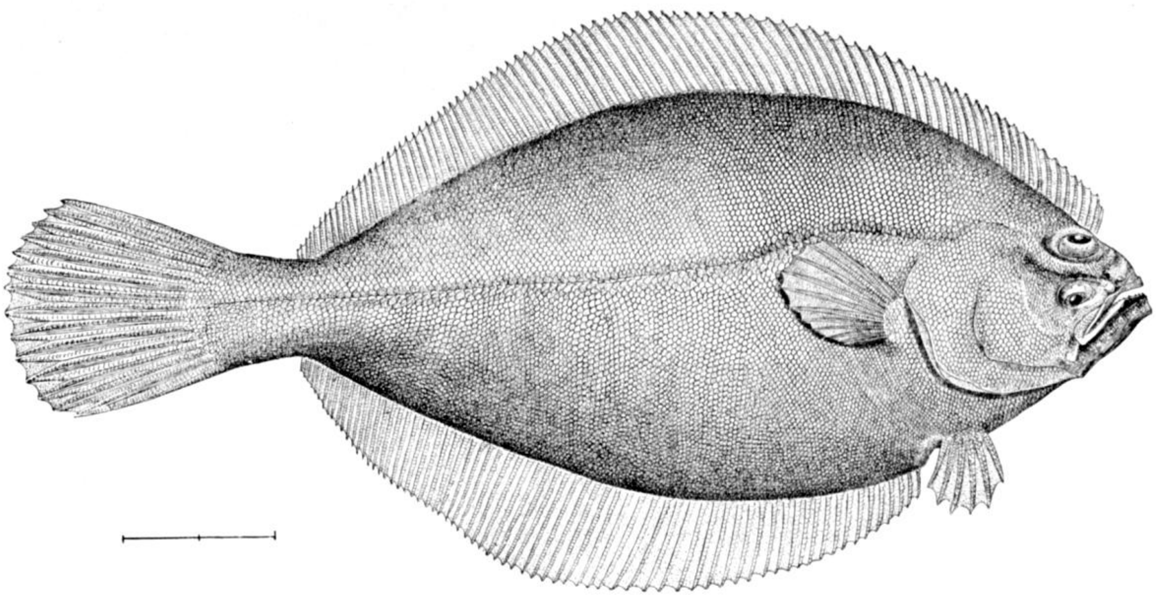
Report of the American Plaice Research Track Working Group



July 5, 2022

PARTICIPANTS

Working Group

NAME	AFFILIATION	E-MAIL
Larry Alade	NEFSC	larry.alade@noaa.gov
Steve Cadrin	SMAST	scadrin@umassd.edu
Jamie Cournane	NEFMC	jcournane@nefmc.org
Alex Hansell	NEFSC	alex.hansell@noaa.gov
Dan Hennen	NEFSC	daniel.hennen@noaa.gov
Lisa Kerr	GMRI	lkerr@gmri.org
Rajeev Kumar	Canada DFO	rajeev.kumar@dfo-mpo.gc.ca
Tyler Pavlowich	NEFSC	tyler.pavlowich@noaa.gov
Dave Richardson	NEFSC	david.richardson@noaa.gov

Contributors and Participants in Working Group Meetings

NAME	AFFILIATION	E-MAIL
Irenene Andrushchenko	Canada DFO	irene.andrushchenko@dfo-mpo.gc.ca
Jamie Behan	GMRI	jbehan@gmri.org
Russ Brown	NEFSC	russell.brown@noaa.gov
Cole Carrano	SMAST	ccarrano@umassd.edu
Toni Chute	NEFSC	toni.chute@noaa.gov
Kiersten Curti	NEFSC	kiersten.curti@noaa.gov
Josh Dayton	NEFSC	joshua.dayton@noaa.gov
Alexander Dunn	NEFSC	alexander.dunn@noaa.gov
Angela Forristall	NEFMC	aforristall@nefmc.org
Robin Frede	NEFMC	rfrede@nefmc.org
Keith Hankowsky	SMAST	khankowsky@umassd.edu
Amanda Hart	GMRI	ahart2@gmri.org
Shakira Goffe	Univ. Maryland	scgoffe@umes.edu

Max Grezlik	SMAST	mgrezlik@umassd.edu
Andy Jones	NEFSC	andrew.jones@noaa.gov
Chris Legault	NEFSC	chris.legault@noaa.gov
Mackenzie Mazur	GMRI	mmazur@Gmri.org
Rich McBride	NEFSC	richard.mcbride@noaa.gov
Lucy McGinnis	SMAST	lmcginnis@umassd.edu
Anna Mercer	NEFSC	anna.mercer@noaa.gov
Tim Miller	NEFSC	timothy.j.miller@noaa.gov
Liz Moore	GMRI	emorre@gmri.org
Paul Nitschke	NEFSC	paul.nitschke@noaa.gov
Jackie Odell	NE Seafood Coalition	jackie@northeastseafoodcoalition.org
Cate O'Keefe	Fishery Applications	okeefe@fisheryapps.com
Rebecca Peters	MEDMR	rebecca.j.peters@maine.gov
Phil Politis	NEFSC	philip.politis@noaa.gov
Eric Robillard	NEFSC	eric.robillard@noaa.gov
Melissa Sanderson	CCCFA	melissa@capecodfishermen.org
Ashley Silver	Univ. Maryland	acsilver@umes.edu
Mark Terceiro	NEFSC	mark.terceiro@noaa.gov
Michele Traver	NEFSC	michele.traver@noaa.gov
Susan Wigley	NEFSC	susan.wigley@noaa.gov

TABLE OF CONTENTS

EXECUTIVE SUMMARY

A research track assessment for American plaice was planned for peer review in 2023. The Working Group was formed in June 2021 and met over the next year to address its terms of reference (TORs). This report represents consensus of the Working Group and includes contributions from Working Group member and participants.

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*."

A review of the scientific literature on American plaice identified that temperature and depth were potential drivers of American plaice distribution. Plaice were observed at greater depths during winter months compared to summer months, and large plaice were observed at deeper depths than small plaice. Occupied depth increased in response to changing ocean conditions.

Temperature, as well as the North Atlantic Oscillation (NAO), were significantly correlated with recruitment rate (recruits per spawner), with the greatest recruitment rates occurring at extreme cold temperature. Warmer temperatures were associated with increased natural mortality, accelerated growth rates, reductions in body size, and earlier ages at maturity of American plaice. Changes in growth rates may indirectly cause natural mortality though potential increases in physiological stress, predation and starvation. Therefore, suitable habitat is expected to decrease with continued changes in ocean conditions.

To identify ecosystem and climate influences on plaice, generalized additive models (GAMs) were used to examine associations between their population dynamics and environmental variables in the Gulf of Maine region, including bottom water temperature, the NAO, the Atlantic Multidecadal Oscillation (AMO), and the Gulf Stream Index (GSI). Mean latitude of survey catches shifted southerly during the 1980s and 1990s then northerly in the last two decades and was significantly related to spawning stock biomass and bottom temperature.

Mean depth of survey catch in the fall survey decreased in the 1980s, gradually shifted to deeper water since the late 1980s, and was positively related to bottom temperature and negatively related to NAO (Figure 1.3). Survey indices of recruitment rate increased since the 1980s and were significantly related to the AMO. Weight-at-age was significantly related to AMO, GSI, spawning stock biomass, and bottom temperature. Mean fish condition from the fall survey was significantly related to spawning stock biomass, bottom temperature, and AMO.

These ecosystem influences were considered in several subsequent ToRs. Decadal shifts in growth were recognized for ToR2 age composition and weight at age, and other environmental factors (temperature, depth) were explored in fishery catch standardization. Factors of survey catchability were investigated in ToR3 for integrating inshore and offshore surveys, and results were considered in the selection of survey indices for the assessment model. Decadal shifts in growth were addressed in ToR4 by applying empirical weight-at-age and exploring environmental covariates of recruitment and survey catchability. Environmental factors were considered for assumptions about recruitment and weight-at-age in reference points (ToR5) and projections (ToR6). Finally, environmental analyses led to several research recommendations in ToR7.

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*."

Since 1980, annual catch of American plaice has ranged from 1,048 mt to 15,540 mt. Commercial landings were the predominant source of fishery removals, averaging 88% of the catch between 1980 and 2019. Most landings were from the Gulf of Maine and the northern edge of Georges Bank. Almost all landings were from trawl trips, and the proportion of landings from the largest vessels gradually increased in the last decade.

Most discarded catch was from the large-mesh trawl fleet, with considerable discards from the shrimp fishery in the late 1980s and early 1990s and relatively few discards from the small-mesh trawl, gillnet and scallop dredge fisheries. Discards were a relatively large proportion of total US catch until 1992, ranging 15% to 40% of total catch 1980-1991, but were less than 10% of total catch since 2014, resulting from a decrease in the minimum legal size.

The age composition of fishery catch was relatively stable since the 1980s. Catch was composed of primarily ages 2-6. Cohort tracking was relatively strong, as measured by positive correlations of catch-at-age by year-class from age-1 to age-10 and several apparently strong year-classes (e.g., 1987, 1993, 2004, 2013) contributed to catch over several years. Mean weight-at-age of the catch was relatively stable for ages 1-6 but decreased over time for older ages since the 1990s.

Several alternative series of standardized fishery catch rates were developed using models that accounted for fishing location, season, vessel tonnage, depth, and price. Results from all approaches produced similar time series that were moderately to strongly correlated with survey biomass indices. Standardized catch rate indices were explored as stock size indices.

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*."

Several fishery independent surveys are available to index American plaice stock size and age composition. Northeast Fisheries Science Center (NEFSC) survey indices suggest that the stock has fluctuated with peak abundances in the early 1960s, late 1970s-early 1980s, and the late 2010s. NEFSC survey indices of abundance at age suggested the same strong year-classes as the fishery catch-at-age (1987, 1993, 2004, 2013) and had good cohort tracking among adjacent ages and years.

Two inshore surveys, the Massachusetts Division of Marine Fisheries trawl survey (1978-2019) and the Maine-New Hampshire trawl survey (2000-2019), sample US state waters in the Gulf of Maine. Age data were not available for either survey. The Massachusetts survey was excluded as an index of abundance in the most recent stock assessment because the inshore index conflicted with other information and appeared to result from plaice shifting to deeper waters and decreasing availability of the resource to the survey (NEFSC 2022).

Data from NEFSC and both inshore surveys were integrated into a single index using spatiotemporal analysis with depth and bottom temperature effects. Over the time series, the geographic distribution of catches from all surveys was variable in both the spring and fall with periods of northeast and southwest movement of the center of gravity. Since the 1960's, the effective area occupied has decreased in the spring and fall by an average rate of 178 and 81 km^2/year, and depth was the strongest correlate for both spring and fall distribution changes. Trends in the integrated index reflected those in NEFSC surveys because they have larger spatial coverage.

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*."

A range of approaches to stock assessment modeling were explored for this research track assessment. Assumed biological parameters were reconsidered. Several forms of integrated assessment model were applied, including conventional statistical catch at age, statistical catch at age with length-based selectivity and discard estimation, and a state-space model with environmental covariates.

Previous stock assessments of American plaice in US waters, assumed a lifetime constant natural mortality rate (M=0.2) based on relative abundance of ages 9+ from an unexploited plaice population in the Gulf of St. Lawrence, Canada. However, American plaice in US waters have significantly different life history than those in Canadian waters. Several life-history based M estimators were applied to the available information for American plaice in US waters.

Estimates based on maximum age and growth were relatively consistent and supported a revised assumption of natural mortality (M=0.3).

The American plaice stock assessment has been conducted using a Virtual Population Analysis (VPA) since 1992. Several forms of statistical catch-at-age model applications were developed for this research track assessment including an Age Structured Assessment Program (ASAP), Stock Synthesis (SS) and the Woods Hole Assessment Model (WHAM). Among these alternative approaches, the Working Group proposed WHAM as the basis for status determination and fishery management advice. WHAM is a state-space age-structured stock assessment model that fits to aggregated catch, stock index, and age composition data, and can include process errors and environmental covariates. Candidate model runs fit the available data well, had relatively high prediction skill and retrospective consistency. The proposed base run fit 1980-2019 fishery catch and age composition, and NEFSC spring and fall survey indices and age composition, modeled as separate series for the Albatross and Bigelow surveys. All candidate model runs indicate that the stock was relatively abundant at the start of the assessment series, decreased in the 1980s from relatively high fishing mortality, and has gradually rebuilt since then to relatively high abundance between 2014-2019, while fishing mortality decreased over the same period. These general results were supported by all WHAM candidate runs as well as ASAP, SS and VPA. Among models and runs with comparable data and assumptions, results were similar.

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*."

Throughout its assessment and management history, the fishery for American plaice in US waters has been managed based on maximum sustainable yield (MSY) proxies derived from dynamic pool reference points from yield and spawning biomass per recruit analyses. The Working Group re-examined the stock-recruit relationship to confirm the justification for using $F_{40\%}$ and $SSB_{F40\%}$ as MSY proxy reference points based on the entire time series of recruitment and recent 5-year estimates of selectivity and observations of weight-at-age.

The proposed WHAM run was used to derive integrated estimates of $F_{40\%}$ (0.42) and $SSB_{F40\%}$ (18,000 mt). Based on these updated reference points and stock assessment results, overfishing from 1980 to 1998 depleted the stock to an overfished state in the late 1980s, but fishing mortality has been less than $F_{40\%}$ since 1998, and the stock rebuilt to be significantly greater than $SSB_{F40\%}$ in 2019. According to these analyses, the current stock is not overfished, and overfishing is not occurring.

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 proposed WHAM run was used to produce integrated projections that account for uncertainty in all estimated parameters, 2020 abundance at age, and future recruitment. Projections assumed future recruitment based on the entire time series of recruitment, recent 5-year estimates of selectivity and observations of weight-at-age, and maturity-at-age from the entire time series, which is consistent with revised reference points and the approach used for other New England groundfish stocks.

Projection results through 2022 are presented for demonstration, but the WHAM application will be updated with 2020-2021 data in the 2022 management track process, and projections will also be updated. Provisional projections assumed 2020-2021catches and projected scenarios included 1) fishing at the estimate of $F_{40\%}$ in 2022 to demonstrate overfishing limit projections, 2) fishing at 75%$F_{40\%}$ in 2022 to demonstrate acceptable catch projections, 3) fishing at the estimate of 2019 F in 2022 to demonstrate a status quo F scenario, and 4) no fishing, as a basis for comparison to harvest scenarios.

All provisional projections resulted in an increase in catch in 2022. Projected spawning stock biomass decreased for all 2022 F>0 scenarios but remained well above $SSB_{F40\%}$. The projected stock decrease is related to decreased recent recruitment.

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*."

All previous research recommendations were addressed by this research track assessment, with the exception of processing age samples from the Massachusetts inshore survey. The Working Group proposes several new research recommendations for improving data and models. Streamlined estimation of commercial catch, including integration of information on from electronic monitoring, would promote reproducibility. The Gulf of Maine scallop fishery should be included in discards as it expands. Spatiotemporal integration of federal and state surveys should be explored further. Several technical aspects of specifying environmental effects in stock assessment models also need development.

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*."

Several empirical and analytical approaches were considered as potential contingency plans if the proposed assessment model is deemed inappropriate for providing management advice, either as a conclusion of research track peer review or subsequently in the management track process. Performance evaluation of alternative approaches by the Index-Based Methods

Working Group suggested that survey-based methods and catch curves may be appropriate for stocks near B_{MSY}.

Efficiency-adjusted swept-area biomass estimates are available for plaice in US waters and could be used as a contingency plan for monitoring the stock. However, the age composition of the fishery and NEFSC surveys as well as model estimates of survey selectivity suggest that older plaice are not fully selected by the surveys. Therefore, catch biomass from the fishery is not directly comparable to survey estimates of biomass for a simple catch/biomass exploitation ratio. Biomass reference points derived from area-swept approaches and dynamic pool models would require selectivity assumptions and information on survey catchability of recruits and adults.

Catch curves were also considered as a contingency approach to stock assessment. However, estimates of total mortality rate from fishery catch-at-age were less than the assumed rate of natural mortality for recent year-classes, producing negative estimates of fishing mortality. Survey catch curves would violate the assumption of constant mortality, because older plaice do not appear to be fully selected by the surveys.

Considering the decades of information available for fishery landings and discards, multiple survey indices, and age composition, data-limited approaches based on surveys or catch curves would not include all the available information. Based on results of simulation testing by the Index-Based Methods Working Group, apparent domed selectivity of surveys for plaice, and challenges deriving reference points for survey biomass, the Working Group recommends that if the proposed assessment approach does not meet the standards of peer review or is rejected in a future management track assessment, an alternative model be developed to integrate information from catch, age composition and indices.

WORKING GROUP PROCESS

Each region of the US developed a stock assessment peer review process to determine best scientific information available to support management of marine fisheries in the region (NOAA 2013). The Northeast Region Coordinating Council (NRCC) consisting of members from the Atlantic States Marine Fisheries Commission, Greater Atlantic Regional Fisheries Office, Mid-Atlantic Fishery Management Council, New England Fishery Management Council, and Northeast Fisheries Science Center, developed an enhanced stock assessment process to improve the quality of assessments, allow more improvement to occur within the routine assessment process, and provide more strategic and longer-term planning for research and workload management. The process involves two tracks of assessment work: 1) a management track that includes routine updates of previously approved assessment methods to support regular management actions (e.g., annual catch limits), and 2) a research track that allows comprehensive research and development of improved assessments on a stock-by-stock or topical basis. The process provides opportunities for input and engagement from stakeholders and research partners, and a longer-term planning horizon to carry out research to improve assessments on both tracks, but particularly the research track. The research track assessment process is the region's approach to implementing the nation's 'next generation stock assessment enterprise' (Lynch et al. 2018). The NRCC develops and negotiates long-term management track cycles for each stock and a five-year research track schedule through time by the NRCC (https://s3.amazonaws.com/nefmc.org/Stock-assessment-process-June2020.pdf).

A research track assessment for American plaice was planned for peer review in 2022 to be followed by a management track assessment in summer 2022, updated with data through 2021 to support catch advice for 2023-2024. The Working Group was formed in June 2021 and met over a series of hybrid meetings:

1. July 17, 2021 (New Bedford, MA and WebEx) – terms of reference, 2019 stock assessment, general approach, tasking, and meeting plans.
2. September 10, 2021 (New Bedford, MA and WebEx) – biological information
3. September 24, 2021 (Gloucester, MA and WebEx) – fishermen perspectives

4. November 15, 2021 (New Bedford, MA and WebEx) – ToR1 environmental effects and ToR2 fishery catch rates
5. November 19, 2021 (New Bedford, MA and WebEx) – ToR2 commercial landings
6. December 6, 2021 (New Bedford, MA and WebEx) – ToR2 commercial discards
7. December 10, 2021 (New Bedford, MA and WebEx) – ToR3 surveys
8. January 28, 2022 (New Bedford, MA and WebEx) – ToR2-3 data decisions and planning for ToR4 assessment models
9. February 16, 2022 (WebEx) – ToR4 model development and planning
10. March 10, 2022 (New Bedford MA and WebEx) – ToR4 model development and planning
11. April 4, 2022 (New Bedford MA and WebEx) – ToR4 model development and planning
12. April 21, 2022 (New Bedford MA and WebEx) – ToR4 model development and planning
13. May 12, 2022 (New Bedford MA and WebEx) – ToR4 model development and planning
14. June 24, 2022 (New Bedford MA and WebEx) – ToR4 completion and ToRs 5-7
15. June 30, 2022 (New Bedford MA and WebEx) – completion of ToRs 5-7 and drafting

The Working Group welcomed participation, input and contributions from non-Working Group members. In advance of meetings, working papers (Appendix A) and presentations were distributed and reviewed for working group discussions and consensus decisions. A shared drive with background documents and working papers is available for Working Group member, participants and reviewers: https://drive.google.com/drive/folders/17e_34Am2w0zG4V9Mwpiv-_VfMEcL-3Yj. The Working Group Chair produced a draft report by compiling information in working papers, meeting minutes and presentations, and the draft report was reviewed by the Working Group.

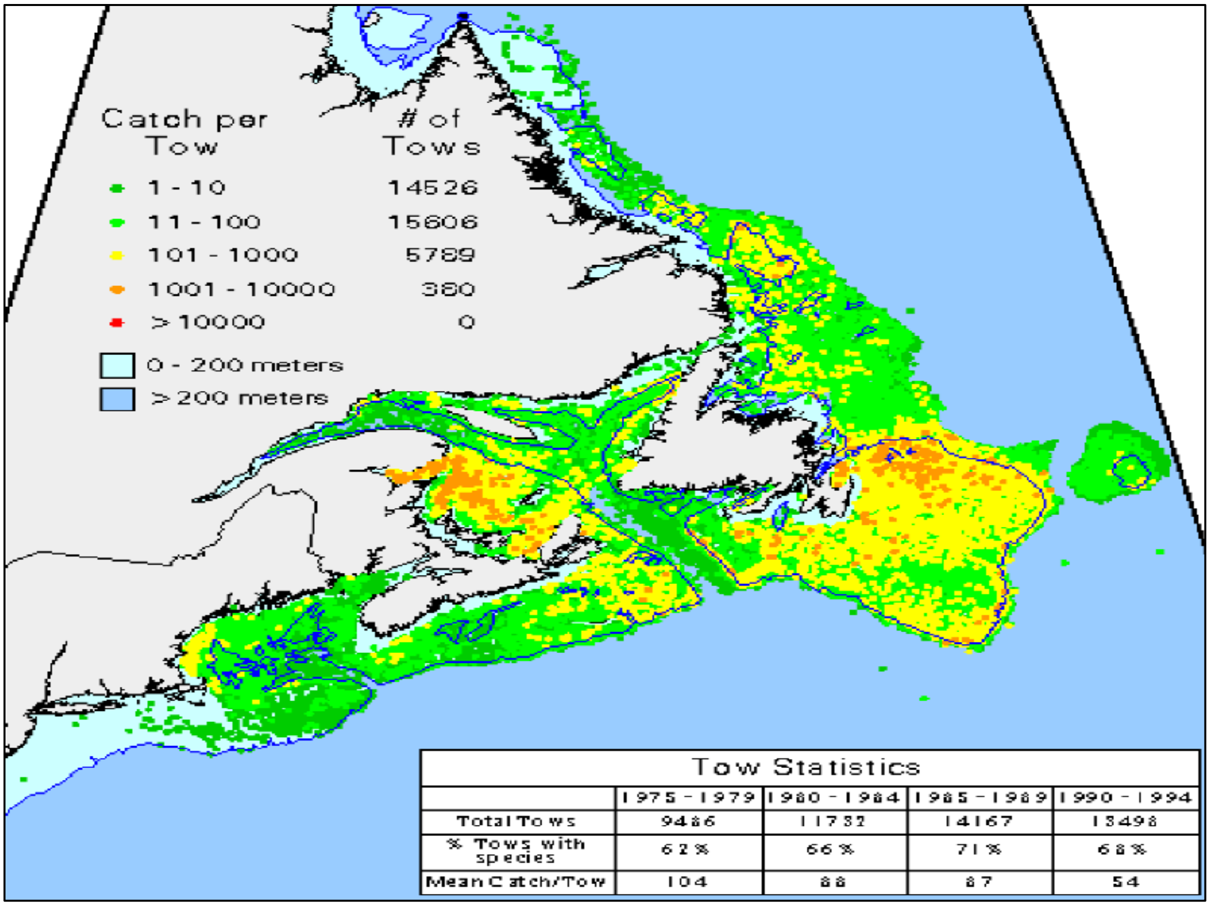
INTRODUCTION

Biology and Ecology

American plaice (*Hippoglossoides platessoides*; Fabricius 1780; commonly referred to as 'dab') is a cold-water demersal flatfish that is widely distributed across the northwest Atlantic Ocean (NEFMC 1985). They occur on the North American continental shelf from Labrador to the Mid-Atlantic Bight (Johnson et al. 1999, Salter 2018; Figure 1). They occupy sandy mud habitats in depths of 10-700 m, and most commonly occur in 50-100 m during spring and slightly deeper (100-180 m) during autumn (Bigelow and Schroeder 1953, Lange and Lux 1978, Johnson et al. 1999).

