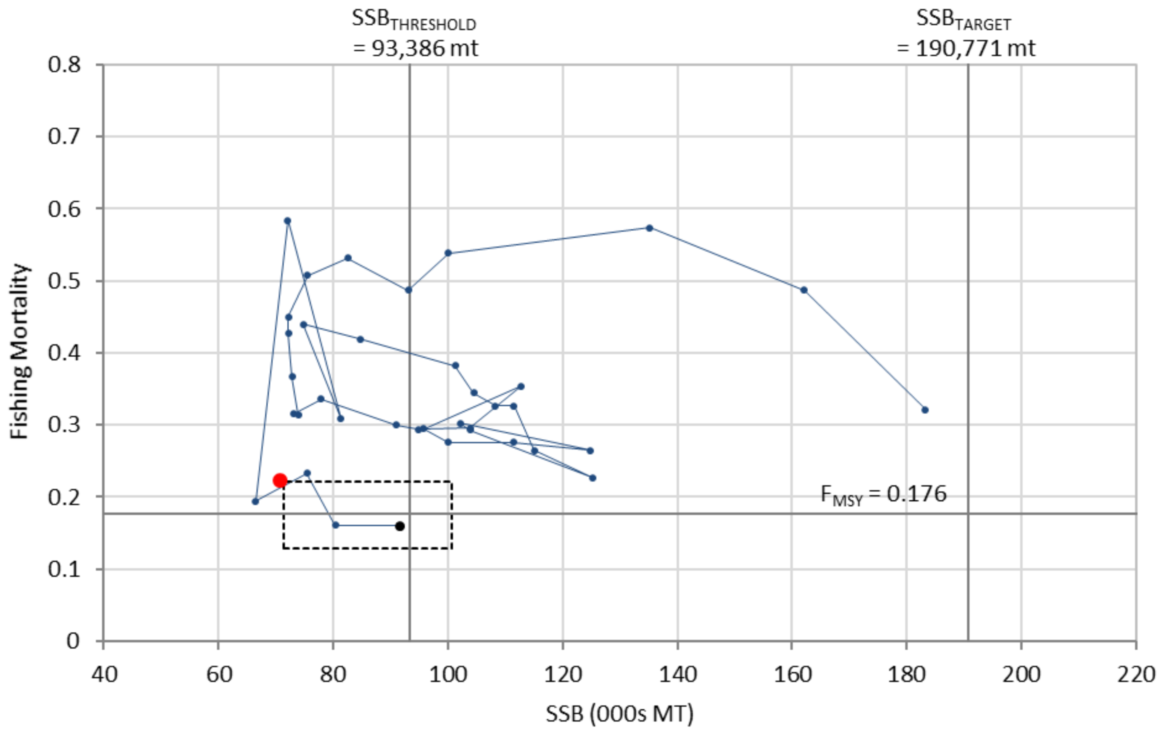


Figure 41. Stock status plot with status determination criteria for the model BF01, the continuity run. Dashed line indicates the 90% confidence region around the terminal year estimates of F and SSB. Red dot shows the retrospective adjusted terminal year values, which fall outside the confidence region and indicate that status should be determined using the retrospectively adjusted values.