American plaice are relatively sedentary but migrate into relatively shallow habitats (<90 m) of the western Gulf of Maine and over southeastern Georges Bank to spawn from January to July, with peak spawning in April and May (Bigelow and Schroeder 1953, Smith et al. 1975, Smith 1985, NEFMC 1985, Johnson et al. 1999). Eggs are buoyant and incubate in 11-14 days in the upper water column to hatch as pelagic larvae (Bigelow and Schroeder 1953, Colton and Temple 1961, Johnson et al. 1999). Juveniles settle to benthic habitats and mature at approximately age-4 (median age at maturity for females in the Gulf of Maine was 3.6 years, O'Brien et al. 1993). Females grow to 77 cm and 4.4 kg, males grow to 65 cm and 3.25 kg, and maximum observed age is 34 years for females and 23 years for males in the Gulf of St. Lawrence (McBride et al. 2018). The oldest observed American plaice in US waters was 24 years (Cadrin 2021 Appendix A Working Paper 15).

American plaice burrow in sediment to escape predators and to ambush prey (Salter 2018). They are opportunistic feeders, feeding on mysids, amphipods, polychaetes, brittle stars, and mollusks as juveniles, then shifting to feed on fish, echinoderms, and bivalves as adults (Bigelow and Schroeder 1953, Johnson et al. 1999). Small plaice (<35 cm) are preyed on by cod (*Gadus morhua*), monkfish (*Lophius americanus*), and spiny dogfish (*Squalus acanthus*; Bigelow and Schroeder 1953, Johnson et al. 1999).



	1975 - 1979	1980 - 1984	1985 - 1989	1990 - 1994
Total Tows	9486	11732	14167	13498
% Tows with species	62%	66%	71%	68%
Mean Catch/Tow	104	88	87	54

Figure 1. Geographic distribution of American plaice based on trawl surveys (from Brown et al. 1996).

Stock Identification

The American plaice resource in US waters (NAFO areas 5-6) appears to be a single phenotypic stock that is separate from the Canadian resource, with some regional variation in growth rates. Survey distributions of juveniles and adults indicate a continuous distribution in the Gulf of Maine-Georges Bank region (Figure 1, Bowen 1987, Brown et al. 1996, Johnson et al. 1999, NEFMC 2016). Juveniles and adults are relatively sedentary, with seasonal movements to shallow spawning habitats in winter and deeper waters in summer (Lange and Lux 1978, Walsh 1994, Johnson et al. 1999, Salter 2018).

The spatial distribution and seasonality of eggs suggest separate spawning grounds in the Gulf of Maine, on Georges Bank and on Browns Bank (Bowen 1987, Johnson et al. 1999). Egg surveys suggest that spawning is about a month or two later on the Scotian Shelf than on

Georges Bank (O'Boyle et al. 1982), but other information suggests similar seasonality of peak spawning (Smith 1985). Eggs and larvae are transported by currents southwest along the New England coast, retained in the Gulf of Maine, transported to Georges Bank, or dispersed off Georges Bank, providing connectivity of early life states among US fishing grounds (Bigelow and Schroeder 1953, Colton and Temple 1961, Johnson et al. 1999). Larvae do not mix between Georges and Browns Banks (Bowen 1987). Walsh (1994b) tested for recruitment synchrony among adjacent management units and reported no correlation between recruitment in US waters and on the Scotian Shelf.

Growth is faster and maturity is earlier in US waters than in Canadian waters, suggesting limited post-larval mixing and persistent phenotypic differences among management units (Lux 1970, Lange and Lux 1978, Neilson and Hurley 1986, Walsh 1994a). Plaice from the western Gulf of Maine and Georges Bank grow faster than those in the Eastern Gulf of Maine and the Scotian Shelf (Esteves and Burnett, 1993). Significant growth differences have also been found between the Gulf of Maine and Georges Bank (Esteves and Burnett 1993, O'Brien et al. 1999), but geographic variation within US waters is much less than the differences between US. and Canadian waters. Analyses updated by the Research Track Working Group showed that geographic differences were not persistent, and differences were inconsistent among decades (Appendix A, Working Paper 3). The Working group also found no persistent geographic differences in length-weight relationships (Appendix A, Working Paper 3) or maturity (Appendix A, Working Paper 4).

The initial Fishery Management Plan recognized separate stocks of American plaice in the Gulf of Maine and on Georges Bank (NEFMC 1985), but it has been assessed and managed as a single US stock (NEFC 1986 and subsequent assessments). The limited larval connectivity between US and Canadian waters, persistent and substantial growth differences between US and Canadian waters, and considerable larval dispersal among US fishing grounds support the assumption of a single stock of American plaice in US waters, with occasional regional variation in growth between the Gulf of Maine and Georges Bank (NEFSC 1999a, 2001a). Therefore, this assessment is of the American plaice fishery and resource in US waters.

Fishery Description

The principal gear used to harvest American plaice is the bottom otter trawl, with relatively small catch from sink gillnets, bycatch in shrimp trawls, and negligible recreational catch (NEFC 1986, NEFSC 1992, Johnson et al. 1999). A mixed species flatfish fishery, primarily targeting winter and yellowtail flounders developed in the late 1880s by fishermen from Provincetown Massachusetts (the northern tip of Cape Cod, Figure 2) using beam trawls, with a switch to otter trawls in the early 1900s (Alexander et al. 1915). The development of otter trawling and the refinement of trawl doors in the 1920s revolutionized the New England groundfish fishery, flatfish markets expanded with the invention of freezers and filet machines, and American plaice were marketed as 'sole' (Jensen 1967).

American plaice was primarily caught as bycatch for most of the historical development of US fisheries, a directed fishery developed in the 1970s and early 1980s, and catches since then have been mostly from mixed-species groundfish trips. American plaice was mostly bycatch in Georges Bank fisheries before 1973, either discarded or landed for bait. Distant-water fleets caught plaice on Georges Bank in the 1960's and early 1970s. A seafood market expanded as a substitute for yellowtail and winter flounders, and a directed fishery developed in the 1970s in the Gulf of Maine (Bigelow and Schroeder 1953, Lange and Lux 1978, NEFMC 1985, NEFSC 1992, Johnson et al. 1999).

Annual landings increased from <2,000 mt during 1960-1963 and around 3,000 mt during 1964-1977 to 7,000 mt in 1977 (Lange and Lux 1978) and peaked at around 11,000 mt from 1980-1982, mostly from directed trips (NEFC 1986). Since then, most catch has been in mixed species groundfish trips (NEFC 1986).

Landed value of American plaice in the US gradually increased from $0.50/lb ($1,100/mt) in the mid-1960s to more than $2.00/lb ($4,400/mt) in the late 1980s and has averaged about $1.75/lb ($3,900/mt) since then (adjusted to 2015 values, Melnychuk et al. 2021). Landed value in the US was $5.3M in 2018 and $4.2M in 2019 (NMFS 2021).

Fisheries for American plaice in US waters were initially managed indirectly from management of multispecies trawl fisheries and regulated catch of other groundfish species. In

1958, the US Fish and Wildlife Service limited minimum trawl mesh to 4.5 in. (114 mm, Jensen 1967). The International Commission for the Northwest Atlantic Fisheries regulated minimum mesh sizes as well as minimum fish sizes, spawning closures and annual quotas for several other New England groundfish in the 1950s-1970s (Kulka 2012). The US Magnuson Act established an exclusive economic zone, excluded distant-water fisheries, and formed regional fishery management councils in 1976. A fishery management plan was developed for cod, haddock and yellowtail flounder by the New England Fishery Management Council in 1977 (Wang and Rosenberg 1997). An international boundary between US and Canadian waters was established in 1984 (Figure 2).

Since 1985, US fisheries for American plaice have been directly managed by the Multispecies Fishery Management Plan of the New England Fishery Management Council (NEFMC 1985; Appendix B). The initial management strategy was based on input controls (days at sea, size limits, gear restrictions, time/area closures) with substantial increases in minimum mesh sizes in 1982 (5 1/8 in., 130 mm), 1983 (5.5 in., 140 mm), 1994 (6 in., 152 mm), 1999 (6 in., 152 mm, diamond mesh or 6.5 in., 165 mm, square mesh) and 2000 (6.5 in., 165 mm, for all trawls). The western Gulf of Maine closure was implemented in 1997. Nordmore grates were required in the Gulf of Maine shrimp fishery in 1992 to decrease groundfish bycatch. The New England groundfish management system transitioned to an output control system (annual catch limits and catch shares) in 2010 (NEFMC 2009). When annual catch limits decreased in 2013 (Figure 3), American plaice was a ‘choke stock’ in the groundfish fishery (Henry et al. 2019).

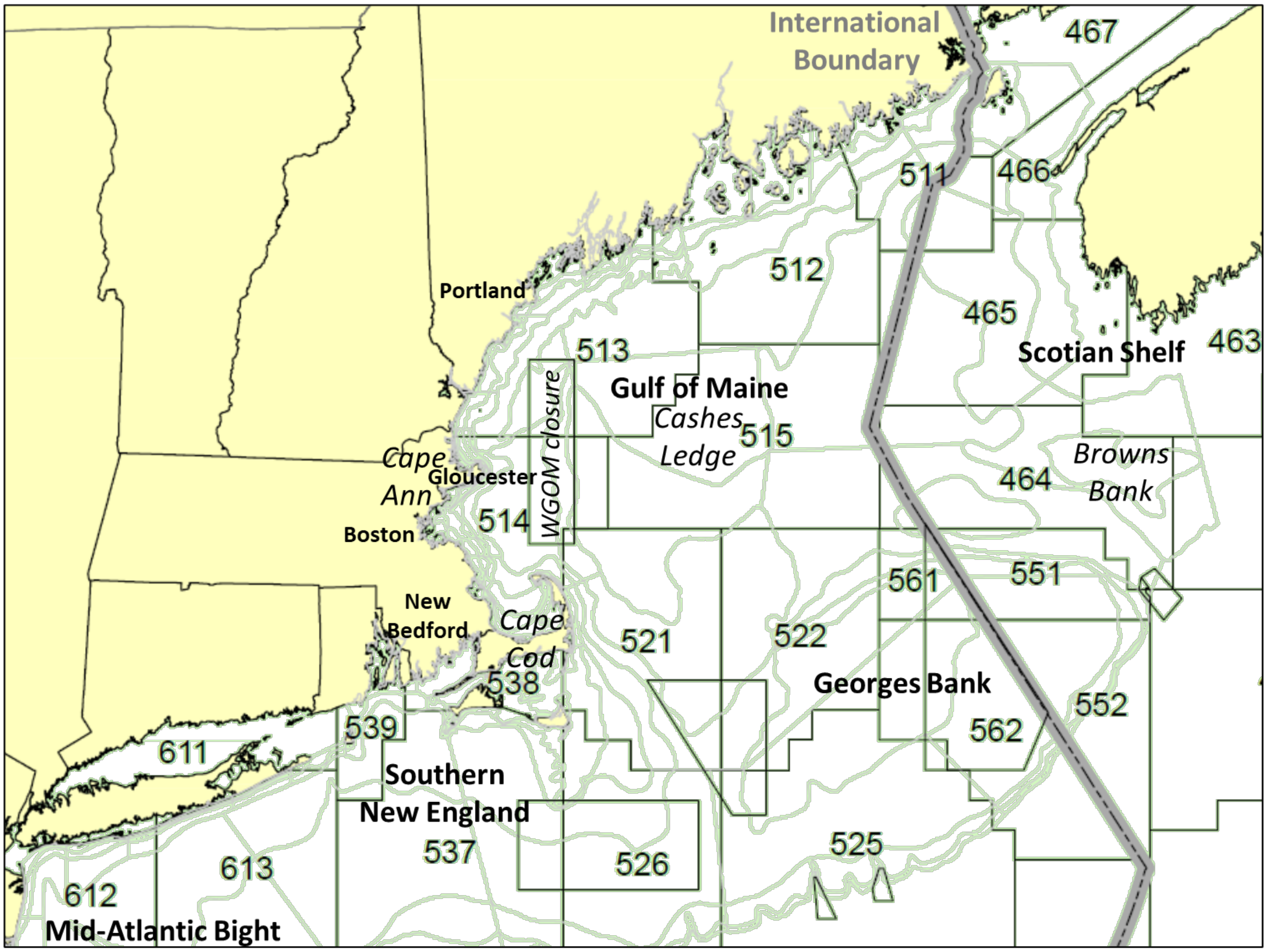


Figure 2. The northeast US continental Shelf with statistical reporting areas (dark green lines, numbered), survey strata (light green lines), the international boundary (bold grey line), geographic regions (bold text) and other locations referred to in text (italics).

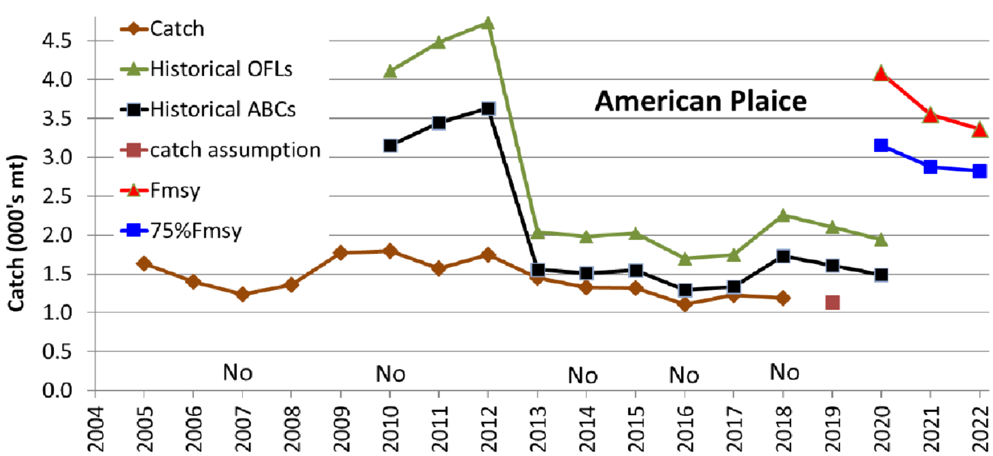


Figure 3. Time series of recent estimated catch of American plaice, overfishing limits (OFLs), Acceptable Biological Catches (ABCs) and projected catch at FMSY and 75%Fmsy (from Groundfish Plan Development Team 2019).

Previous Stock Assessments

American plaice in US waters have been assessed since the 1970s. The earliest stock assessments of American plaice in US waters were based on descriptive analysis of fishery landings, catch rates and survey indices (Lange and Lux 1978, NEFC 1986). The 1978 stock assessment reported increases in fishery catch rates and survey indices from 1973 to 1977 (Lange and Lux 1978), and the 1986 stock assessment concluded that the stock declined in the early 1980s (NEFC 1986).

The first analytical assessment was developed for the 14^{th} Northeast Stock Assessment workshop (NEFSC 1992). Estimates of landings, discarded catch from the large mesh trawl and shrimp trawl fleets, catch at age, and Northeast Fisheries Science Center (NEFSC) offshore survey abundance indices 1980-1991 were used for a calibrated Virtual Population Analysis (VPA), assuming a constant natural mortality rate (M) of 0.2 per year. Estimates of annual fishing mortality rate (F) on fully-recruited ages (ages 6+) decreased from about 0.7 during 1985- 1987 to about 0.5 during 1988-1991, which was greater than commonly used management reference points ($F_{0.1}$=0.17, F_{max}=0.28, $F_{20\%}$=0.49). Estimates of spawning stock biomass (SSB) declined from 48,000 mt in 1980 to 13,000 mt in 1991, and the 1987 year-class was relatively large. The VPA assessment was updated and revised over the next 20-year series through Stock Assessment Workshops, Groundfish Assessment Review Meetings, and operational or management track updates (Figure 4).

In 1998, management reference points were revised for all northeast US stocks to comply with the mandate to end overfishing and rebuild stocks (Applegate et al. 1998). The overfishing definition for American plaice was revised from $F_{20\%}$ (0.49, the fishing mortality expected to conserve 20% of maximum spawning potential) to $F_{0.1}$ (0.18, the fishing mortality that produces one tenth of the maximum increase in yield-per-recruit from F=0) as an interim approach until the stock assessment could be updated, because the stock-recruit series was too short to estimate F_{MSY} (the F expected to produce maximum sustainable yield). A SSB_{MSY} proxy was based on average recruitment and spawning potential at $F_{0.1}$.

The assessment reviewed by the 28th Northeast Stock Assessment Workshop included several revisions to the assessment (NEFSC 1999b, O'Brien et al. 1999). Catch-at-age for 1985-1993 was regionally stratified to account for growth differences between the Gulf of Maine and Georges Bank, and MADMF inshore surveys were included as additional abundance indices in the VPA calibration. The model had no residual patterns, and retrospective analyses indicated relatively consistent estimates with no patterns of inconsistency. The VPA estimated a substantial decrease in SSB from 49,000 mt in 1980 to 7,800 mt in 1989, and increased to 13,500 mt in 1997, which was 56% of the SSB_{MSY} proxy (24,200 mt). Estimated F (ages 5-8) peaked at 0.79 in 1995 and was 0.47 in 1997, greater than estimated reference points ($F_{0.1}$=0.19, F_{max}= =0.35, and $F_{20\%}$=0.40). There was no apparent stock-recruit relationship, so the MSY proxy was based on $F_{0.1}$ (NEFSC 1999b, O'Brien et al. 1999).

The stock assessment was updated for the 32nd Northeast Stock Assessment Workshop (NEFSC 2001b, O'Brien and Esteves 2001). Estimates of SSB decreased from 47,000 mt in 1980 to 7,500 mt in 1989 then increased and stabilized at 14,000 mt, which was 58% of the SSB_{MSY} proxy (24,200 mt). Estimates of F (ages 5-8) peaked at 0.64 in 1994-1995 and decreased to 0.26 in 1999, which was greater than estimated reference points ($F_{0.1}$=0.19, F_{max}=0.35). Retrospective analyses indicated relatively consistent estimates of SSB and F, but there was a pattern of recruitment estimates being revised upward as the assessment was updated (NEFSC 2001b, O'Brien and Esteves 2001).

Management reference points were re-evaluated for all New England groundfish in 2002, and MSY proxies for American plaice were based on the $F_{40\%}$ (0.17) because the 1980-1999 stock recruit relationship was negative (i.e., decreasing recruitment with increasing SSB; NEFSC 2002b). The VPA was updated by the initial Groundfish Assessment Review Meeting (NEFSC 2002a). The estimate of 2001 SSB was 13,800 mt (48% of the SSB_{MSY} proxy, 28,600 mt), and 2001 F (ages 5-8) was estimated to be 0.43. The retrospective pattern in recruitment estimates continued, but there were no retrospective patterns in estimates of SSB and F (NEFSC 2002a).

The assessment was updated in 2005 (Mayo and Terceiro 2005). The estimate of 2004 SSB was 14,100 mt (49% of the SSB_{MSY} proxy, 28,600 mt), and the estimate of 2004 F was 0.15

(90% of the $F_{40\%}$, 0.17). Retrospective analysis indicated a weak pattern of F estimated revised downward, and SSB estimates revised upward (Mayo and Terceiro 2005).

The stock assessment was revised in 2008 to reduce the emerging retrospective pattern by expanding the age structure from ages 1-9+ to ages 1-11+, with aggregate age-9-11+ survey indices (NEFSC 2008). The revised VPA had a moderate retrospective pattern, with Mohn's (1999) rho statistics of -0.31 for F, 0.43 for SSB and 0.6 for recruitment, and stock status and catch advice was based on rho-adjusted estimates. The adjusted estimate of 2007 SSB was 11,100 mt (50% of the SSB_{MSY} proxy, 22,200 mt), and the adjusted estimate of 2007 F (ages 6-9) was 0.09 (47% of F40%, 0.19).

The assessment was revised in 2012 with revised estimates of landings and discards (large mesh trawl, shrimp trawl, small mesh trawl, gillnet and scallop fleets), time-varying maturity, and an updated version of the VPA calibration software (NEFSC 2012). There was a retrospective pattern in estimates of SSB (rho=0.62), F (rho = -0.35) and recruits (rho=1.24). Retrospective-adjusted estimate of 2010 SSB was 10,800 mt (59% of the SSB_{MSY} proxy, 18,400 mt), and 2010 F was 0.13 (72% of F40%, 0.19).

The stock assessment was updated in 2015 (NEFSC 2015). The retrospective pattern decreased for estimates of SSB (rho=0.33) and F (rho = -0.32). Retrospective adjusted estimate of 2014 SSB were 11,000 mt (84% of the SSB_{MSY} proxy, 13,107 mt) and 2014 F was 0.12 (59% of $F_{40\%}$, 0.20).

The assessment was updated in 2017 (NEFSC 2017b). The retrospective pattern continued for estimates of SSB (rho=0.35) and F (rho = -0.33). Retrospective adjusted estimate of 2016 SSB were 13,400 mt (99% of the SSB_{MSY} proxy, 13,500 mt) and 2016 F was 0.11 (51% of $F_{40\%}$, 0.22). Information on NEFSC survey efficiency and catchability and derived estimates of area-swept biomass were considered for comparisons to VPA estimates (NEFSC 2017b).

The most recent stock assessment revised the assessment by excluding Massachusetts inshore surveys from the VPA calibration to improve model diagnostics and agreement with survey biomass estimates and to account for an apparent shift to deeper habitats (NEFSC 2019). The retrospective pattern improved for estimates of SSB (rho=0.27) and F (rho = -0.20).

Retrospective adjusted estimate of 2018 SSB were 17,700 mt (116% of the SSB_{MSY} proxy, 15,300 mt) and 2018 F was 0.09 (34% of $F_{40\%}$, 0.26).

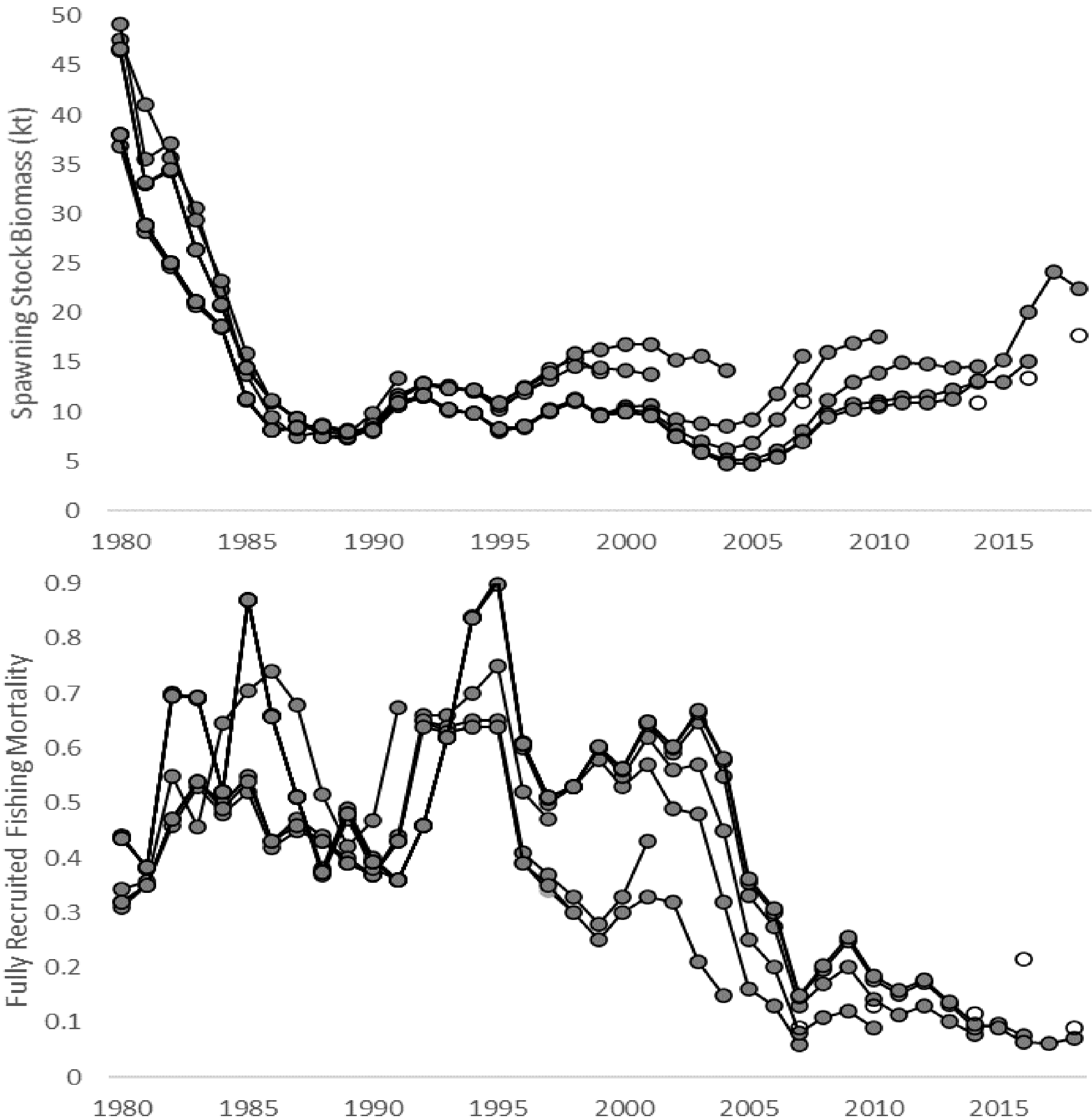


Figure 4. Estimates of spawning stock biomass (top) and fully-recruited fishing mortality (bottom) from the series of American plaice stock assessments. Open circles indicate retrospective-adjusted estimates.

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*."

Contributors:

Jamie Behan, Lisa Kerr, Amanda Hart, Alex Hansell, Tyler Paklovitch, and Steve Cadrin

A review of the scientific literature was undertaken to characterize existing research that has identified ecosystem and climate influences on the American plaice stock dynamics (Behan et al. 2021, Working Paper 14 Appendix A). The aim of this review was to characterize the state of knowledge on this topic such that these relationships could be considered, as appropriate, in addressing other TORs in the research track stock assessment process. Due to the limited literature currently available for American plaice, in addition to the primary study areas of interest (Gulf of Maine and Georges Bank), this review also includes literature from the Scotian Shelf, Grand Bank, and Gulf of St. Lawrence regions.

To help identify ecosystem and climate influences on American plaice recruitment, distribution, and growth, generalized additive models (GAMs) were used to examine associations between American plaice population dynamics and environmental variables in the Gulf of Maine region including sea surface and bottom temperature anomalies, the Atlantic Multidecadal Oscillation (AMO), the North Atlantic Oscillation (NAO), and the Gulf Stream Index (GSI; Behan and Kerr 2022, Working Paper 16 Appendix A). Ecosystem influences were considered in ToR2 by recognizing decadal shifts in growth and standardizing fishery catch rates, in ToR3 by investigating factors of survey catchability as well as integrating inshore and offshore surveys, in ToR4 by recognizing decadal shifts in growth and exploring environmental covariates of recruitment and survey catchability, in ToRs5-6 by considering evidence for alternative stanzas of recruitment to inform biological reference points and projections, and in ToR7 by proposing additional research recommendations.

Background

Along the Atlantic coast of North America, the range of American plaice extends from southern Labrador to the Mid-Atlantic Bight (Johnson et al. 1999, Salter 2018) with the majority of biomass in US waters in the Gulf of Maine and Georges Bank regions. Ocean properties in the Gulf of Maine are highly influenced by the Labrador current, which flows in a general northeast to southwest direction into the Gulf of Maine region (Nye et al. 2010; Brickman et al. 2021). More specifically, sources of water in the Gulf of Maine include the relatively cooler, fresher, surface Scotian Shelf Water which enters via the Nova Scotian Shelf, and warmer, saltier Slope Water that enters the Gulf of Maine at intermediate and greater depths through the Northeast Channel (Townsend et al. 2015; Pettigrew et al. 2005). The flow of water through the Gulf of Maine is cyclonic, or counter-clockwise (Townsend et al. 2015; Chang et al. 2016), where the flow is driven in part by density differences between the Scotian Shelf water and the Slope Water (Brooks 1985). The strength of the Labrador current is tied to both seasonality and temperature differences such that when the Labrador current is strong, water temperatures in the Gulf of Maine decrease and when the strength of Labrador current is weak, a higher proportion of warm water coming from the Gulf Stream enters the Gulf of Maine, increasing Gulf of Maine water temperatures (Nye et al. 2010). The position of the Gulf Stream has shifted northward (Nye et al. 2010; Pershing et al. 2015; Brickman et al. 2021) and the Gulf of Maine has seen temperature warming rates much higher than that of the global average in recent decades (Pershing et al. 2015). While some attribute the warmed temperatures to this northward shift in the Gulf Stream (Nye et al. 2010; Saba et al. 2016), others attribute this to the strength of the Gulf Stream (Alexander et al. 2019). Temperature is considered to be one of the most influential environmental factors for many species across the Northwest Atlantic and has been associated with effects on growth, maturity, survival, metabolism, recruitment, and other aspects of life history (Nye et al. 2010; Runge et al. 2010; Lesser 2016). Many species along the Northwest Atlantic coast have thermal preferences between 5 and 15°C, and thermal habitat within this range has decreased (Nye et al. 2010) or is expected to decrease in the future (Kleisner 2017).

Nutrient profiles in the Gulf of Maine have also changed in the past few decades. Deep (>100m) waters in the Gulf of Maine have become both fresher and colder with increasing silicate and lowering nitrate levels (Townsend et al. 2010, 2015). Although changes in the predominating form of slope water that enters the Gulf of Maine (Labrador Slope Water or Warm Slope Water) are correlated with the North Atlantic Oscillation (NAO, Smith et al. 2001; Greene and Pershing 2003; Thomas et al. 2003; Townsend et al. 2006), Townsend et al. (2010) found that the increasing silicate levels of the Gulf of Maine were not correlated with the current NAO cycle, but rather better coincided with the melting of the Arctic ice sheet. Changes in nutrient concentrations are important driving forces in determining the species composition of phytoplankton in an ecosystem (Townsend et al. 2010). Even though the Gulf of Maine's food web complexity has increased from the mid-2000s, indicating a more stable ecosystem than prior decades (Han et al. 2021), changes in the composition of species that serve as the foundation of this food web, such as the decline in the copepod Calanus finmarchicus in the GOM, may also result in significant impacts on higher trophic levels or changes in primary productivity (Nye et al. 2010; Pershing et al. 2021).

Other observed and expected changes in the Gulf of Maine region include increases in precipitation (Wake et al. 2006), acidification (Nye et al. 2010), sea level rise (Rhamstorf 2007), species invasions, and/or dominance by more "warm water" species (Harris and Tyrrell 2000; Nye et al. 2010; Kleisner et al. 2017). The forecasted increase of warm water species in the Gulf of Maine is expected to decrease suitable habitat (Kleisner et al. 2017) and increase natural mortality (Jorgensen and Holt 2012) of species in the Gulf of Maine.

Although ecosystem profiles are transforming along the Northeast U.S Shelf, the temperatures at which many species are still observed have not significantly changed overtime (Nye et al. 2010). This suggests that species are shifting their distributions to seek out their ideal habitat conditions. Shifts in spatial distributions are hypothesized to be one of the first responses to unfavorable changes in species' environments (Walther et al. 2002; Parmesan and Yohe 2003). Thus, it is important to identify how various ecosystem and climate influences are affecting distributions, stock dynamics, and other life history characteristics of species in the Gulf of Maine, so that informed conservation or management decisions can be implemented.

Ecosystem Influences on Distribution and Habitat Use

Temperature and depth are important drivers of American plaice distribution. High plaice densities have been correlated with depths between 60-90 m and mean temperatures <2 °C in the Gulf of St. Lawrence (i.e., Magdalen Shallows, Swain et al. 1998). A spatial analysis of plaice catch from 2002-2010 surveys in the Gulf of St. Lawrence revealed high densities of plaice were associated with areas of moderate depth (66.1 m), and relatively lower salinity (<32) and temperature (max monthly temps between 1-3 °C, and annual temperatures 0.5-1.5 °C at mean depth), whereas low densities were associated with highly saline and warm, deep channel locations (Chouinard et al. 2014). However, upon analysis of plaice catch from a longer time series (1971-2010), warm, deep channel locations were not associated with low densities (Chouinard et al. 2014). In relation to seasonality, American plaice in the Gulf of St. Lawrence occupied deeper and warmer water in the winter (374-426 m and 5.2-5.4 °C) compared to summer (58-67 m and -0.1 to 0.3 °C; Swain et al. 1998). This seasonal migration pattern has also been observed in other coastal Canadian waters (Powles 1965), as well as in the Gulf of Maine and Georges Bank regions (Methratta and Link 2007). In the Gulf of Maine, similar seasonal depth migration patterns can be seen in Figure 1.1.

Size and life stage are also important factors determining habitat use. Methratta and Link (2007) found that smaller American plaice were more abundant in shallow habitats whereas large plaice were most abundant in the deep portions of the Gulf of Maine. In other studies, depth played a major role in determining spatial patterns and abundance of juvenile plaice in the Grand Bank (Walsh et al. 2004). Juvenile plaice were found to be abundant on sand/shell hash sediments, less abundant on muddy sand & rock/sand sediments, and almost absent from mud sediments (Walsh et al. 2004). Adult and juvenile American plaice showed distribution overlap which can lead to possible spatial competition between these two age groups (Walsh et al. 2004).

Several studies have explored development of species distribution models for American plaice. Swain et al. (1998) found that habitat associations in the Gulf of St. Lawrence were not consistently stronger when comparing depth and temperature variables, but using both depth and temperature in models improved plaice density prediction capabilities. Nye et al. (2009) found the relationship between area occupied and abundance was not strong. Adding location improved fit of generalized additive models and better described influences of environmental variables on distribution (Walsh et al. 2004). In that study, temperature, depth, and sediment were the top predictor variables for American plaice and yellowtail flounder distribution (Walsh et al. 2004).

Hare et al. (2016) categorized American plaice as having 'high' sensitivity to climate change, and high associated potential for climate-related distribution change. From 1985-2018 an influx of cold water into the Grand Bank resulted in a 200 km southward shift in the distribution of American plaice, resulting in a contraction of distribution, and declines in biomass (Robertson et al. 2021). Nye et al. (2009) found that American plaice in the Northeast US waters shifted to deeper water at a rate of 0.53 m/year from 1968 to 2007 associated with changing ocean conditions (e.g., the North Atlantic Oscillation and Atlantic Multidecadal Oscillation). However, Fredston-Hermann et al. (2020) found sea surface temperature to be a better predictor of range edge positions for many demersal species, including American plaice, than sea bottom temperature. Through the use of a high-resolution global climate model, Kleisner et al. (2017) estimated an increase in surface and bottom temperatures by 3.7 ℃ and 3.9 ℃, respectively, over the 80-year forecast period in the Gulf of Maine. This projected warming is expected to result in a loss in suitable thermal habitat for American plaice in the US Northeast Continental Shelf marine ecosystem (Kleisner et al. 2017).

Exploratory analyses by the Working Group showed substantial changes in mean latitude of catch for both the spring and fall NEFSC trawl surveys, and changes in mean depth of catch for the fall survey (Figure 1.2). Mean latitude of survey catches shifted southerly during the 1980s and 1990s then northerly in the last two decades. Mean depth of catch in the fall survey decreased in the 1980s and has gradually shifted to deeper water since the late 1980s. Mean latitude was significantly related to estimates of spawning stock biomass and temperature indicators, with the greatest amount of variance explained by bottom temperature. Mean depth of catch in the fall survey was positively related to bottom temperature and spawning biomass and negatively related to NAO (Figure 1.3).

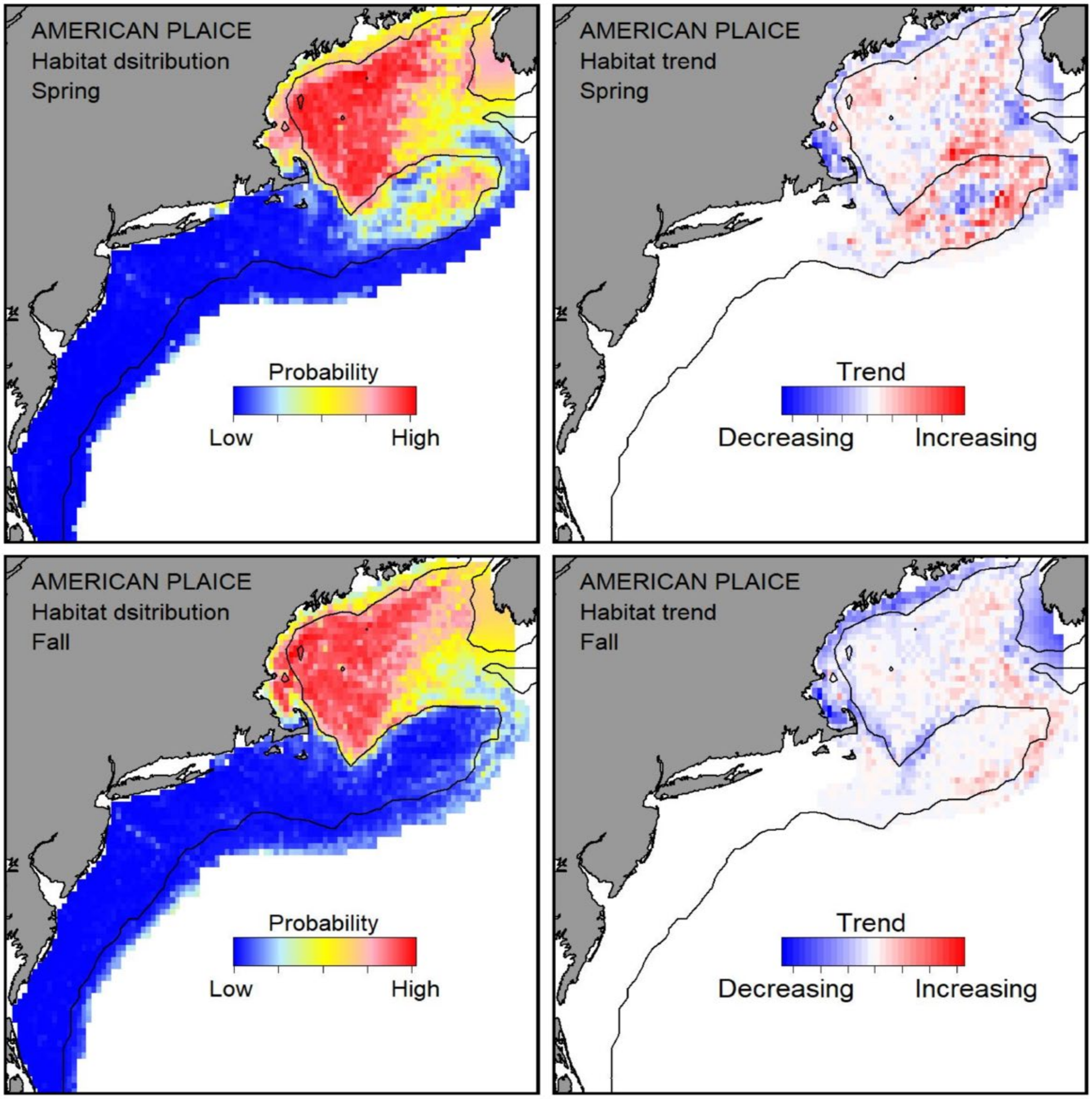


Figure 1.1. Geographic distribution of American plaice in the Gulf of Maine region. Top row: Spring habitat probability and trends. Bottom row: Fall habitat probability and trends. Modeled using machine learning techniques (from NOAA Fisheries 2021).

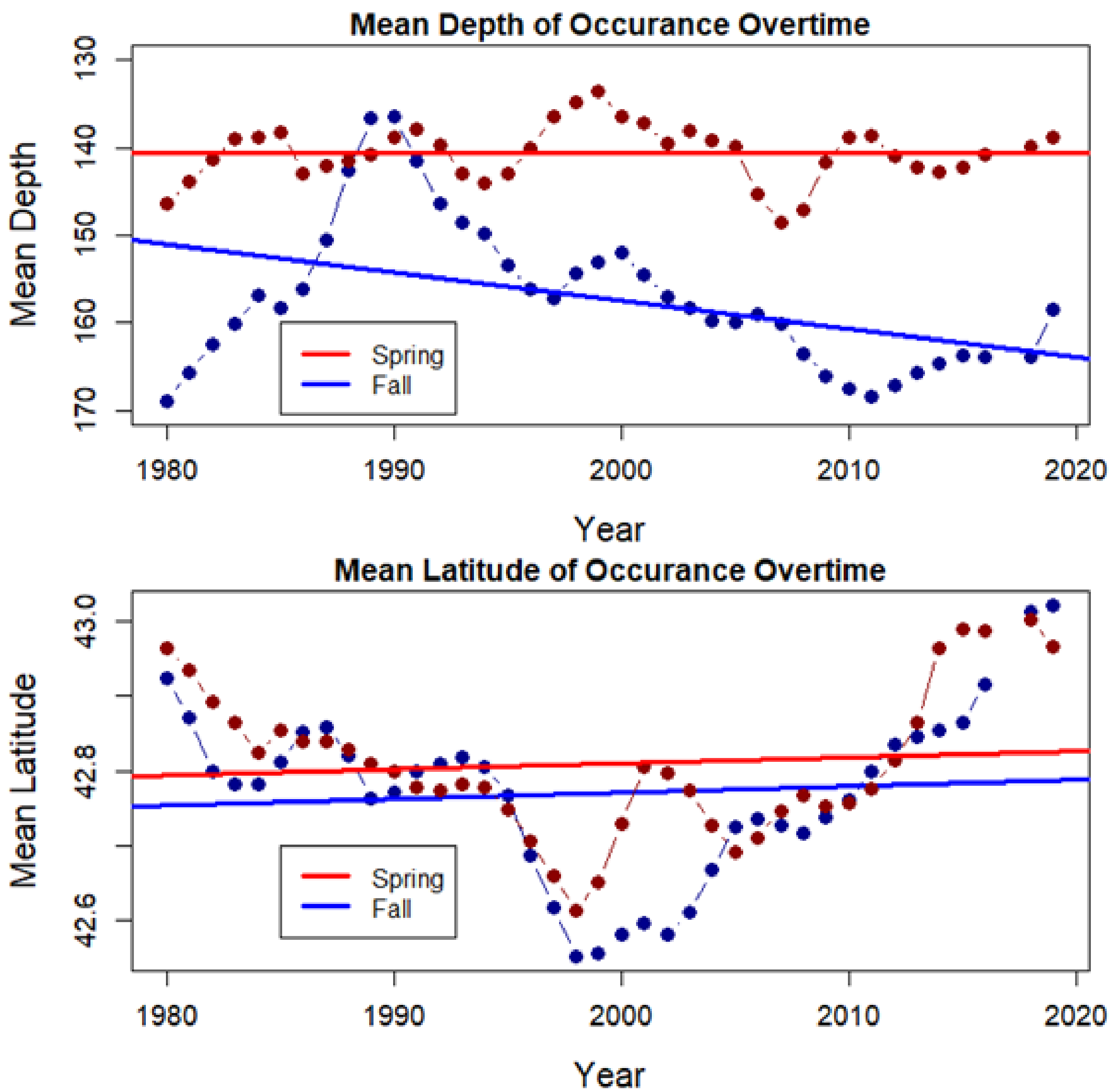


Figure 1.2. Mean Depth and latitude of catches from NEFSC spring and fall bottom trawl surveys.

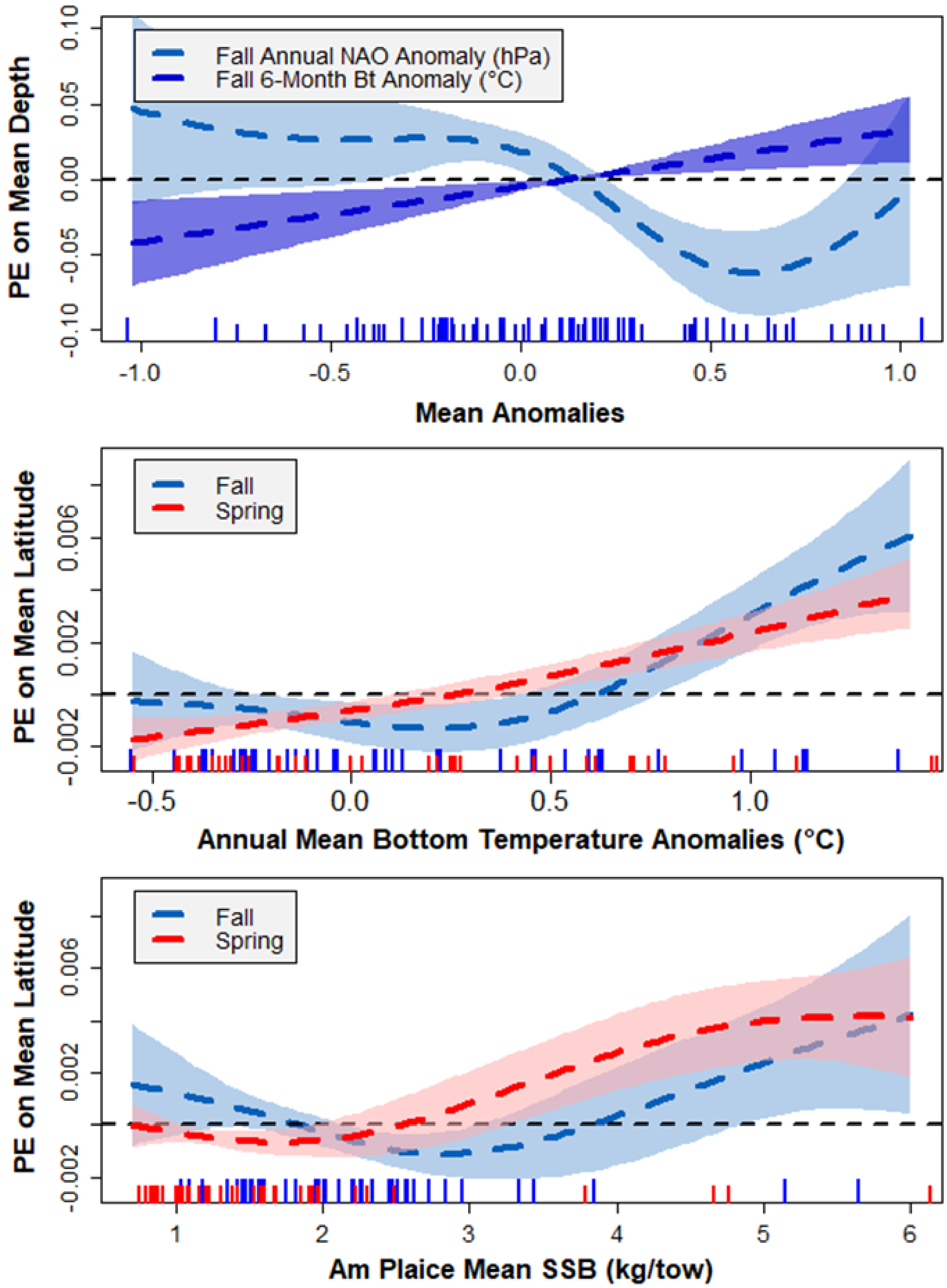


Figure 1.3. Environmental effects on mean latitude and mean depth of survey catches.