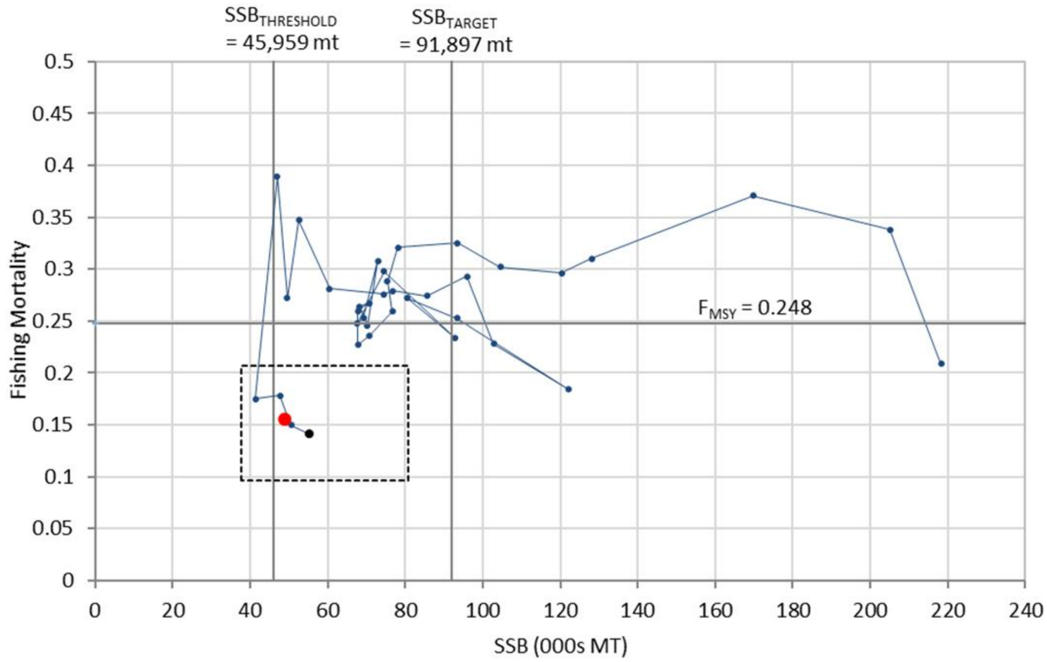


Figure 42. Stock status plot with status determination criteria for the final WHAM model for this assessment. Dashed line indicates the 90% confidence region around the terminal year estimates of F and SSB. Red dot shows the retrospective adjusted terminal year values, which fall within the confidence region and indicate that a retrospective adjustment is not necessary to determine stock status.