Ecosystem Influences on Recruitment

American plaice eggs are buoyant and incubate for 11-14 days in the upper water column, hatching as pelagic larvae (Bigelow and Schroeder 1953, Colton and Temple 1961, Johnson et al. 1999). Thus, sea surface temperatures are an important consideration for plaice recruitment. The thermal limit for survival and incubation of American plaice eggs is 14°C (Howell and Caldwell 1984), and recent Gulf of Maine sea surface temperatures in late spring-early summer (the end of the spawning season) have exceeded that threshold in some areas (e.g., Figure 1.4). Brodziak and O'Brien (2005) identified a significant relationship between reproductive rate (recruits per spawner) of American plaice and regional temperature anomalies in the Gulf of Maine, as well as with the North Atlantic Oscillation, with the greatest recruitment rates occurring at extreme cold temperatures (Brodziak and O'Brien 2005, Figure 1.5). Other variables such as salinity were not found to be a significant driver of plaice recruitment (Walsh et al. 2004).

Adult plaice spatial structure influences the proportion of nursery area utilized by juvenile plaice. Walsh et al. (2004) noted that an increase in adult American plaice stock size typically correlated with a wider distribution of juveniles which resulted in exposure to a broader range of environmental conditions. In addition to adult distribution influencing juvenile survival, larval distributions and the influence of sea surface temperature on survival may impact future adult population distributions (Fredston-Hermann et al. 2020). Nye et al. (2009) noted that potential changes in currents due to climate impacts may lead to transportation of larvae to suboptimal nursery habitats (Nye et al. 2009).

Past examinations of stock-recruit relationships for American plaice have noted weak relationships with stock size and identified temporal variation in recruitment (i.e. non-stationarity). One study found constant recruitment, the geometric mean of recruitment over the range of the stock size, more appropriate to describe the stock-recruitment data than one in which recruitment was proportional to stock size for 8 different flatfish stocks, including American plaice in the Gulf of Maine and Georges Bank (Iles 1994). A major caveat of a model utilizing constant recruitment however, is that it assumes positive recruitment at a stock size of zero, which is biologically implausible (Iles 1994). Another study found models that utilized autocorrelated stock recruitment relationships (SRRs) were more similar to observed patterns, however autocorrelated SRRs also tended to demonstrate increased risk of overfishing (Zhang et al. 2020). In addition, the SRRs in Zhang et al.'s (2015) study also demonstrated strong temporal variation that is likely associated with both environmental and anthropogenic effects, which suggests non-stationarity in recruitment dynamics. Exploratory analyses by the Working Group showed that survey indices of recruitment rate (R/SSB; Figure 1.6) were significantly related to the AMO (Figure 1.7).

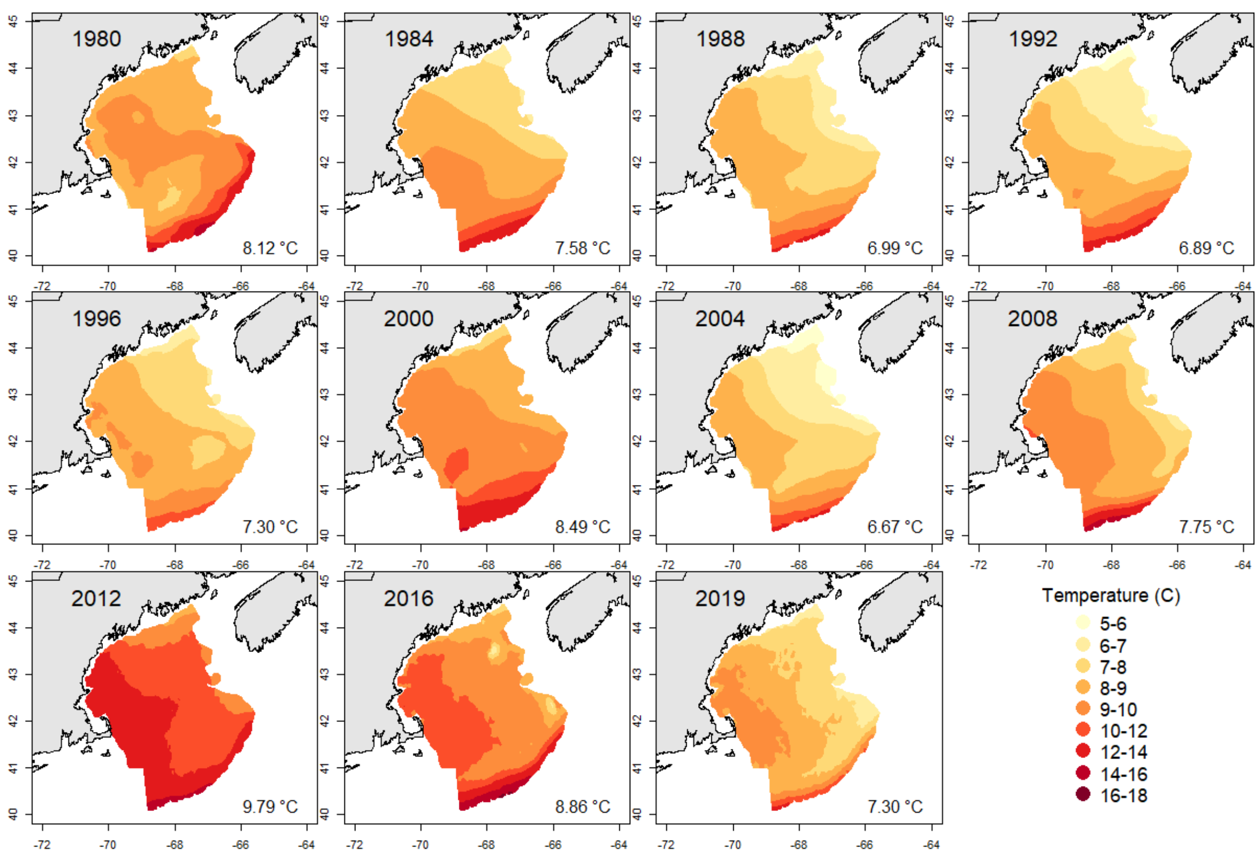


Figure 1.4. Spatial representation of 4-month mean sea surface temperatures (March - June) across the Northeast Fisheries Science Center (NEFSC) bottom trawl survey, with restricted strata from 1980-2019 (offshore strata 13-30, 36-40). Data were sourced from the Finite Volume Ocean Community Model (FVCOM). The corresponding year for each plot panel is listed in the upper left corner and the average sea surface temperature across the spatial area is listed in the bottom right corner of each panel. This 4-month period represents the period of the year where plaice larvae are most likely to be at or near the sea surface, before they descend towards the sea floor (Huntsman 1918; Johnson 2004).

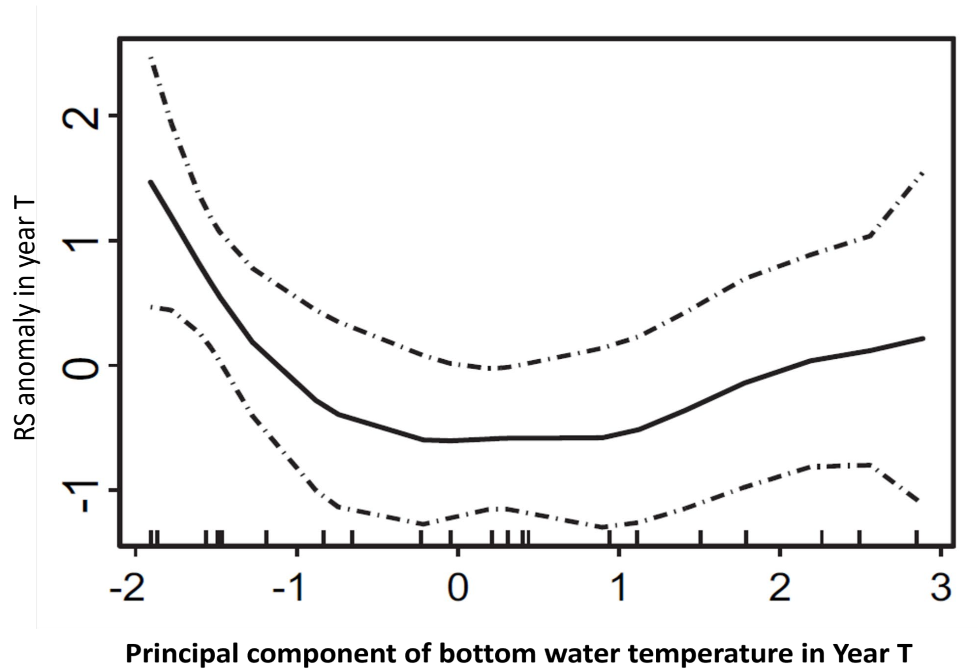


Figure 1.5. Estimated effects of the first principal component of water temperature anomalies on recruits per spawner anomalies of American plaice (from Brodziak and O'Brien 2005).

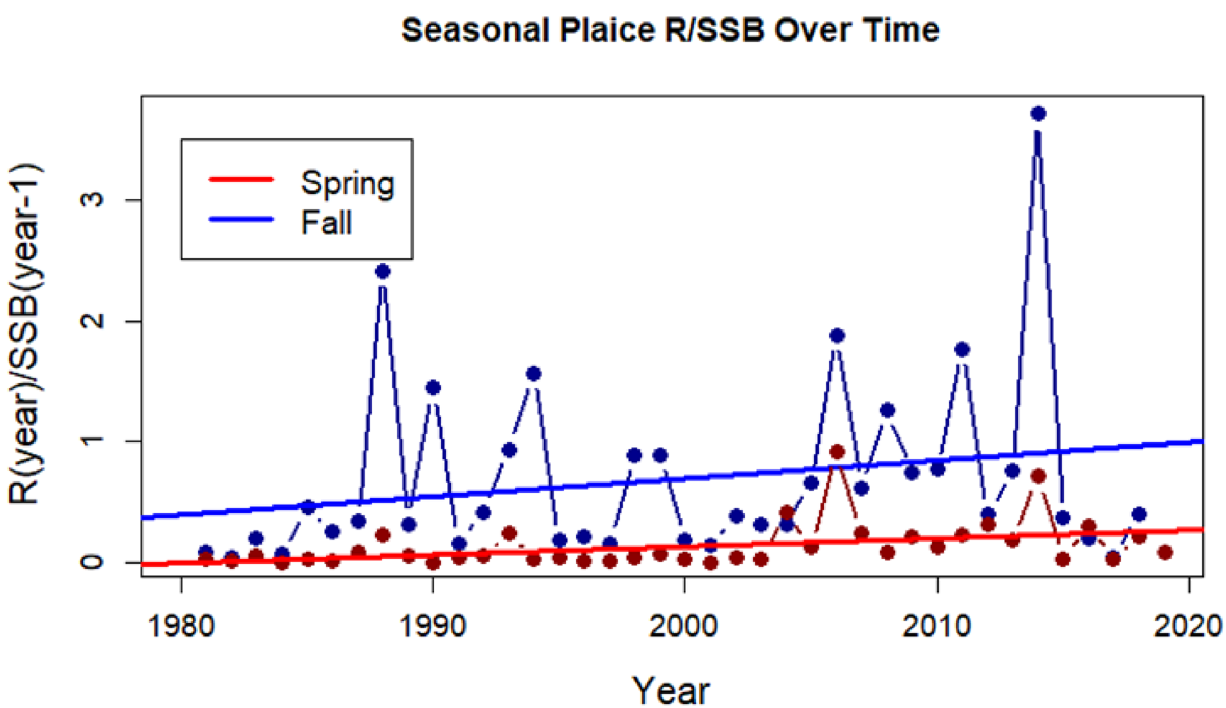


Figure 1.6. Survey indices of recruitment rate (R/SSB).

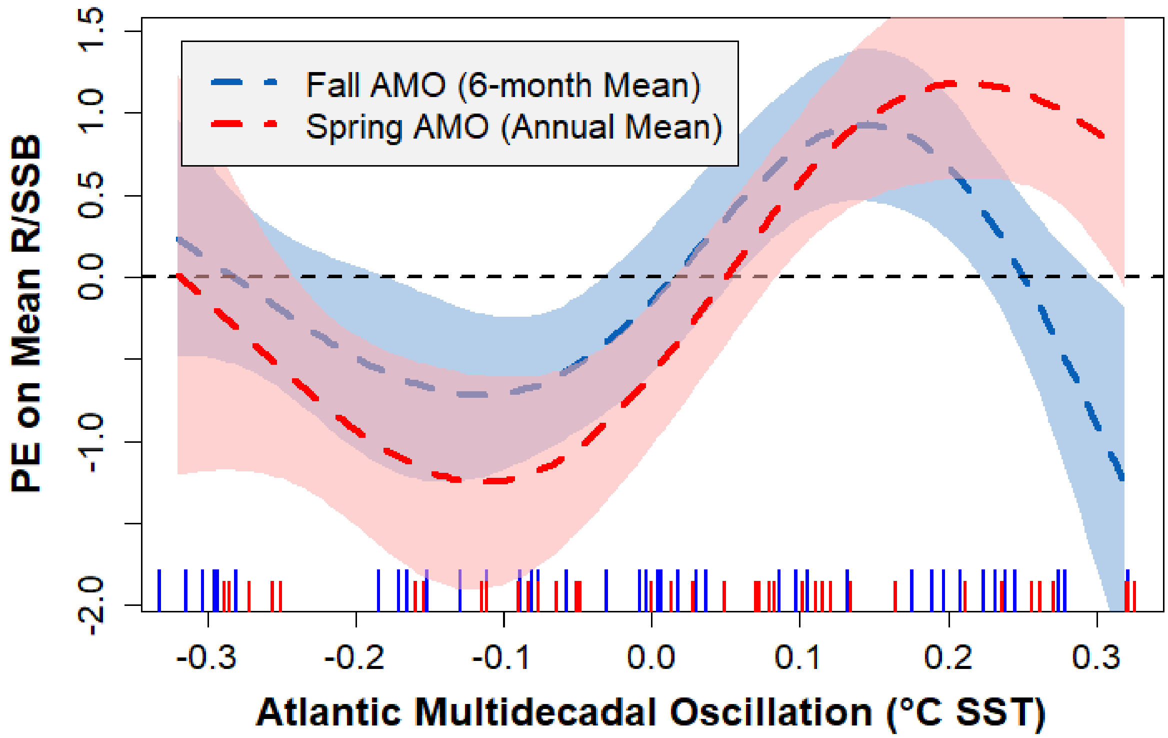


Figure 1.7. AMO effects on recruitment rate (R/SSB).

Ecosystem Influences on Growth and Maturity

Warmer temperatures have been associated with accelerated growth rates, reductions in body size, and earlier ages at maturity in marine fish (Levangie et al. 2021). Powles (1965) found that slower growth rates in plaice typically occur in deeper waters in the Gulf of St. Lawrence which is thought to be associated with colder temperatures compared to shallow water habitat. Conversely, increased growth rates have been correlated with areas of higher temperatures in Canadian waters (Swain and Morgan 2001; Salter et al. 2018). In the northwest Atlantic, Johnson et al. (1999) reported that the fastest growth rates occur in the Gulf of Maine. However, a more recent study found plaice on Georges Bank have faster growth rates than plaice in the Gulf of Maine (Salter et al. 2018).

In the Gulf of St. Lawrence, female plaice have been found in warmer waters than male plaice at certain times of year and have also been determined to have a higher growth rate than males (Swain et al. 1998). This could be a trade-off between foraging rate and predation risk. Females may risk a higher foraging rate than males to obtain more food and prefer warmer temperatures (Swain et al. 1998). An examination of age at maturity of American plaice in the Grand Bank demonstrated an overall pattern of decline from the 1960s to mid-2010s. From 1960-1970, age at 50% (A50) maturity for female American plaice in the Grand Bank was typically >9, but from 1970-1990, estimates demonstrated declines in A50 throughout the area (Zheng et al. 2020). After 1990, model estimates from Zheng et al. (2020) show a slight increasing trend in female age at 50% maturity, but did not return to historical high values. Zheng et al. (2020) also noted that the size at maturation of American plaice was larger in the southern part of the Grand Bank than the northern part between 1978-2005 (Zheng et al. 2020). Although latitudinal patterns were not significant throughout the time series of this study, temperature was thought to play a role in driving maturity indirectly through its effects on metabolism and growth (Zheng 2020).

Fishing pressure can directly impact and interact with environmental conditions to influence growth and maturation patterns. While increased fishing pressure has been thought to promote growth and earlier maturation due to reduction in resource competition, decreases in age and length at maturity coincided with timing of prolonged intensive fishing from 1970s to early 1990s in the Grand Bank (Zheng et al. 2020). Similarly, growth rates may also be affected by population abundance, as Rijnsdorp and Van Leeuwen (1992, 1996) found decreased growth rates in plaice when abundance was high. However, another study found that fishing pressure did not affect the rate of growth in plaice from 1957-1961 (Powles 1965).

Exploratory analyses by the Working Group showed that weight at all ages (Figure 1.8) was significantly related to AMO (except age 1), weight at younger ages were significantly related to GSI (ages 2-4 for fall survey and ages 2-6 for spring survey) and spawning stock biomass (ages 3-5 in fall and 5-6 in spring), and weight at some ages were significantly related to bottom temperature (ages 1 and 9 in fall, ages 10 and 11 in spring). Response curves showed consistency between similar significant variables (Figure 1.9). Mean condition factor from the fall survey (Figure 1.10) was significantly related to spawning stock biomass, bottom temperature, and AMO (Figure 1.11).

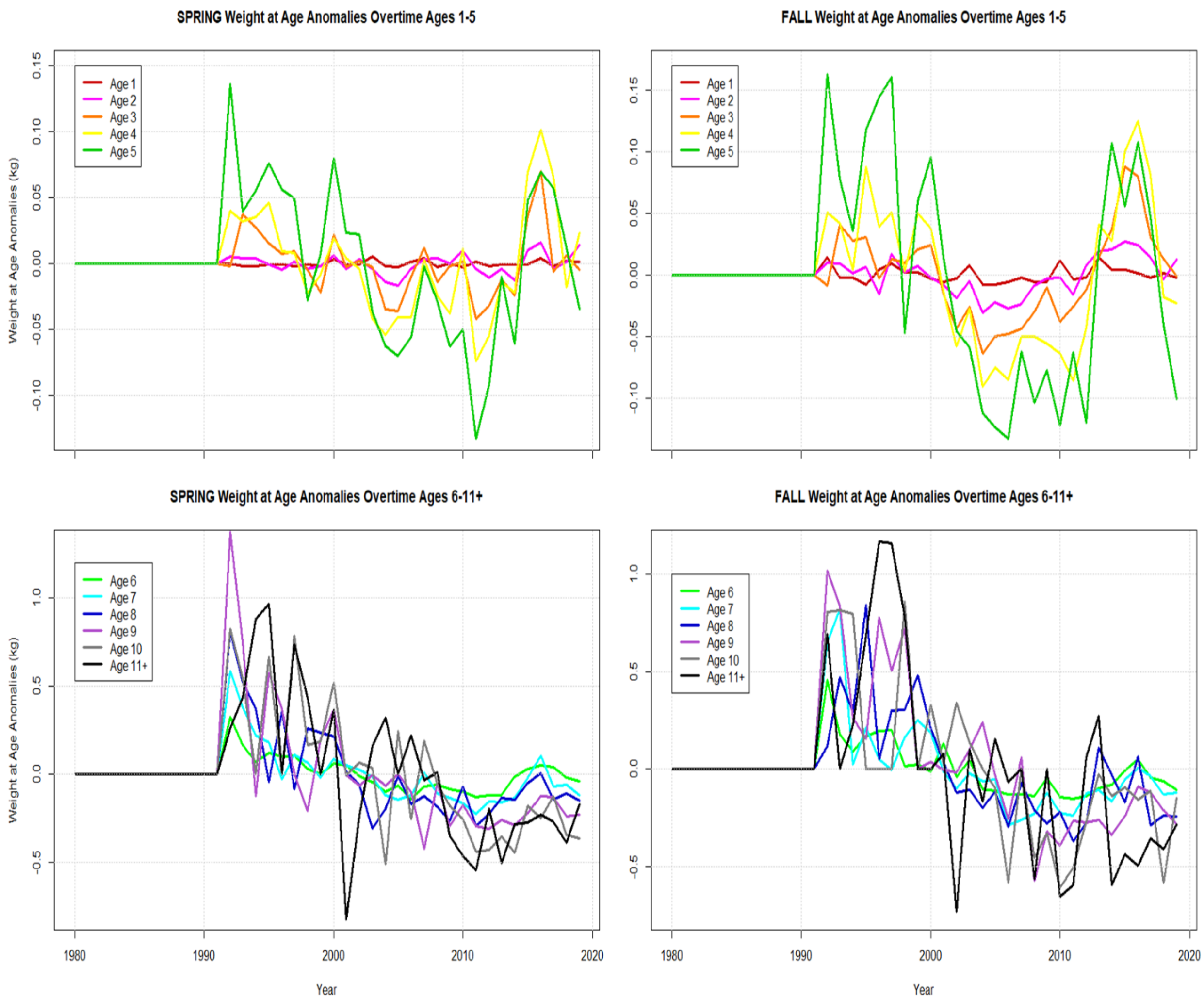


Figure 1.8. Weight at age anomalies from spring and fall surveys.

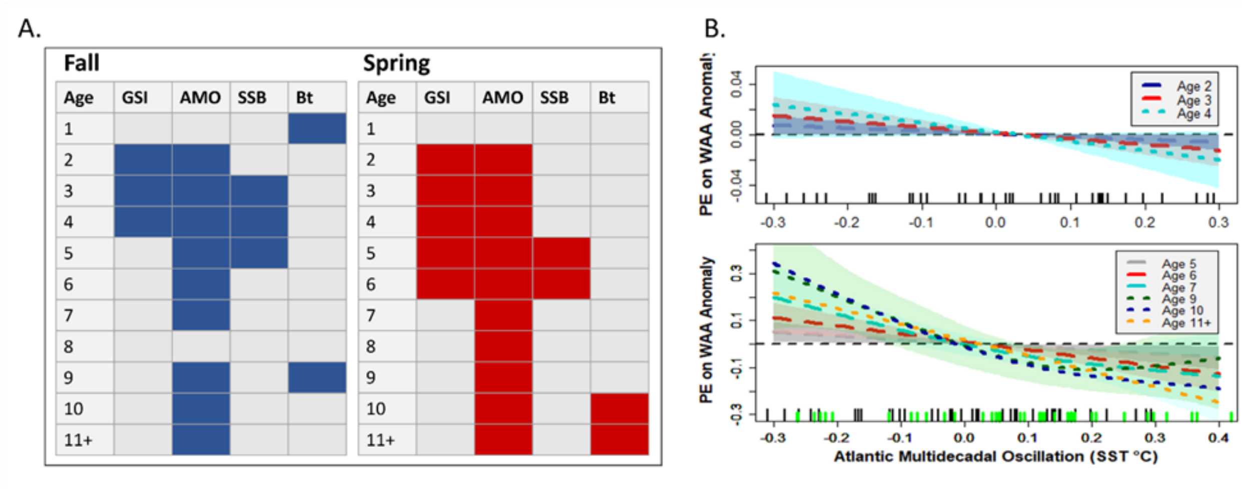


Figure 1.9. Environmental effects on weight at age. Significant environmental covariates at each age for fall in spring, denoted by colored cells (A), and GAM relationship curves between AMO and fall WAA anomaly data (B).

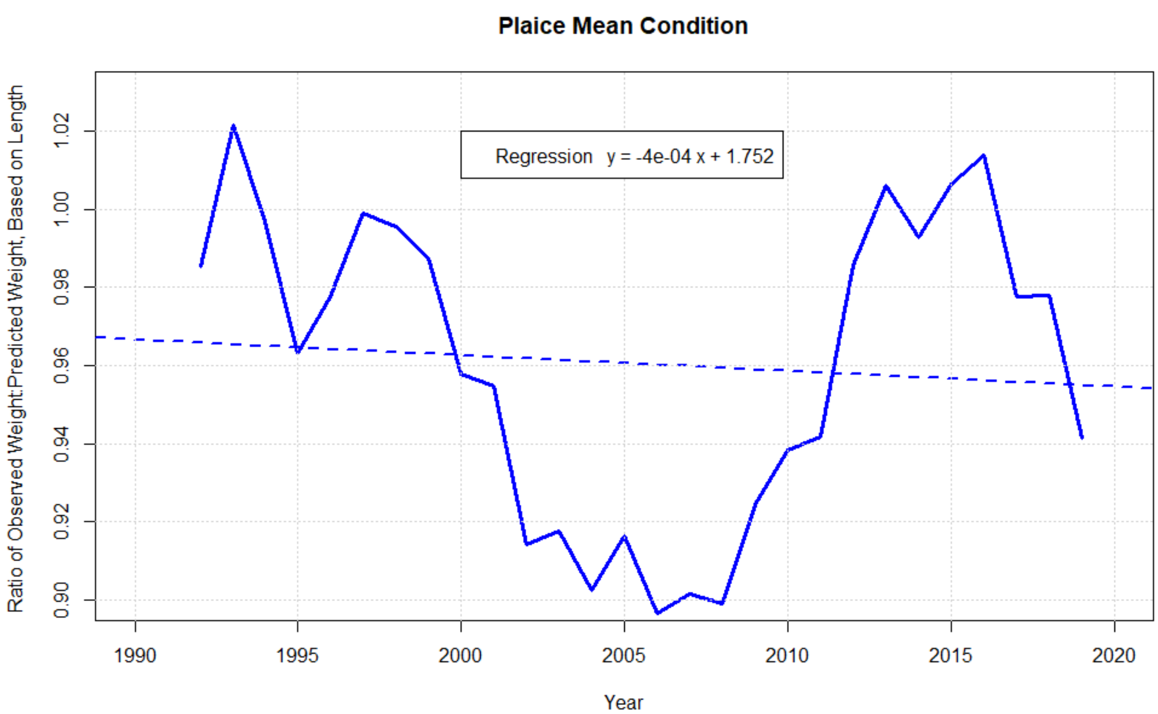


Figure 1.10. Mean condition factor from the fall survey.

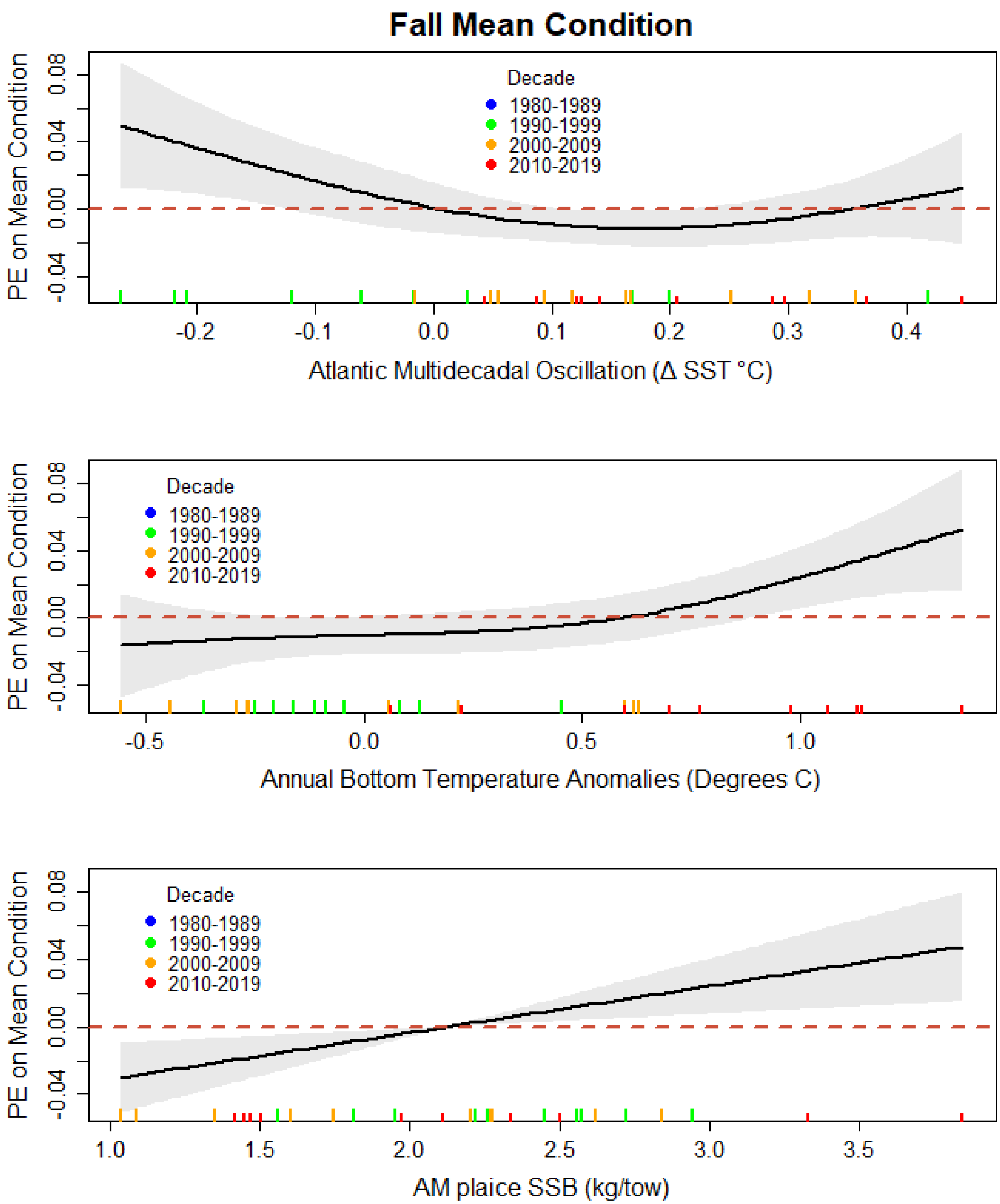


Figure 1.10. Environmental effects on condition factor.

Ecosystem Influences on Natural Mortality

Warmer conditions may increase mortality at the southern extent of a species' range, particularly in early life stages, by means of eggs hatching at lethal temperatures and by reduced growth rates which may increase predation and starvation (Nye et al. 2009; Levangie et al. 2021). In addition, unusually cold temperatures have also been associated with increases in natural mortality levels of American plaice in Canadian waters (Committee on the Status of Endangered Wildlife in Canada 2009). In a multispecies analysis of the Scotian Shelf system, Levangie et al. (2021) estimated an increase in natural mortality rate of American plaice associated with warming and a decrease in maximum size.

For juvenile and small plaice, cod are main predators, but mortality from cod predation has decreased due to declines in large cod abundance (Powles 1965). A predation study was conducted between juvenile Atlantic cod, Atlantic herring *Clupea harengus* (predators) and American plaice larvae (prey) at natural temperature extremes (8 and 13°C), to assess the effects of water temperature on predator-prey interactions (Fuiman and Batty 1994). They found cod and herring to have a 96% and 83% predation success rate on plaice larvae, respectively, and that predator capture and prey evasion rates were not affected by water temperature (Fuiman and Batty 1994).

Summary

A review of the scientific literature on American plaice and exploratory analyses identified the following:

- Temperature and depth are important drivers of American plaice distribution.
    - Plaice have been observed at greater depths during winter months compared to summer months.
    - Larger plaice have been observed at greater depths than smaller plaice.
    - Occupied depth has increased in response to ocean warming.
- Temperature, as well as NAO, were identified as having a significant influence on recruits per spawner of American plaice. The greatest historical recruitment rates occurred at extreme

cold temperature, but recent updated analyses indicate a general trend of increasing recruitment with increasing ocean temperature, except for the most recent year-classes which deviate from this trend.

- Warmer temperatures have been associated with accelerated growth rates, reductions in body size, and earlier ages at maturity of American plaice.
- Increases in the natural mortality rate of American plaice has been associated with ocean warming.
- Suitable habitat is expected to decrease with continued changes in ocean thermal conditions.

These ecosystem influences were considered in ToR2 by recognizing decadal shifts in growth and standardizing fishery catch rates, in ToR3 by investigating factors of survey catchability as well as integrating inshore and offshore surveys, in ToR4 by recognizing decadal shifts in growth and exploring environmental covariates of recruitment and survey catchability, in ToRs5-6 by considering evidence for alternative stanza of recruitment to inform biological reference points and projections, and in ToR7 by proposing additional research recommendations.

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*."

Contributors:

Larry Alade, Steve Cadrin, Dan Hennen, Tyler Pavlovich, Paul Nitschke, Toni Chute, Kiersten Curti, Keith Hankowsky, Max Grezlik, Lucy McGinnis, Alex Hansell, Andy Jones, ACate O'Keefe and Melissa Sanderson

The Working Group attempted to derive the longest time series of total catch. Recent stock assessment updates reported commercial landings and discards since 1980 (NEFSC 2015, 2017, 2019), but previous assessments reported landings since 1960 and discards since 1980 (NEFSC 1992, 1999, 2001, 2002, 2005, 2008, 2012), and the fishery management plan describes landings back to the 1940s but does not report annual statistics (NEFMC 1985).

The final year of fishery data in this assessment is 2019, because of delays with the Catch Accounting Monitoring System to derive discards and age composition for 2020. The methods developed by this research track assessment will be applied to a management track assessment in summer 2022, updated with data through 2021.

Since 1980, total catch of American plaice has ranged from 1,048 mt to 15,540 mt (Table 2.1, Figure 2.1). Commercial landings have been the predominant source of fishery removals averaging 88% of the catch between 1980 and 2019. Estimates of commercial landings, commercial discards, age composition, catch rates, and fishery perspectives are described below.

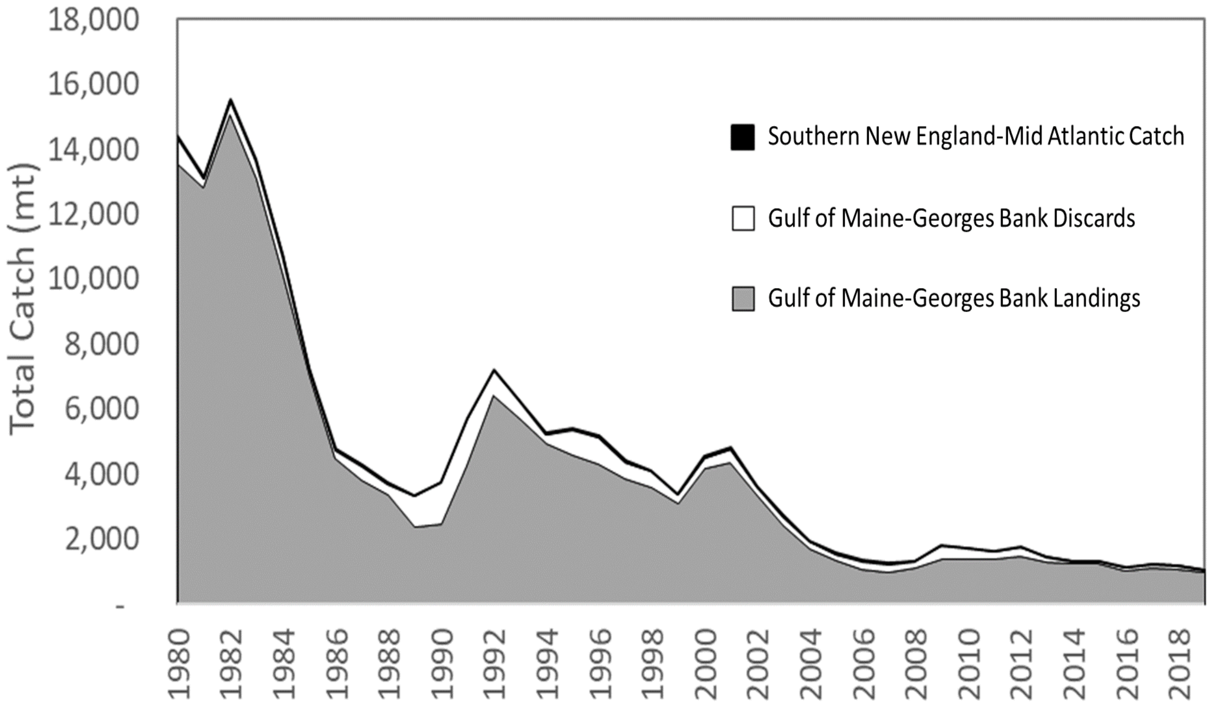


Figure 2.1. Estimates of American Plaice landed and discarded catch (1980-1988 discards are from NEFSC 1992).

Table 2.1. Estimates of American Plaice catch (mt) by disposition (landed and discarded). Estimates of 1980-1988 discards are from NEFSC (1992).

	US Landings	US Discards	Russian catch in US waters	Total
1960	1,309		-	
1961	1,522		-	
1962	1,927		-	
1963	2,182		24	
1964	3,611		-	
1965	3,343		112	
1966	3,343		279	
1967	3,236		1,018	
1968	3,254		338	
1969	3,432		412	
1970	2,594		945	
1971	2,176		340	
1972	1,794		429	
1973	1,602		447	
1974	2,076		20	
1975	2,423		148	
1976	3,509		3	
1977	7,068		128	
1978	9,503		-	
1979	11,330		-	
1980	13,549	845	-	14,394
1981	12,881	288	-	13,169
1982	15,126	414	-	15,540
1983	13,141	534	-	13,675
1984	10,134	619	-	10,753
1985	7,029	236	-	7,265
1986	4,472	290	-	4,762
1987	3,801	463	-	4,264
1988	3,382	349	-	3,731
1989	2,353	985	-	3,338
1990	2,445	1,317	-	3,762
1991	4,261	1,470	-	5,731
1992	6,416	786	-	7,202
1993	5,720	556	-	6,276
1994	4,977	287	-	5,264
1995	4,617	801	-	5,418
1996	4,363	804	-	5,166
1997	3,890	518	-	4,409
1998	3,631	495	-	4,126
1999	3,108	289	-	3,397
2000	4,196	351	-	4,547
2001	4,400	413	-	4,814
2002	3,399	223	-	3,622
2003	2,426	305	-	2,732
2004	1,710	246	-	1,956
2005	1,337	229	-	1,567
2006	1,094	258	-	1,352
2007	988	259	-	1,247
2008	1,100	233	-	1,333
2009	1,380	426	-	1,806
2010	1,404	334	-	1,738
2011	1,368	254	-	1,622
2012	1,462	302	-	1,765
2013	1,297	168	-	1,465
2014	1,239	92	-	1,331
2015	1,231	86	-	1,317
2016	1,009	108	-	1,117
2017	1,134	102	-	1,235
2018	1,087	109	-	1,196
2019	970	78	-	1,048

Commercial Landings

Estimates of American plaice commercial landings are available back to 1960 after the development of an integrated system for monitoring fishery statistics (Rounsefell 1948). In 1994, the monitoring system transitioned from port interviews to fishermen logbooks (vessel trip report, VTR; NEFSC 1996). Landings estimates for 1960-1993 were from previous stock assessments (NEFSC 1999b, O'Brien et al. 1999), and landings for 1994-2019 were re-estimated for this research track assessment. A standardized procedure has been used to assign area and fishing effort from logbooks to dealer-reported landings for 1994-2019 (Wigley et al. 2008).

Landings from domestic and foreign fleets in US waters averaged approximately 3,000 to 4,000 mt per year in the 1960s, peaked at 15,000 mt in 1982, and has been approximately 1,000 mt per year since 2005 (Figure 2.2). Almost all landings were from trawl trips, with <1% from gillnet and scallop dredges. Landings are from a range of vessel sizes, and the portion of landings from the largest vessels (tonnage class 4) gradually increased in the last decade. Landings were mostly in spring summer (quarters 2-3) in the 1990s but have been more evenly distributed among seasons in recent years (Figure 2.3). American plaice are landed in four market categories: small, medium, large and unclassified. Large plaice dominated annual landings in the late 1970's through the late 1980's, followed by the expansion of the small market in the 1990's through 2010's (Figure 2.4). In the last decade, large fish became less common, with mediums small and unclassified accounting for approximately 85% of the total landings.

Gloucester and New Bedford were the primary ports where plaice were landed in the 1960s and early 1970s, then Portland became a major port for plaice landings in the 1990s and early 2000s, and most landings in the last decade were from Gloucester and Boston. Most landings have been from the Gulf of Maine (statistical areas 512-515) and the northern edge of Georges Bank (statistical areas 521, 522 and 561), but the distribution of landings extends to the Mid Atlantic Bight, particularly in earlier years (Figure 2.5). The proportion of landings from the Gulf of Maine has increased in the last decade, particularly from statistical area 513 (western Gulf of Maine).

Although commercial landings are generally well estimated from electronic dealer reports, there have been some notable misreporting violations. For example, in 2015, 73 mt of landed American plaice were reported as haddock; and in 2016, 2 mt of landed American plaice were reported as haddock (USA vs. Rafael and Freitas 2016). An analysis of fishery monitoring data identified 350 fishing trips from 2011 to 2015 (<1% of total groundfish trips in that period) that may have misreported stock area for up to 2.5 million pounds of groundfish, including American plaice (USCG 2019). The extent of misreporting magnitude and time series is unknown. It may be limited to the one vertically integrated operation with violations or more pervasive. Incentives for misreporting results are greater when the relative availability of American plaice is greater than the relative catch allocations so that catches of plaice constrain fishermen's ability to catch other groundfish stocks. An analysis of market prices and lease prices suggests that incentives for misreporting American plaice started in 2014 because of increasing lease prices (Henry et al. 2019) resulting from the substantial reduction in annual catch limits in 2013 based on the 2012 stock assessment (NEFSC 2012). Lease prices of American plaice increased from 2012 to 2015 ($0.05 in 2012, $0.26 in 2013, $0.63 in 2014, $0.98 in 2015; in 2010 dollars, Murphy et al. 2018) Annual catch limits for American plaice substantially increased in 2021, based on the 2019 stock assessment (NEFSC 2019), reducing the incentives for misreporting.

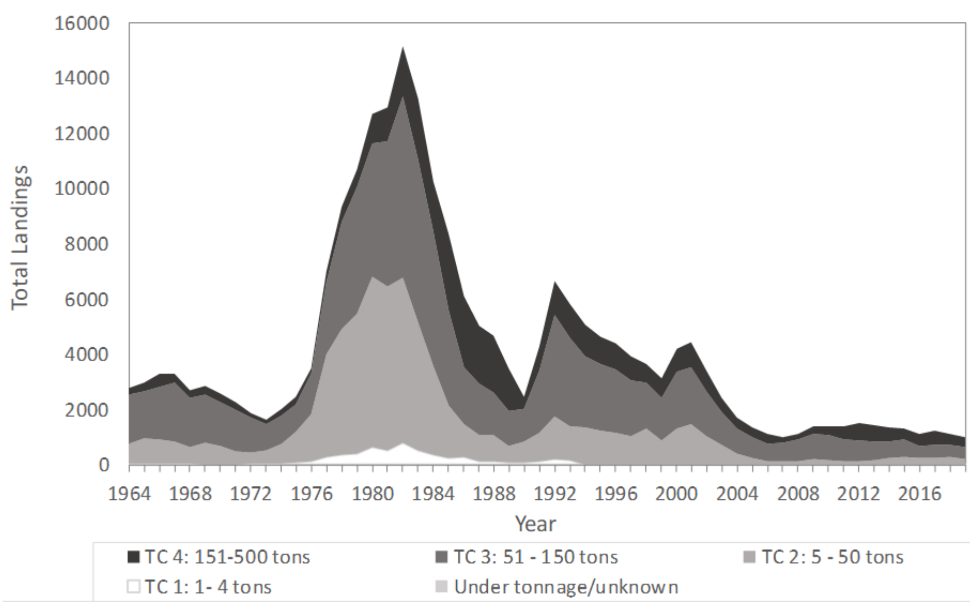


Figure 2.2. Commercial landings of American plaice by vessel tonnage class.

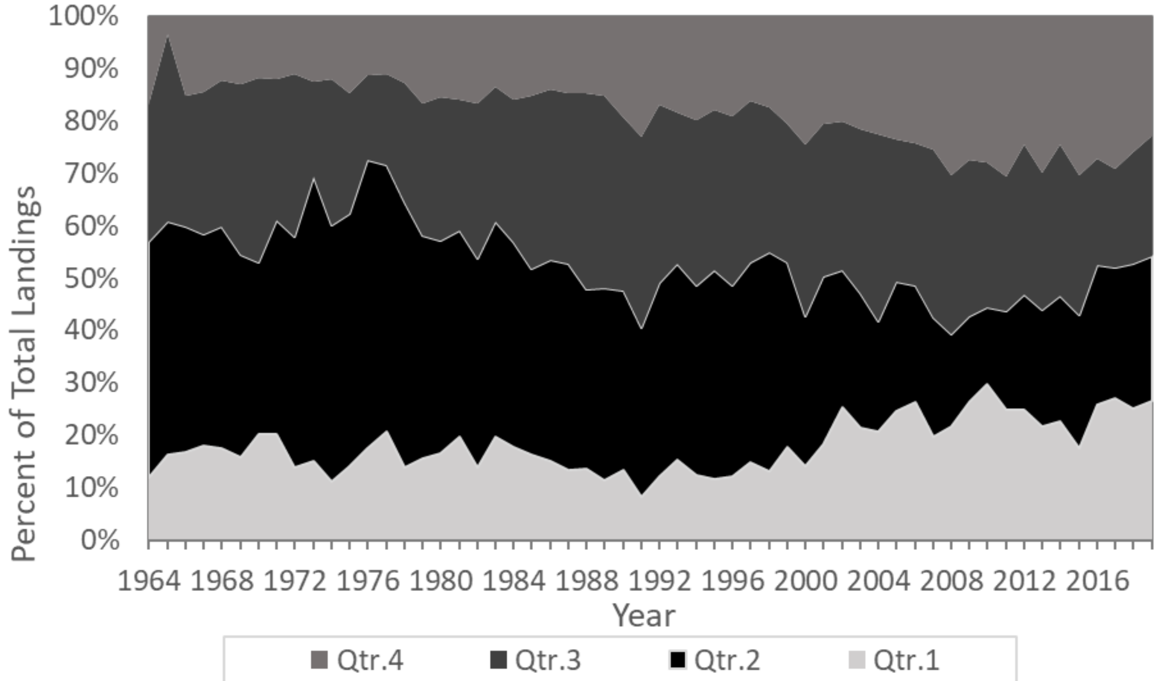


Figure 2.3. Proportion of American plaice commercial landings by quarter-year.