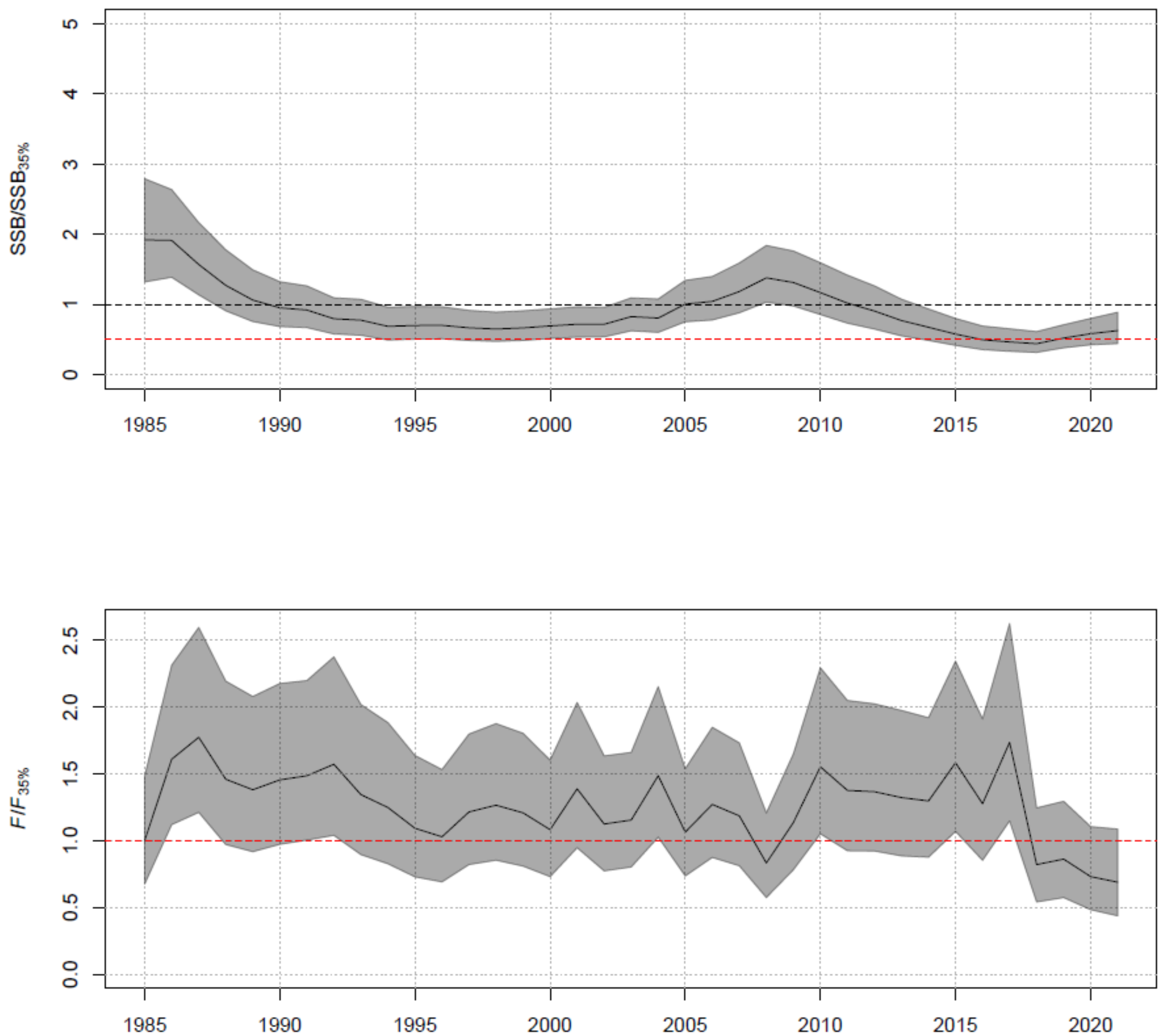


Figure 43. Final model SSB and fishing mortality in relation to SSB35% and F35%, the status determination criteria. The current bluefish stock is not-overfished and over-fishing is not occurring.