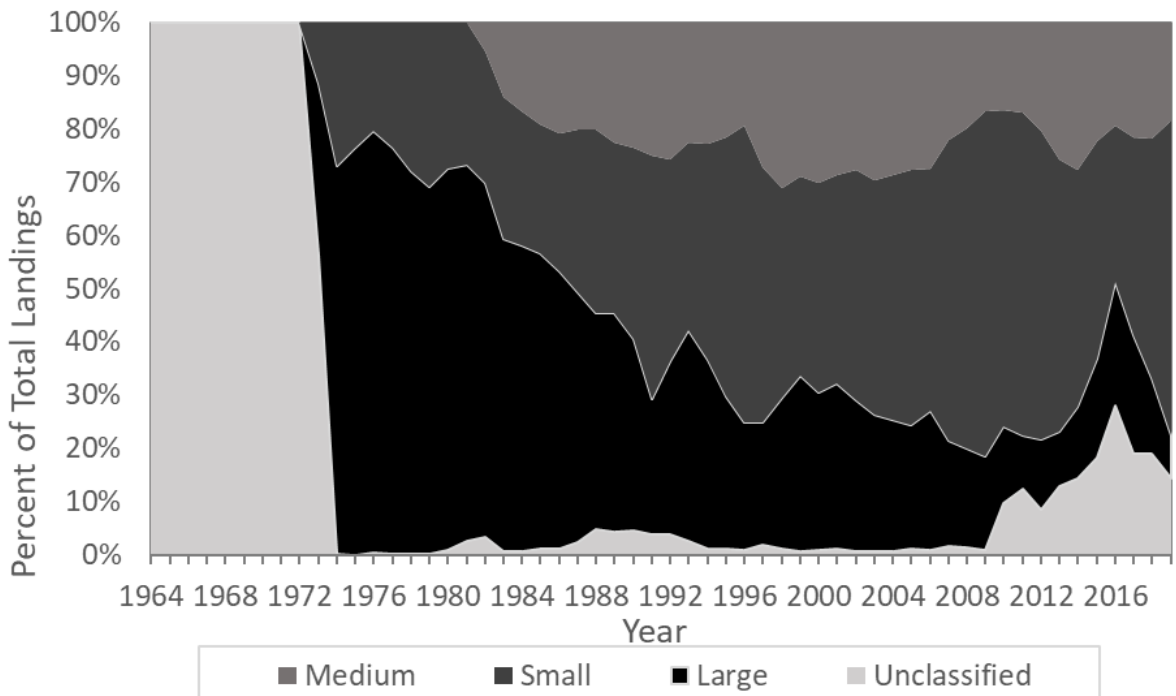


Figure 2.4. Proportion of American plaice commercial landings by market category.

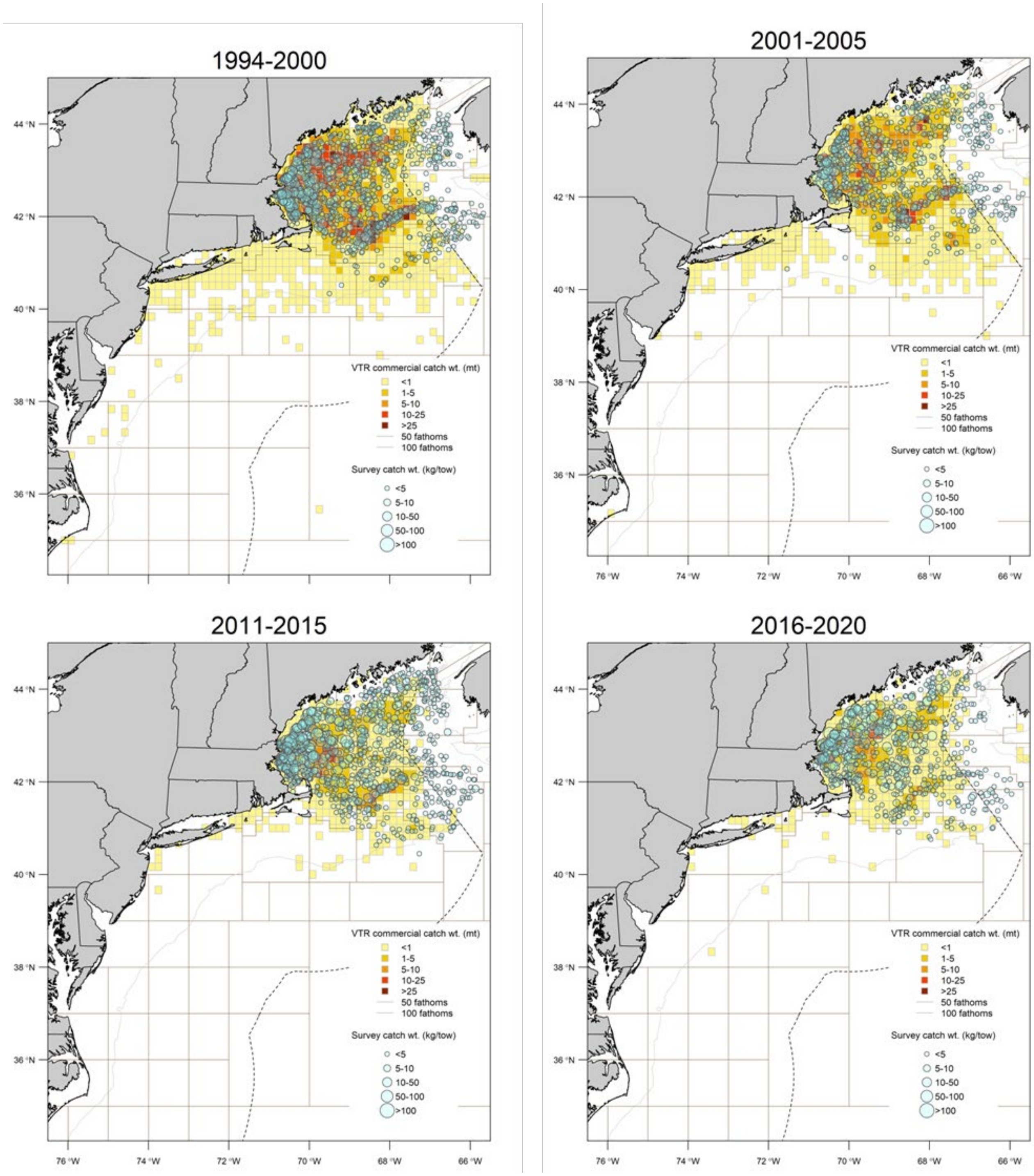


Figure 2.5. Commercial landings of American plaice (squares) and NEFSC Spring and Fall survey catches (circles) since 1994 by five year time blocks derived from the spatial information provided on VTRs.

Landings from the Electronic Monitoring Maximized Retention Program

Some fishermen have participated in electronic monitoring programs since 2018. Electronic monitoring replaces or supplements at-sea observers with integrated camera systems on vessels and has been used to collect fishery-dependent data across a range of fisheries in several countries (van Helmond et al. 2020). Potential benefits of electronic monitoring include reduced uncertainty in estimates of catch, increased fleet coverage, accurate discard estimates, and precise catch per unit effort time series. Additionally, electronic monitoring has the potential to improve timeliness and quality of data collection, processing, and analyses, as well as incentivize accountability to support sustainable fisheries (McElderry 2006).

The development of electronic monitoring programs for the New England groundfish fishery began with initial pilot testing in 2010 (Fitz-Gerald et al. 2019, van Helmond et al. 2020). Phase I determined the baseline data required to monitor groundfish quota, including detection of fishing events, counting fish, and species identification. Phase II developed methods to obtain fish weights and improve species identification through catch handling protocols. Phase III tested catch handling protocols to simulate an operational EM program and identify the components necessary to support different EM approaches, including an audit program and maximized retention program (Fitz-Gerald et al. 2019).

The maximized retention program began under an experimental fishing permit in the 2018 fishing year. The goal of the program is to document allocated groundfish through dealer reports that are verified by dockside monitoring. All allocated groundfish caught by vessels in the program, regardless of size or condition, are retained at-sea, landed and recorded by a dealer. The dealer report is the official catch record for each fishing trip and includes the documentation of undersized and damaged groundfish using maximized retention program market codes. All electronic monitoring data are reviewed to confirm that allocated groundfish are not discarded at sea. Offloads for fishing trips following maximized retention protocols are observed by NOAA dockside monitoring to verify dealer weights and collect biological data normally collected at sea, with a focus on collecting lengths from the exempted portion of the catch. Data from the program are used to improve catch accounting for quota monitoring.

Since 2018, the maximized retention program improved data collection, refined catch handling protocols and developed review guidance by recruiting up to eight high-volume vessels to participate in the program. As of May 1, 2021, there were six vessels participating under the maximized retention program, and the program is expected to be operational and available as a monitoring option in the Sector Operations Plans for the 2022 fishing year. In addition, review rates are expected to decrease to 50% under an operational program and dockside monitoring is expected to transition to a third-party model.

Landed catch from the maximized retention program included three additional disposition categories (1246-terminal, 1247-undersize, and 1248-terminal/undersize mix). The total annual catches in these categories were 0.53 mt in 2018, 4.73 mt in 2019, and 0.48 mt in 2020. The few groundfish discarded at-sea in the maximized retention program are counted and identified to the species level. Of the 440 maximized retention trips that have occurred since the start of the program through 2020, only 230 individual American plaice were observed to be discarded in the program.

Commercial Discards

Most American plaice are discarded because they are smaller than the minimum size regulation. With increases to the commercial minimum sizes during the early 1980's and 1990's, discarding due to undersize fish accounts for approximately 67% of total fish discards.

Discard Estimates from At-Sea Observers

Historical discards (1980-1988) were estimated by previous assessments using fishery and survey size distributions, with assumed size selectivity and retention (NEFSC 1992). Discards for 1989-2019 were re-estimated by expanding discards rates (discards/kept) sampled by at-sea observers to all trips with a standardized bycatch reporting method (Wigley et al. 2006). Discard estimates from observer data in previous assessments were stratified by quarter-year, three fleets (large mesh ≥ 5.5 inch trawl and shrimp trawl) and two regions (Gulf of Maine and Georges Bank). For this research track assessment, the working group evaluated alternative stratifications that produced nearly identical estimates, but some stratifications resulted in lower variance, particularly early in the time series. Based on precision, discard estimates were stratified by five fleets (large mesh >5.5-inch trawl, small mesh trawl, shrimp trawl, gillnet, and scallop dredge) in two regions (Gulf of Maine-Georges Bank, Southern New England-Mid Atlantic Bight; Table 2.1.6).

Estimates of discards peaked at almost 1,500 mt in 1991, gradually decreased to approximately 200 to 400 mt per year in the 2000s, and further decreased to approximately 100 mt per year since 2014 (Figure 2.6). Most discards were from the large-mesh trawl fleet, with considerable discards from the shrimp fishery in the late 1980s and early 1990s and relatively small contributions to total discards from the small-mesh trawl, gillnet and scallop dredge fisheries. Discards were a relatively large portion of total US catch until 1992, ranging 15% to 40% of total catch 1980-1991), then decreased to approximately 5-20% of annual catch since then and have been less than 10% of total catch since 2014 (Figure 2.1, Table 2.1), resulting from a decrease in the minimum legal size from 14 in. (35.6 cm) to 12 in. (30.5 cm) in May 2013 (Appendix B).

The Groundfish Plan Development Team identified an 'observer effect' in which observed trips were significantly shorter and had other significant differences when compared to unobserved trips (Demarest 2019). An investigation of incentives for illegal discarding, based on market prices and lease prices, Henry et al. (2019) found that there were no incentives to discard American plaice before 2014, and there were discard incentives on <10% of groundfish trips since then (1% in 2014, 4% in 2015, 3% in 2016 and 8% in 2017).

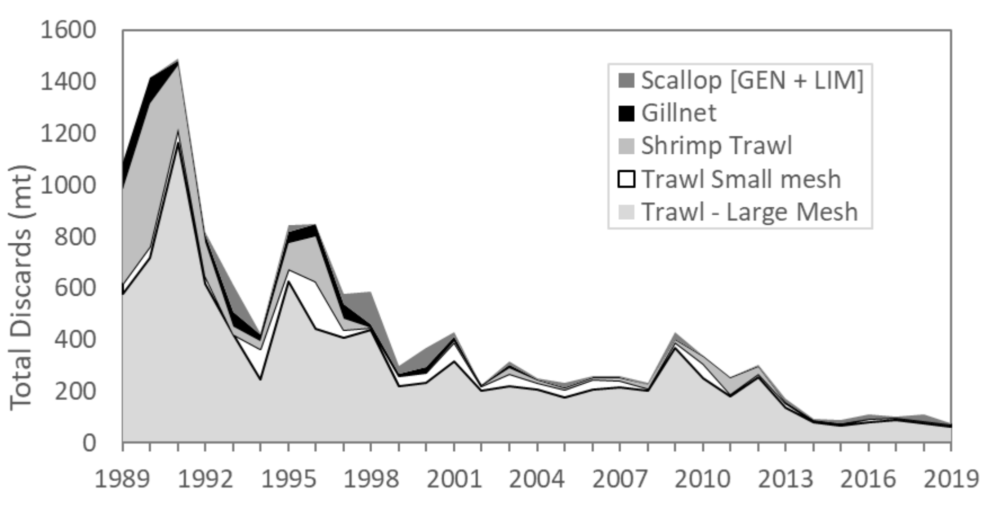


Figure 2.6. Discard estimates of American plaice and 95% confidence limits.

Table 2.2. Number of fishing trips observed from 1989 to 2019, summarized by region and gear type (large-mesh trawl, small-mesh trawl, shrimp trawl, gillnet, scallop dredge) from at-sea monitors and observers.

	Gulf of Maine-Georges Bank					S.New England-Mid Atlantic			
Year	LgMesh	SmMesh	Shrimp	Gillnet	Dredge	LgMesh	SmMesh	Gillnet	Dredge
1989	56	45	40	106		9	75	1	
1990	46	22	31	149		17	61	6	
1991	72	41	53	955	1	13	117	5	1
1992	62	28	82	1031	10	22	68	155	8
1993	33	12	82	619	11	23	33	154	11
1994	26	2	77	122	7	26	25	283	19
1995	49	34	73	127	7	68	70	411	20
1996	24	43	35	74	15	36	86	351	25
1997	19	7	17	48	11	18	61	342	19
1998	8	1		121	10	15	37	281	16
1999	33	15		102	62	15	41	87	8
2000	94	6		124	228	33	48	103	28
2001	139	14	3	74	17	61	60	92	88
2002	207	49		100	13	50	49	74	87
2003	374	41	15	385	14	27	98	106	109
2004	426	96	12	948	60	204	310	149	236
2005	1102	157	17	815	110	348	261	179	221
2006	519	47	20	174	131	165	180	78	97
2007	529	33	14	232	180	244	251	102	179
2008	679	20	19	219	212	229	178	90	432
2009	732	50	12	326	79	295	404	91	409
2010	878	69	15	1671	99	373	486	207	239
2011	1287	54	1	1806	137	482	439	256	252
2012	1322	47	19	1459	219	369	307	233	201
2013	813	105	24	747	270	500	480	86	202
2014	792	88		1119	191	599	620	254	281
2015	575	111		572	176	499	492	493	351
2016	414	83		285	174	452	870	704	454
2017	389	164		281	255	540	1315	820	389
2018	426	118		197	329	485	1001	633	227
2019	654	65		344	308	619	963	828	239

Discard Estimates from Electronic Monitoring

Amendment 23 to the New England groundfish Fishery Management Plan includes a target at-sea monitoring coverage rate of 100% for all groundfish sector trips for four years to establish a baseline of accurate and precise catch for deriving an effective coverage for the future (NEFMC 2021). Groundfish sectors can choose to achieve the target 100% at-sea monitoring coverage rate by hosting an observer or using electronic monitoring. The Working Group considered estimates of discards from electronic monitoring programs. Although catch from electronic monitored trips are a small portion of American plaice catch for 2018-2019, participation in these programs is expected to increase. Therefore, future assessments should consider deriving discards from electronic monitoring for vessels in those programs.

The audit model program was conducted under an Experimental Fishing Permit starting in 2016. The goal of the program was to validate fishermen self-reported discards of groundfish species by auditing electronic Vessel Trip Reports in comparison to electronic monitoring data collected from the same trips. Between 2016 and 2019, the audit model program evolved to incorporate improved data collection and review protocols. Electronic monitoring data were used in place of at-sea monitoring data for a subset of trips in 2016, then for all audit program trips in fishing year 2017. In fishing year 2018, the audit component was introduced with review of 100% of electronic monitoring data to compare to electronic Vessel Trip Reports. In fishing year 2019, the review rate was reduced to 50% of the electronic monitoring data from all trips in the program, and electronic Vessel Trip Reports data were applied as the official catch record for trips that passed the audit.

Starting in fishing year 2021, the audit model was operationalized to allow New England groundfish sectors the option to adopt the audit model program as part of their Sector Operations Plans (NOAA 2020). The operationalized audit model program requires participants to record the estimated weight of all discards on an electronic Vessel Trip Reports and follow catch handling protocols at sea. To facilitate collection of discard data from video footage, participants must place all groundfish species under a camera on a measuring strip prior to discarding. As of May 1, 2021, eight sectors included an electronic monitoring component in their fishing year

2021-2022 Sector Operations Plans, and there are currently 22 active sector vessels using electronic monitoring to meet monitoring requirements.

Data from the audit model program have been used primarily to account for discards in sector quota monitoring. All regulated groundfish discards are handled and discarded in camera view and data is collected for species, length, weight, count, and catch disposition. Descriptions of the audit program catch handling protocols, video reviewer guidance, and data collection fields, are included in NOAA's "Electronic Monitoring Audit Model Program Reviewer Guidance Manual_V18," (NOAA, 2021b; https://apps-nefsc.fisheries.noaa.gov/NEMIS/index.php/docs/guidance). Details of the data fields collected from video footage are included on the Northeast Electronic Monitoring Information System application programming interface (NEMIS API; https://apps-nefsc.fisheries.noaa.gov/NEMIS/index.php/docs).

To date, audit model electronic monitoring data streams have not been incorporated in stock assessments for New England groundfish species. Analyses conducted during the experimental phase of the audit model program indicate that aggregate and trip-level discard estimates from electronic monitoring and Northeast Fisheries Observer Program data were similar in 2017 and 2018. The average difference in estimates of discarded weight by species between the datasets was small (~3.6 pounds per trip in 2017 and ~15.9 pounds per trip in 2018) and not statistically significant in either year at the fleet and trip levels (Cadrin et al. 2020). Audit model data for American plaice were provided by Teem Fish Monitoring and the Cape Cod Commercial Fishermen's Alliance (Table 2.3). Total discard weight was calculated by multiplying the number of fish by the average fish weight in pounds. Total landed weight comes from dealer reports for all audit model vessels combined.

Table 2.3. Audit Model electronic monitoring data for fishing years 2019 and 2020, including all recoded discards and total landed weight from all Audit Model EM vessels.

Fishing Year	Vessels	Trips	Discards (number)	Mean Wt (kg)	Discards (mt)	Landings (mt)
2019	5	26	627	0.21	0.13	31.50
2020	8	82	9030	0.18	0.66	15.26

The Working Group decided to estimate discards for all trips using at-sea observer samples, including electronic monitoring trips. Therefore, discard estimates include all trips with no duplication of discard estimates. Although all trips, including electronic monitoring trips, are sampled by the Northeast Fishery Observer Program, electronic monitoring trips are exempted from the additional At-Sea Monitors. Therefore, electronic monitoring trips are not sampled by observers at the same intensity as other trips. As the number of participating vessels increases, discards should be estimated directly from electronic monitoring data. However, the integrated Catch and Monitoring System is still in development.

Discard Mortality

Previous plaice assessments assumed 100% discard mortality (i.e., all fish discarded are dead). Previous post-release mortality field studies of plaice from the Gulf of Maine shrimp fishery and some Canadian fisheries suggest that plaice can survive after being caught in trawls and discarded (Powles 1969, Ross and Hokenson 1997, Benoit et al. 2012). Estimates of discard survival varied (60-100%) as a function of deck time and fleet. General consensus from fishermen who participated in the research track workshops was that discard mortality of plaice is relatively high, and assuming 100% discard mortality is realistic. Data from Electronic Monitoring trips that caught plaice indicated that pre-release processing time was approximately one hour, which corresponds to 85-90% mortality (Benoit et al. 2012). Considering that the field estimates of discard mortality did not account for other factors such as bird predation, air temperature and depth of tow, the Working Group decided to maintain the 100% discard mortality assumption for the research track assessment.

Age Composition

Age was determined by examination of the thin-section and cut surfaces of one otolith (Dery 1988). Otoliths age determination was validated by marginal increment analysis (Powles 1965, 1966; Pitt 1967). Age determination has been validated for American plaice in Canadian waters, but not for US waters, where otolith patterns can be more difficult to interpret (Lux 1969, Dery 1988). Two readers independently processed fall 2009 NEFSC survey samples and reconciled 96% of samples (Appendix A, Working Paper 2). Since 2004, 79 precision and accuracy tests included 3,583 otoliths subsampled from a total of 19,614 samples and produced an average coefficient of variation (CV) of 3.1% (ranging 1.8% to 5.8%) and an average agreement of 75% (ranging 52% to 85%; https://apps-nefsc.fisheries.noaa.gov/fbp/QA-QC/ap-results.html). For the last five years of samples, 25 precision and accuracy tests included 2,195 subsamples from the total of 13,888 samples and produced a 4.10% CV and 67.3% agreement, which were relatively less precise (CV was the 90th percentile) and was the least consistent among the 21 species recently tested (https://www.fisheries.noaa.gov/resource/data/accuracy-and-precision-fish-ages-northeast).

Patterns of length at age were explored to investigate changes in growth over time and regional variation for determining age composition in the fishery and in surveys (Appendix A, Working Paper 3). Data from the NEFSC bottom trawl surveys, 1980-2018, in the Gulf of Maine and Georges Bank survey strata. Observations of length at age were visualized as scatterplots and as fitted von Bertalanffy functions by subregion and decade. The number of old fish and age of the oldest fish was stable over time and space, with the exception of relatively few old fish in the southwestern Gulf of Maine in the 1980s. Results suggest that growth has tended towards faster rate of growth and smaller maximum size (asymptotic length decreased approximately 20 cm; Figure 2.7). There does not appear to be much difference in growth between subregions, geographic differences are inconsistent (i.e., relatively faster growth in an area one decade, then slower growth in the area during the next decade), and geographic variation in growth is much less than differences among decades. Therefore, the Working Group decided to determine annual age composition for the entire stock, with empirical age-length keys and no regional stratification.

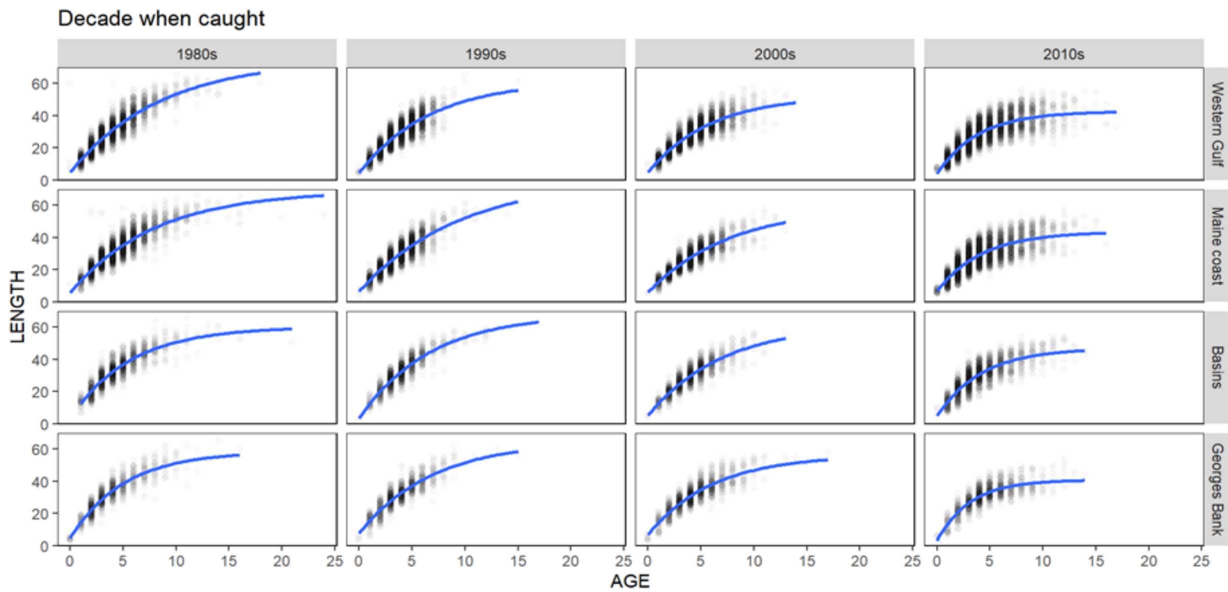


Figure 2.7. Length at age observations of American plaice and von Bertalanffy functions by decade and subregion.