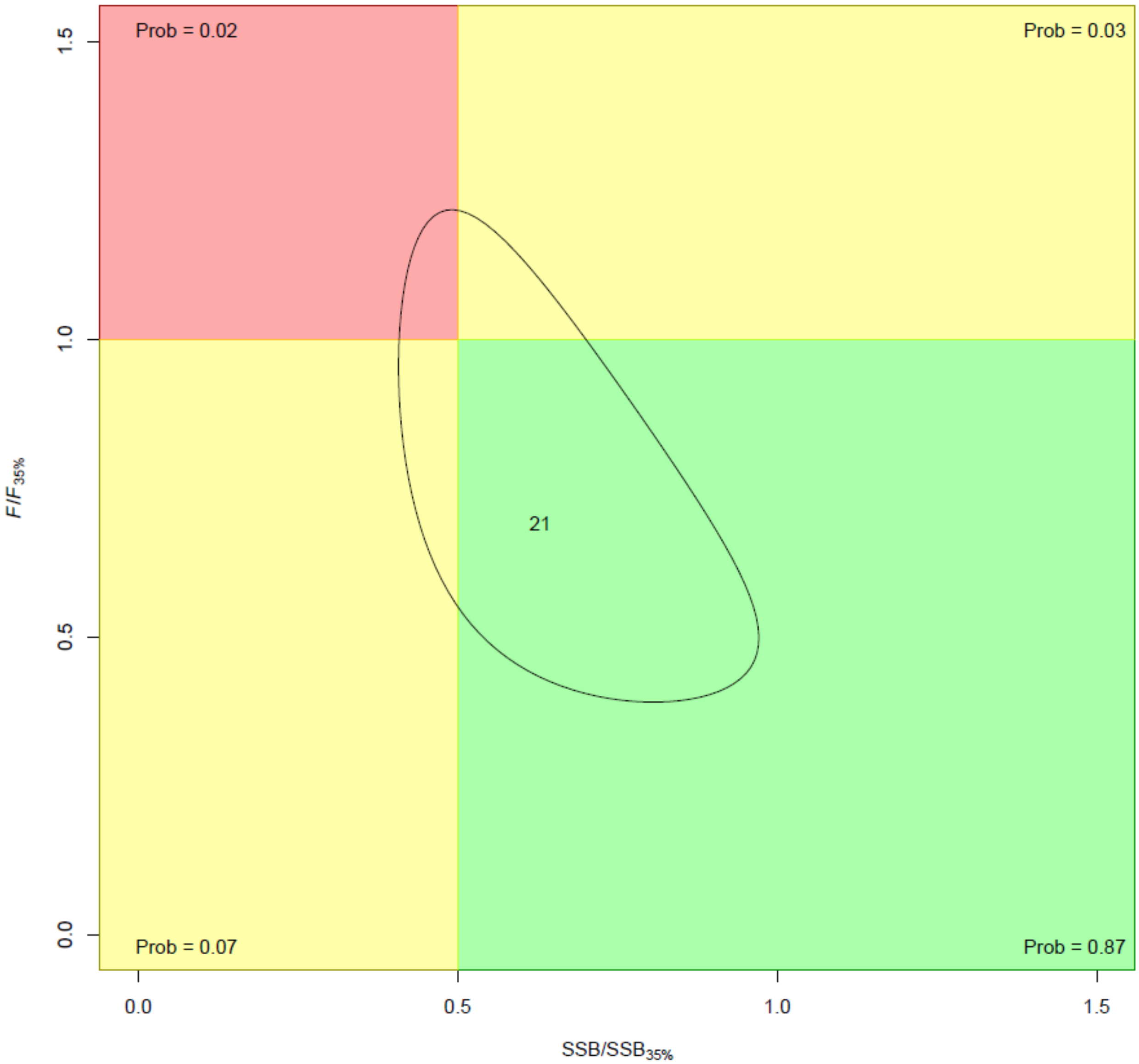


Figure 44. Kobe plot from final model (BF28W_m7) with stock status probability ellipse showing an 87% probability the bluefish stock is not overfished and over-fishing is not occurring. "21" indicates the 2021 point estimates of $SSB/SSB_{35\%}$ and $F/F_{35\%}$.

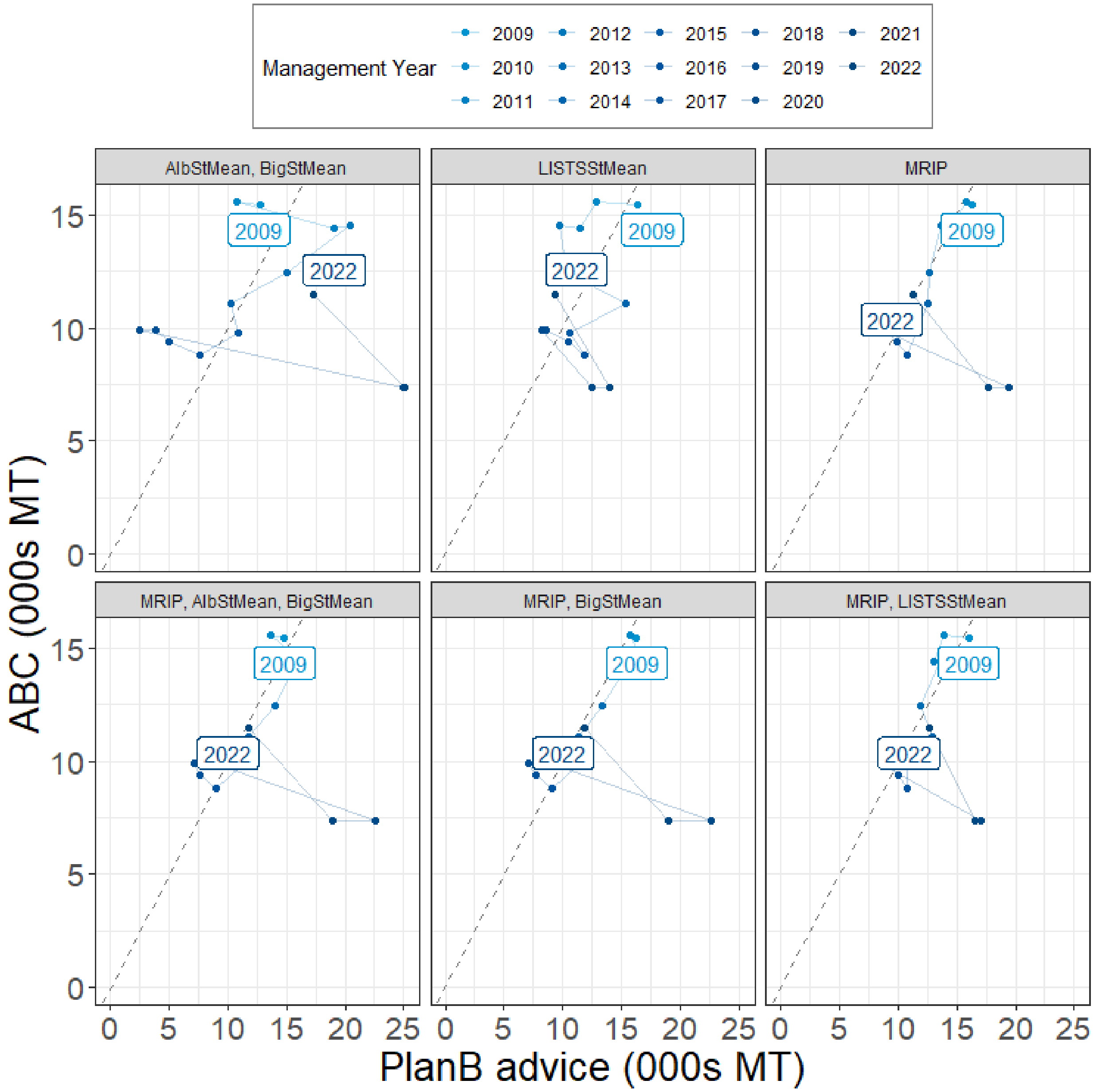


Figure 45. Results of applying a hypothetical Ismooth analysis compared to the actual ABC recommended for use in management. The dashed line is the 1-1 line.

12 LIST OF RELEVANT WORKING PAPERS

Working papers are available on the NEFSC data portal for this assessment, and at the hyperlinks below.

1. Tyrell et al. 2022. Bluefish Ecosystem and Socioeconomic Profile.
2. Valenti 2022a. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from American Littoral Society Fish Tagging Data
3. Tyrell 2022. Bluefish VAST Index Exploration.
4. Gaichas et al. 2022. Vector Autoregressive Spatio-Temporal (VAST) modeling of piscivore stomach contents, 1985-2021.
5. Truesdell et al. 2022. Life History Analyses for Bluefish.
6. Tyrell and Truesdell 2022. Natural mortality of bluefish.
7. Celestino et al. 2022a. Index of abundance exploration and development by the Bluefish Working Group's Fishery Independent Data Group.
8. Wood 2022a. Commercial and Recreational Data Collection and Analysis.
9. Drew 2022a. Recreational Data Changes for Bluefish, 2012-2021.
10. Drew 2022b. The Spatial Distribution of Bluefish (*Pomatomus saltatrix*): Insights from MRIP Data.
11. Valenti 2022b. Catch-and-Release Recreational Angling Mortality of Bluefish (*Pomatomus saltatrix*): Updated Analysis for 2022
12. Drew 2022c. Development of the Composite YOY Index for Bluefish.
13. Drew 2022d. A Fishery-dependent CPUE index for bluefish derived from MRIP data.
14. Celestino et al. 2022b. Development of Bluefish Age-Length Keys.
15. Wood 2022b. Bluefish Model Bridge-Building in ASAP.
16. Wood 2022c. ASAP diagnostic plots.
17. Wood 2022d. WHAM diagnostic plots.
18. Truesdell 2022. Alternative assessment plan.