Length-weight relationships were also explored to investigate geographic variation and changes over time (Silver et al. 2021, Working Paper 3 Appendix A). Length-weight observations from NEFSC surveys in the Gulf of Maine and Georges Bank, 1992-2021, were grouped into regions, sex, season, and 5-year periods to explore spatial, seasonal, and temporal trends. Length-weight relationships were relatively stable among regions, seasons, and time periods (Figures 2.8-2.9). Therefore, the Working group decided to use a single, stock-wide length-weight relationship for determining catch in numbers and age composition (W in kg, L in cm): $W = 2.5511\text{e}-6\ L^{3.331}$

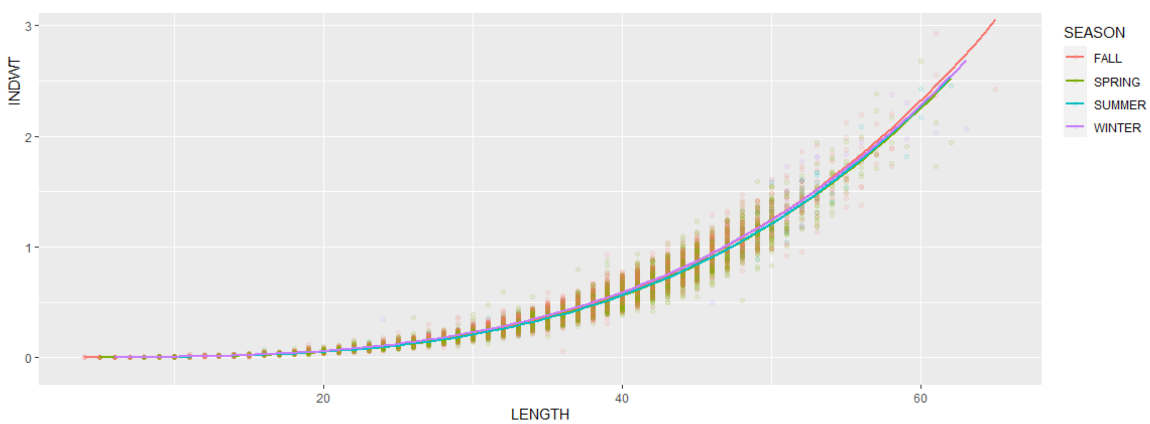


Figure 2.8. Observations and predicted relationships of length and weight of American plaice by season.

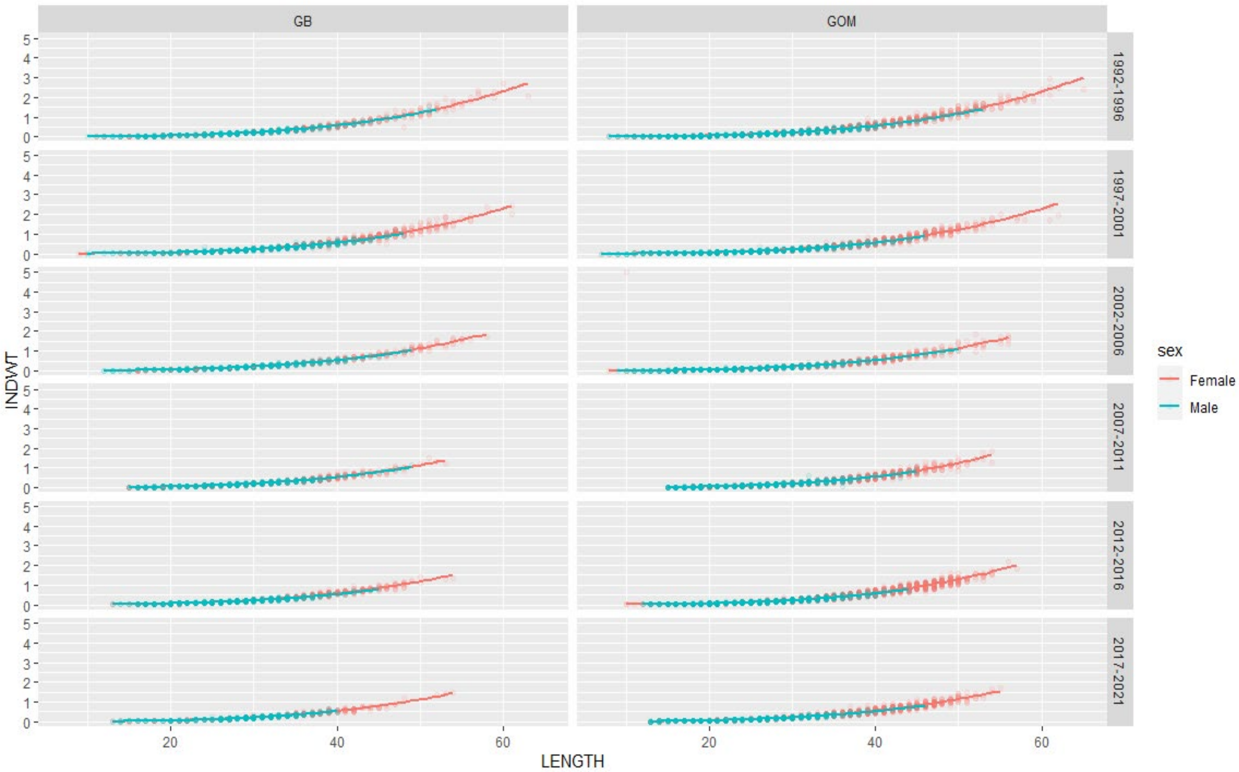


Figure 2.9. Observations and predicted relationships of length and weight of American plaice by region (GB: Georges Bank; GOM: Gulf of Maine and 5-year periods.

Age Composition of Commercial Landings

Plaice landings have been sampled for length and age composition since 1975, but ages were not routinely sampled until the mid-1980s, and age sampling was stratified by quarter, market category and statistical area. Age composition for 1980-1993 were from previous stock assessments (NEFSC 1999b, O'Brien et al. 1999) in which 1980-1984 size frequencies were expanded to landings by quarter-year and three market categories ('small'+'peewee', 'medium', and 'large'+'jumbo'; Figure 2.10) using length-weight equations (Lux 1969), and age composition was estimated using seasonal survey age-length keys. Age composition was also regionally stratified for 1980-1994 to account for regional differences (NEFSC 1999b, O'Brien et al. 1999). Age composition was re-estimated for 1994-2019 based on available samples (Tables 2.4-2.5) by expanding size frequencies to landings by half-year and the same three market categories ('small'+'peewee', 'medium', and 'large'+'jumbo'; Figure 2.11) using a single length-weight relationship. Sampling intensity since 1994 has generally been strong, exceeding the NAFO/ICNAF standard 200 mt per 100 lengths in all years except 1995.

Landings in numbers were derived by expanding sample weights to landings in the Gulf of Maine-Georges Bank region by half-year and market category. Sampled weights were estimated by applying the revised length-weight equation (Silver et al. 2021, Working Paper 3 Appendix A). Semi-annual age-length keys were applied to half-year landings at age to derive landings-at-age at age. Landings at age for the Gulf of Maine-Georges Bank region were then expanded to include landings the Southern New England-Mid Atlantic region, which contributed a small portion of total landings (Figure 2.1) and was not well sampled. Bootstrapped CVs for landings-at-age were generally less than 30% for ages 4-, but greater for older and younger ages (e.g., CV>40% for ages 8+).

Age composition of landings have been relatively stable over time, mostly ages-3-8, with less catch at ages 2-3 after mesh size increases in the late 1990s (Figure 2.12, Table 2.6). Landings of plaice older than age-10 increased in the last decade. Weights-at-age from commercial landings were relatively heavy in the 1980s and 1990s, decreased in the 2000s and have been relatively light in the last decade (Table 2.7).

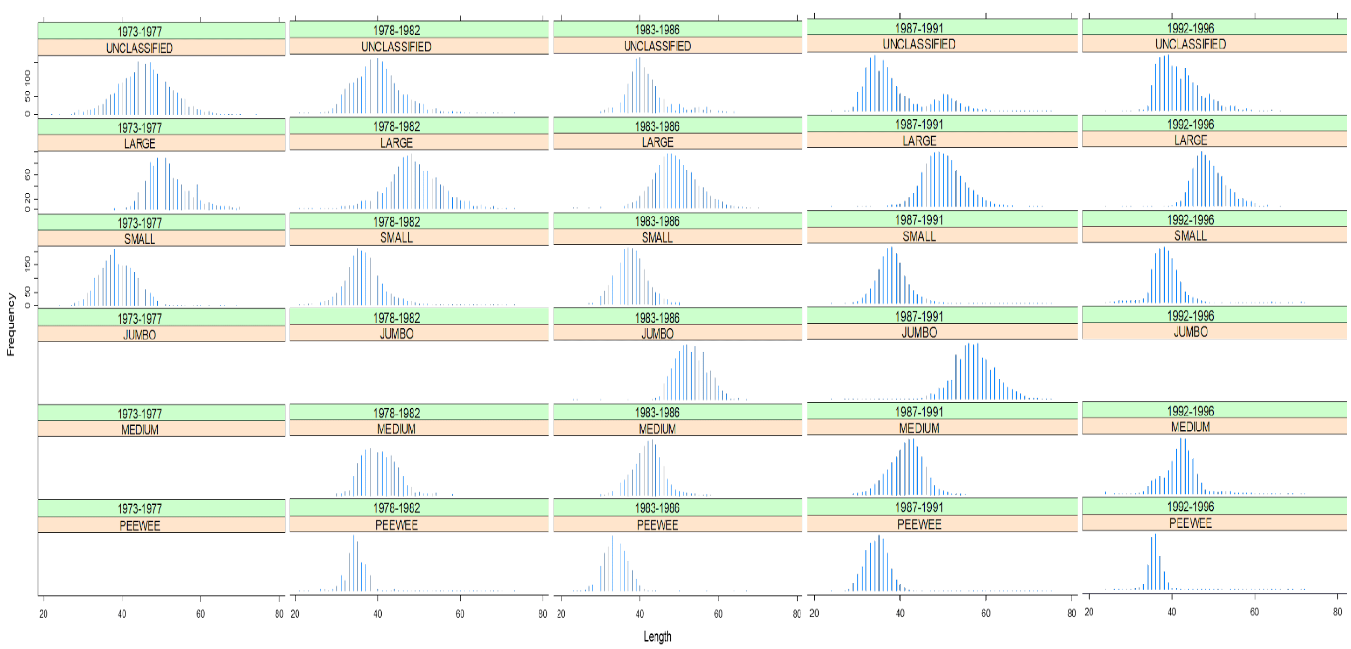


Figure 2.10. Length frequency of landed American plaice by market category, 1973-1996.

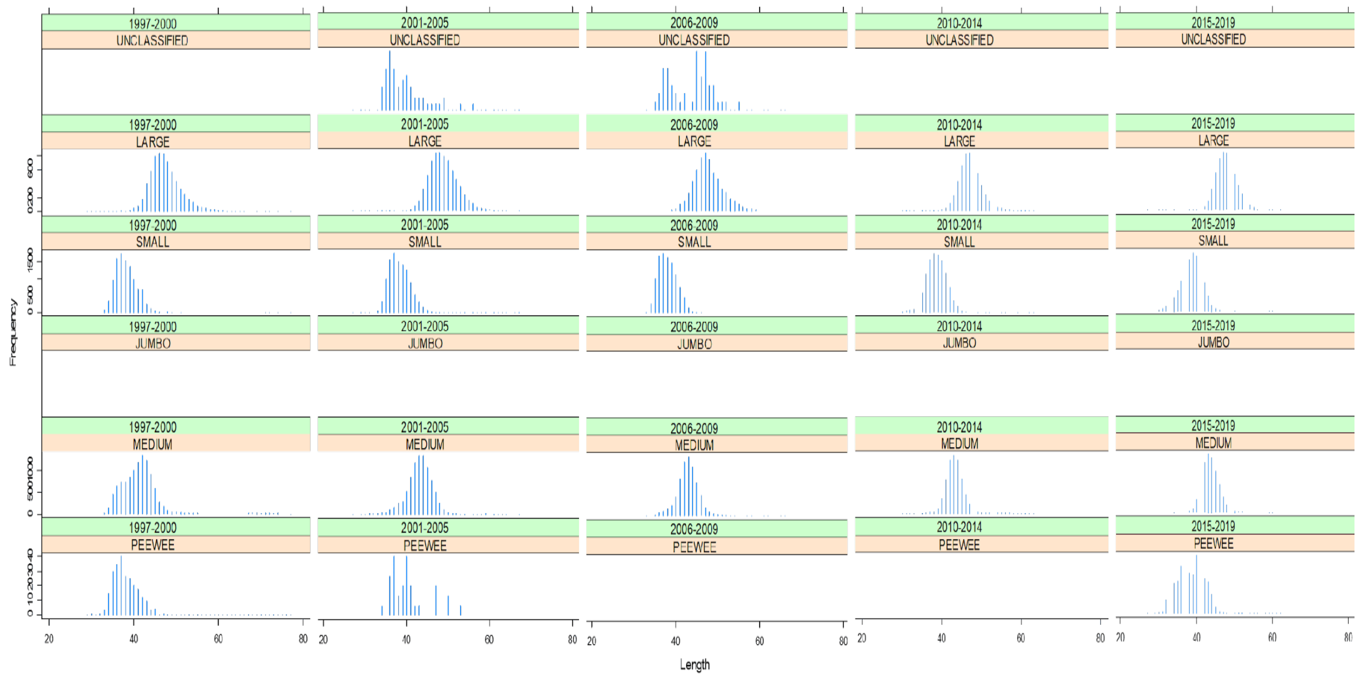


Figure 2.11. Length frequency of landed American plaice by market category, 1997-2019.

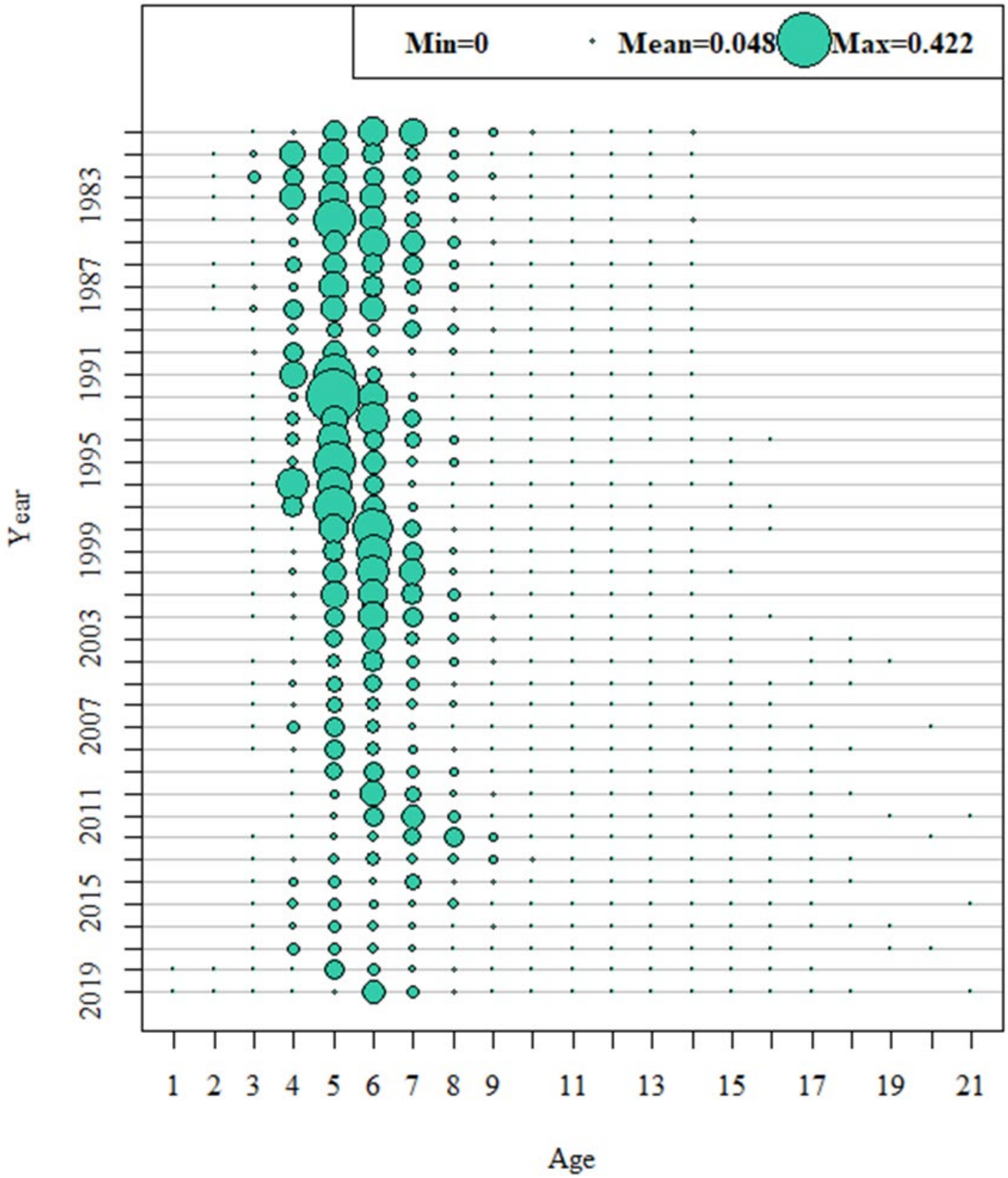


Figure 2.12. Age composition of American plaice landings.

Table 2.4. Number of American plaice sampled for length from commercial landings by market category and half-year.

	Unclassified		Large		Small		Medium		EM Undersize	
Year	Half1	Half2	Half 1	Half 2	Half 1	Half 2	Half 1	Half 2	Half 1	Half 2
1994			404	818	136	586	556	786		
1995			461	119	301	230	353			
1996			756	942	654	583	451	1532		
1997			1080	776	1147	675	1303	1785		
1998			588	518	701	364	1244	1455		
1999			1328	752	1841	1357	1904	1472		
2000			830	840	3872	748	965	468		
2001			944	829	1306	506	994	1112		
2002			863	731	649	573	616	650		
2003	61		1815	2035	991	814	849	1359		
2004			2257	1558	816	872	1055	1486		
2005			1812	889	823	835	840	844		
2006			1145	1152	605	763	1096	923		
2007	50		967	1123	672	736	751	1067		
2008			1128	1325	749	859	1435	1492		
2009			869	1240	738	954	1423	1700		
2010			1109	1538	769	730	1304	1263		
2011			1482	1285	834	720	1164	1259		
2012			1262	1325	982	776	1516	1315		
2013			965	1260	612	716	1268	1233		
2014			1174	1102	452	836	1039	906		
2015			1215	1497	855	1057	1305	1735		
2016			1468	1505	1494	1183	1312	1431		
2017			1644	1519	1439	1362	1498	1599		
2018			1262	677	1365	608	1313	691		1059
2019			636	311	717	350	1200	400	2597	2776

Table 2.5. Number of American plaice otoliths aged from commercial landings by market category and half-year.

	Large		Small		Medium	
Year	Half 1	Half 2	Half 1	Half 2	Half 1	Half 2
1994	119	152	64	106	85	202
1995	100	26	49	20	75	
1996	173	187	87	72	98	319
1997	215	145	178	99	162	174
1998	137	156	127	74	257	339
1999	257	148	403	209	385	236
2000	183	176	737	113	193	67
2001	194	151	223	98	126	160
2002	178	123	110	75	128	59
2003	431	449	193	188	181	256
2004	566	104	175	133	244	156
2005	263	260	111	153	144	225
2006	326	278	100	138	223	209
2007	228	307	146	121	181	279
2008	324	246	57	59	306	317
2009	252	312	78	100	272	319
2010	401	374	71	96	315	262
2011	448	354	68	119	247	304
2012	297	444	94	158	190	354
2013	414	380	104	145	369	358
2014	350	255	104	124	257	208
2015	324	343	165	178	310	392
2016	480	381	312	216	373	355
2017	438	367	292	284	383	419
2018	313	71	247	48	363	74
2019	297	185	148	65	316	137

Table 2.6. Landings at age of American plaice (in thousands of fish).

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980			22	770	3129	3903	3629	1185	1139	850	1380
1981		587	1332	4332	5101	3619	2381	1574	645	440	621
1982		113	2134	3495	4296	3482	3293	2038	1256	737	718
1983		1	437	3732	4267	3807	2251	1270	697	449	911
1984		3	253	1297.5	4819	2865	1913	577	274	307	769
1985			60	786	2066	2787	2213	1081	438	267	182
1986		1	198	1082	1502	1462	1307	631	255	105	99
1987		15	343	486	1703	1271	891	541	187	62	60
1988		1	446	1148	1456	1427	543	270	177	88	55
1989			76	451	686	504	749	469	193	103	116
1990			202	846	1049	500	290	349	193	96	161
1991			23	1862	2835	1112	321	165	202	98	105
1992			46	739	4872	2563	813	191	131	118	93
1993			123	1029	2037	2452	1382	265	287	151	125
1994			62	970	2296	1440	1162	681	286	163	132
1995			145	784	2949	1703	771	690	139	53	50
1996			159	2386	2531	1502	479	262	138	41	45
1997			3	1441	2769	1554	600	182	88	60	64
1998			15	278	1759	2437	1138	323	57	50	25
1999			2	295	1199	1858	1084	382	130	35	32
2000			142	442	1532	2189	1671	535	141	53	20
2001			13	376	1912	2103	1523	867	307	60	30
2002			3	333	1074	1608	1140	509	267	154	178
2003				165	814	1042	657	488	247	164	118
2004			8	170	504	781	456	351	216	77	80
2005			5	215	547	589	417	186	103	47	49
2006			1	141	443	428	311	206	104	56	43
2007			36	349	560	397	200	99	61	30	26
2008			31	173	614	455	289	178	95	71	62
2009				105	597	706	444	329	125	43	101
2010				87	311	846	590	270	163	111	95
2011				73	242	655	763	394	105	58	46
2012			10	80	248	366	626	685	323	84	96
2013			30	155	330	482	355	379	312	166	72
2014			54	305	409	240	509	179	162	128	132
2015			75	330	395	297	241	325	108	138	134
2016			38	190	330	287	188	105	156	62	173
2017			6	346	371	318	232	112	98	103	138
2018	0	2	6	70	574	367	219	180	76	64	197
2019	1	6	3	60	168	658	342	174	80	41	104

Table 2.7. Landed weight (kg) at age of American plaice.

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980			0.285	0.352	0.443	0.656	0.829	1.039	1.183	1.374	1.807
1981		0.133	0.207	0.343	0.454	0.783	0.891	0.982	1.130	1.254	1.620
1982		0.200	0.291	0.334	0.429	0.577	0.961	1.138	1.196	1.552	1.900
1983		0.184	0.341	0.460	0.582	0.682	0.828	1.043	1.244	1.446	1.816
1984		0.180	0.331	0.423	0.605	0.683	0.895	1.192	1.133	1.369	1.888
1985			0.221	0.270	0.362	0.544	0.852	1.167	1.377	1.665	2.199
1986		0.191	0.267	0.322	0.410	0.591	0.842	1.174	1.491	1.747	2.326
1987		0.201	0.284	0.386	0.475	0.627	0.895	1.177	1.483	1.732	2.427
1988		0.151	0.282	0.360	0.473	0.646	0.893	1.231	1.396	1.717	2.402
1989			0.339	0.393	0.489	0.586	0.739	0.858	1.334	1.463	1.954
1990			0.384	0.420	0.522	0.660	0.826	0.968	1.089	1.305	1.814
1991			0.333	0.453	0.543	0.715	0.963	1.161	1.276	1.541	2.129
1992			0.473	0.424	0.538	0.739	0.953	1.240	1.319	1.640	2.216
1993			0.416	0.451	0.518	0.649	0.945	1.234	1.394	1.577	2.498
1994			0.364	0.441	0.517	0.600	0.773	1.087	1.274	1.393	2.101
1995			0.355	0.412	0.518	0.648	0.720	1.023	1.279	1.643	1.695
1996			0.385	0.408	0.511	0.657	0.879	1.115	1.313	1.558	2.036
1997			0.292	0.419	0.470	0.634	0.810	1.082	1.300	1.457	2.007
1998			0.322	0.392	0.429	0.548	0.786	1.018	1.348	1.417	2.442
1999			0.255	0.419	0.471	0.564	0.718	0.922	1.122	1.356	1.499
2000			0.406	0.407	0.480	0.566	0.729	0.944	1.095	1.322	1.827
2001			0.509	0.428	0.453	0.536	0.699	0.820	1.054	1.416	1.975
2002			0.292	0.402	0.457	0.520	0.670	0.883	1.046	1.233	1.615
2003				0.397	0.438	0.549	0.689	0.851	0.988	1.064	1.599
2004			0.339	0.434	0.468	0.554	0.682	0.780	0.953	1.034	1.585
2005			0.427	0.406	0.476	0.567	0.685	0.862	0.936	1.074	1.620
2006			0.322	0.412	0.472	0.567	0.680	0.809	0.911	0.994	1.520
2007			0.390	0.408	0.468	0.560	0.681	0.837	0.979	1.102	1.790
2008			0.397	0.437	0.475	0.519	0.617	0.625	0.741	0.808	1.246
2009				0.448	0.474	0.538	0.570	0.594	0.720	0.887	1.133
2010				0.458	0.505	0.520	0.587	0.635	0.629	0.589	1.044
2011				0.466	0.531	0.559	0.571	0.630	0.687	0.715	1.084
2012			0.467	0.440	0.475	0.545	0.576	0.589	0.648	0.671	1.060
2013			0.371	0.441	0.489	0.532	0.589	0.608	0.635	0.629	1.140
2014			0.399	0.441	0.503	0.563	0.585	0.668	0.705	0.793	1.078
2015			0.408	0.473	0.531	0.586	0.631	0.648	0.803	0.729	0.952
2016			0.417	0.486	0.549	0.638	0.693	0.741	0.765	0.890	0.930
2017			0.434	0.458	0.548	0.637	0.707	0.793	0.843	0.863	1.186
2018	0.184	0.204	0.445	0.501	0.602	0.691	0.656	0.787	0.744	0.778	0.934
2019	0.075	0.165	0.417	0.492	0.554	0.584	0.662	0.750	0.774	0.755	1.092

Age Composition of Commercial Discards

The time series of discarded catch was re-estimated for this assessment by half-year and the fleets used for the standardized bycatch reporting method (Wigley et al. 2006). Size distribution of discards was sampled at sea (Tables 2.8-2.9) and varied by fleet. Sampling intensity was less than 100 mt per length sample in most years. Discarded size distributions from the large-mesh trawl and gillnet fleets were similar (Figure 2.13), with most discards slightly smaller than the legal size (increased from 12 in, 30.5 cm, to 14 in, 35.6 cm, in 2013; Appendix B). Discards from the shrimp fleet and other small-mesh trawl fisheries were smaller, and discards from the scallop fishery were larger, including legal sizes.

Age composition of discards was derived using survey age-length keys by half-year. Half-year periods with less than 100 lengths sampled were characterized by samples in the same half-year, fleet and decade. Age composition of discarded American plaice are dominated by ages 2-5 (Table 2.10 and Figure 2.14). Mean weights at age of discards reflect smaller fish discarded in the 1980s and relatively stable weight at age since then (Table 2.11).

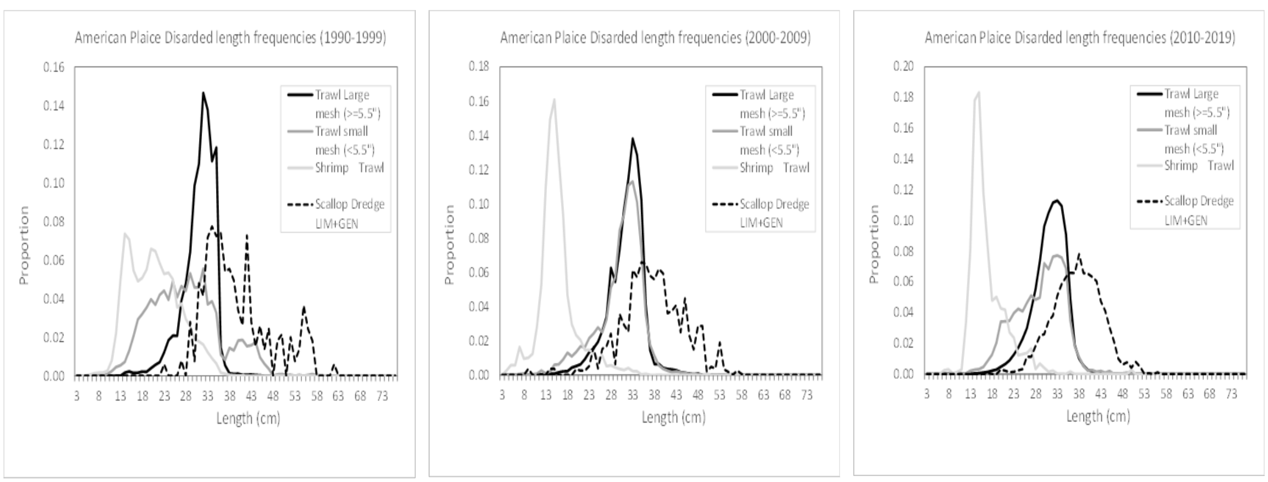


Figure 2.13. Length frequency of discarded American plaice by fleet and decade.

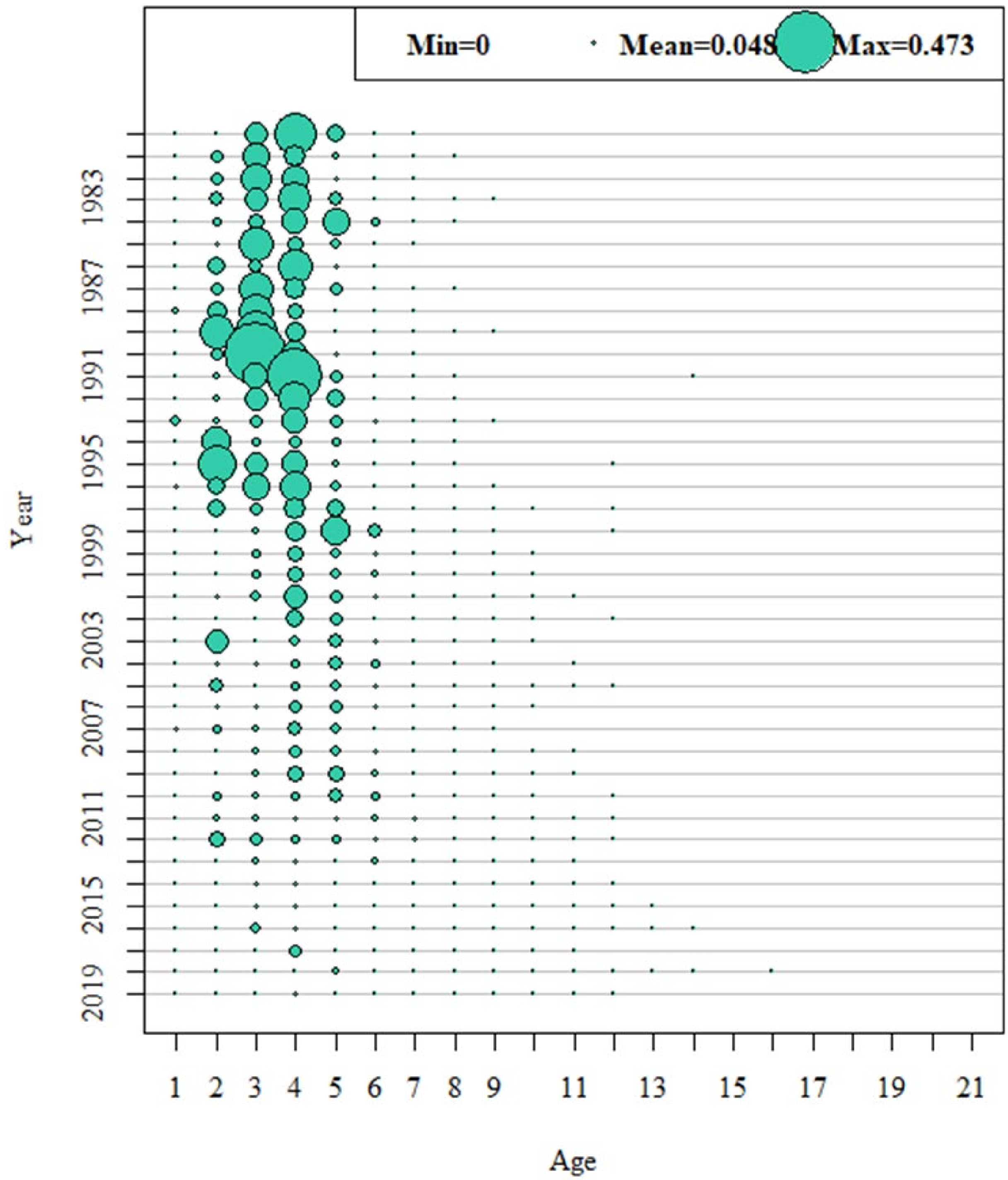


Figure 2.14. Age composition of American plaice discards.

Table 2.8. Number of tows or sets sampled for American plaice discard lengths.

	Large-Mesh Trawl		Small-Mesh Trawl		Shrimp Trawl		Gillnet		Scallop Dredge	
Year	Half1	Half2	Half1	Half2	Half1	Half2	Half1	Half2	Half1	Half2
1989	253	426	75	337	439	47		3		
1990	98	108		93	361	35	96	21		
1991	144	319	15	29	487	29	72	24		
1992	151	44	10	20	424	81	229	50		
1993	89	61	9		763	18	414	5	1	1
1994	90	80			425	43	42	35	4	
1995	368	220	63	546	474	117	84	34	1	1
1996	81	10	87	851	351	62	95	26		1
1997	77	53	23		184		32	3	16	
1998	81	8	7	3			75	10	22	12
1999		118		61			44	12	7	39
2000	68	52	188	43			41	16	3	1
2001	22	101	50	15			50	1		
2002	86	684	1	607			28	9		1
2003	774	1044	202	302	110		161	17		24
2004	597	1249	234	998	129		181	137	11	40
2005	2334	2354	284	670	125		39	49	35	50
2006	1456	1036	118	123	90	61	9		17	7
2007	1304	1485	111	270	69		38	10	22	10
2008	1739	1702	59	95	130	39	29	31	71	13
2009	1452	1283	60	248			57	7	112	54
2010	1069	485	18	100	5	25	45	21		1
2011	1143	897	61	296			60	29	1	10
2012	1119	986	81	217	72		94	37	47	19
2013	666	525	35	271	65		42	14	163	28
2014	676	233	41	102			37	4	41	15
2015	277	361	16	66			11	14	117	29
2016	73	114	3	42			3		106	
2017	130	289	21	292			3	2	41	3
2018	486	111	18	109			15	2	215	10
2019	308	533	12	22			7	3	75	9

Table 2.9 Number of American plaice discard lengths sampled.

	Large-Mesh Trawl		Small-Mesh Trawl		Shrimp Trawl		Gillnet		Scallop Dredge	
Year	Half1	Half2	Half1	Half2	Half1	Half2	Half1	Half2	Half1	Half2
1989	1575	3068	829	2651	4938	639		3		
1990	671	2104		823	3479	137	475	91		
1991	1931	3892	40	111	3073	164	191	43		
1992	1250	344	79	170	1847	253	986	64		
1993	931	624	25		3064	104	1589	7	1	1
1994	188	1580			2055	339	159	74	5	
1995	2767	1302	597	2763	3516	556	395	52	1	1
1996	513	33	430	6145	1970	254	308	55		1
1997	510	132	76		802		123	7	18	
1998	359	13	11	4			256	10	37	16
1999		495		112			166	20	7	57
2000	186	180	760	210			161	18	4	1
2001	61	447	146	40			202	1		
2002	464	3220	1	2303			83	11		1
2003	2613	4603	767	1069	860		489	25		39
2004	1863	5356	808	4400	606		391	270	18	75
2005	7881	11184	931	2828	728		51	69	61	105
2006	5547	5310	459	617	333	265	9		25	7
2007	5626	7149	306	870	400		78	14	35	10
2008	7192	9865	306	184	457	149	56	40	109	14
2009	5657	5968	178	885			89	7	192	136
2010	4887	2334	99	386	10	62	110	25		1
2011	4176	3367	151	1132			88	32	1	11
2012	3708	3534	416	730	484		162	40	63	20
2013	2600	1635	78	1301	347		77	16	391	29
2014	1915	508	73	181			62	5	62	15
2015	673	716	24	98			14	17	231	36
2016	127	338	4	78			3		234	
2017	477	807	56	678			3	2	46	3
2018	1547	354	43	170			44	2	536	10
2019	724	1439	14	26			10	3	90	10

Table 2.10. American plaice discards at age (in thousands of fish).

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	5	99	1050	1902	810	30	3				
1981	5	395	860	724	237	30	20	8			
1982	10	490	1214	1079	208	117	4				
1983	15	662	1041	1442	648	104	18	1	4		
1984	3	367	737	1125	1213	380	22	3			
1985	65	158	1158	551	339	85	15				
1986	59	638	540	1202	198	14	0	0			
1987	38	575	1498	953	580	66	5	2			
1988	314	785	1393	685	141	18	10				
1989	27	2618	3140	1539	305	95	16	6	7		
1990	94	1098	5375	2216	495	78	12	0			
1991	3	430	1712	3505	808	52	5	2	0	0	1
1992	25	347	1022	1494	822	86	21	11	0		
1993	405	243	477	877	437	168	25	6	8	0	
1994	124	881	291	399	268	48	14	1	0	0	0
1995	149	2307	1368	1452	390	54	6	7	0	0	1
1996	265	1003	1494	1751	574	220	57	10	2	0	
1997	62	710	489	849	671	123	36	10	3	4	2
1998	6	33	284	704	1065	534	55	21	2	0	5
1999	4	57	196	385	256	135	45	15	5	3	
2000	1	39	234	388	247	174	61	20	9	3	
2001	14	178	336	708	379	138	74	15	11	2	7
2002	2	17	86	410	271	69	28	15	7	3	3
2003	40	684	120	305	409	152	49	22	16	3	0
2004	10	141	108	222	326	195	42	12	9	0	1
2005	39	360	101	231	277	145	32	19	3	2	2
2006	58	110	120	286	301	137	99	20	2	4	0
2007	136	221	179	347	271	111	30	6	4	0	0
2008	21	56	169	293	229	126	50	10	9	2	4
2009	49	120	207	445	481	218	91	32	19	10	1
2010	103	248	183	228	405	264	110	30	22	10	5
2011	44	190	162	136	139	191	125	82	26	6	11
2012	7	475	378	272	276	171	138	109	46	7	4
2013	17	95	137	122	93	154	38	34	37	13	5
2014	9	44	100	97	50	26	45	10	14	11	9
2015	2	77	100	85	53	21	10	14	4	2	4
2016	1	11	218	97	43	18	5	6	7	4	4
2017	1	31	32	219	70	24	12	5	1	3	2
2018	3	6	60	33	149	56	24	14	7	5	10
2019	2	33	16	96	40	73	17	9	3	2	3

Table 2.11. Discarded weight (kg) at age of American plaice.

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	0.030	0.076	0.152	0.233	0.277	0.312	0.392				
1981	0.032	0.070	0.109	0.157	0.177	0.191	0.171	0.219			
1982	0.018	0.095	0.122	0.146	0.190	0.178	0.239				
1983	0.013	0.033	0.119	0.165	0.190	0.211	0.215	0.310	0.240		
1984	0.004	0.044	0.103	0.165	0.203	0.234	0.285	0.257			
1985	0.018	0.058	0.077	0.123	0.145	0.208	0.196				
1986	0.016	0.042	0.091	0.145	0.185	0.194	0.295	0.359			
1987	0.013	0.042	0.096	0.157	0.216	0.278	0.297	0.116			
1988	0.016	0.046	0.120	0.158	0.193	0.256	0.212				
1989	0.012	0.047	0.149	0.249	0.252	0.259	0.415	0.668	0.746	0.000	
1990	0.017	0.048	0.132	0.222	0.263	0.309	0.355	0.427	0.000	0.000	
1991	0.022	0.049	0.155	0.262	0.319	0.406	0.730	1.030	1.727	1.926	1.599
1992	0.019	0.045	0.115	0.259	0.309	0.350	0.384	0.329	0.951	0.000	1.697
1993	0.017	0.038	0.199	0.261	0.326	0.417	1.523	0.993	1.934	1.164	1.327
1994	0.015	0.027	0.217	0.271	0.317	0.321	0.318	1.114	1.379	1.651	1.681
1995	0.015	0.036	0.159	0.265	0.309	0.383	0.670	1.181	1.497	1.332	1.801
1996	0.019	0.031	0.093	0.213	0.302	0.404	0.430	1.024	1.448	1.801	1.801
1997	0.014	0.029	0.157	0.243	0.290	0.372	0.457	1.122	0.343	0.400	1.859
1998	0.040	0.068	0.159	0.232	0.266	0.330	0.420	1.174	0.741	1.659	1.197
1999	0.035	0.100	0.200	0.246	0.299	0.309	0.361	0.496	2.004	1.827	
2000	0.004	0.109	0.202	0.271	0.363	0.360	0.498	0.932	1.028	1.078	
2001	0.011	0.026	0.149	0.246	0.289	0.330	0.373	0.513	0.452	0.622	0.352
2002	0.021	0.074	0.141	0.230	0.284	0.297	0.345	0.367	0.453	0.663	1.055
2003	0.019	0.028	0.160	0.237	0.284	0.350	0.349	0.331	0.600	0.623	0.849
2004	0.012	0.025	0.082	0.205	0.291	0.335	0.424	0.576	0.733	1.594	1.127
2005	0.009	0.023	0.098	0.219	0.281	0.344	0.390	0.570	0.877	1.180	3.885
2006	0.015	0.040	0.124	0.218	0.281	0.327	0.340	0.505	0.720	0.419	1.327
2007	0.006	0.034	0.142	0.243	0.326	0.362	0.354	0.490	0.351	1.420	1.088
2008	0.009	0.044	0.162	0.238	0.280	0.313	0.366	0.495	0.468	0.790	0.497
2009	0.012	0.063	0.207	0.252	0.286	0.333	0.351	0.370	0.385	0.409	1.123
2010	0.029	0.045	0.134	0.238	0.275	0.292	0.307	0.325	0.353	0.324	0.594
2011	0.015	0.026	0.083	0.195	0.242	0.281	0.288	0.302	0.305	0.569	0.604
2012	0.017	0.033	0.074	0.199	0.240	0.274	0.283	0.314	0.346	0.337	0.697
2013	0.044	0.098	0.162	0.226	0.292	0.273	0.292	0.334	0.329	0.379	0.556
2014	0.039	0.128	0.197	0.220	0.256	0.295	0.280	0.343	0.309	0.314	0.519
2015	0.027	0.133	0.201	0.251	0.287	0.334	0.350	0.385	0.392	0.522	0.474
2016	0.036	0.131	0.224	0.281	0.317	0.364	0.484	0.453	0.465	0.415	0.508
2017	0.019	0.097	0.186	0.255	0.309	0.325	0.325	0.324	0.552	0.455	0.497
2018	0.035	0.080	0.180	0.230	0.315	0.340	0.363	0.439	0.468	0.445	0.565
2019	0.026	0.105	0.174	0.248	0.280	0.330	0.356	0.386	0.516	0.326	0.892

Age Composition of Total Catch

Age composition of fishery catch has been relatively stable since the 1980s, with ages 2-6 contributing most catch (Table 2.12, Figure 2.15). Several apparently strong year-classes (e.g., 1987, 1993, 2004, 2013) contributed to catch over several years. Cohort tracking is relatively strong in the total catch at age, as measured by positive correlations of catch-at-age by year-class from age-1 to age-10 and moderate to strongly positive correlations among adjacent years and ages (r=0.53-0.93, r>0.9 for ages 4-8; Figure 2.16). Mean weight-at-age of the catch was relatively stable for ages 1-6 but decreased over time for older ages since the 1990s (Table 2.13, Figure 2.17).

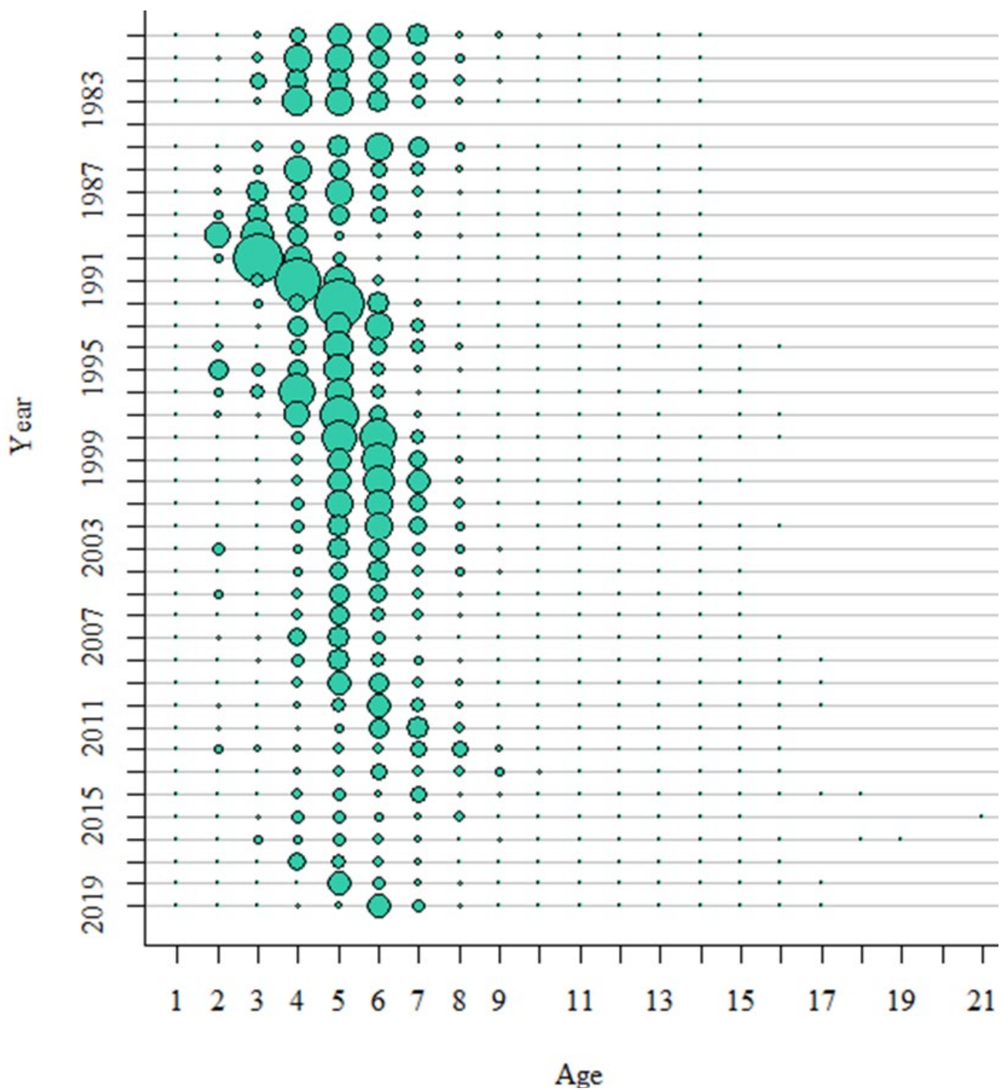


Figure 2.15. Age composition of American plaice catch.

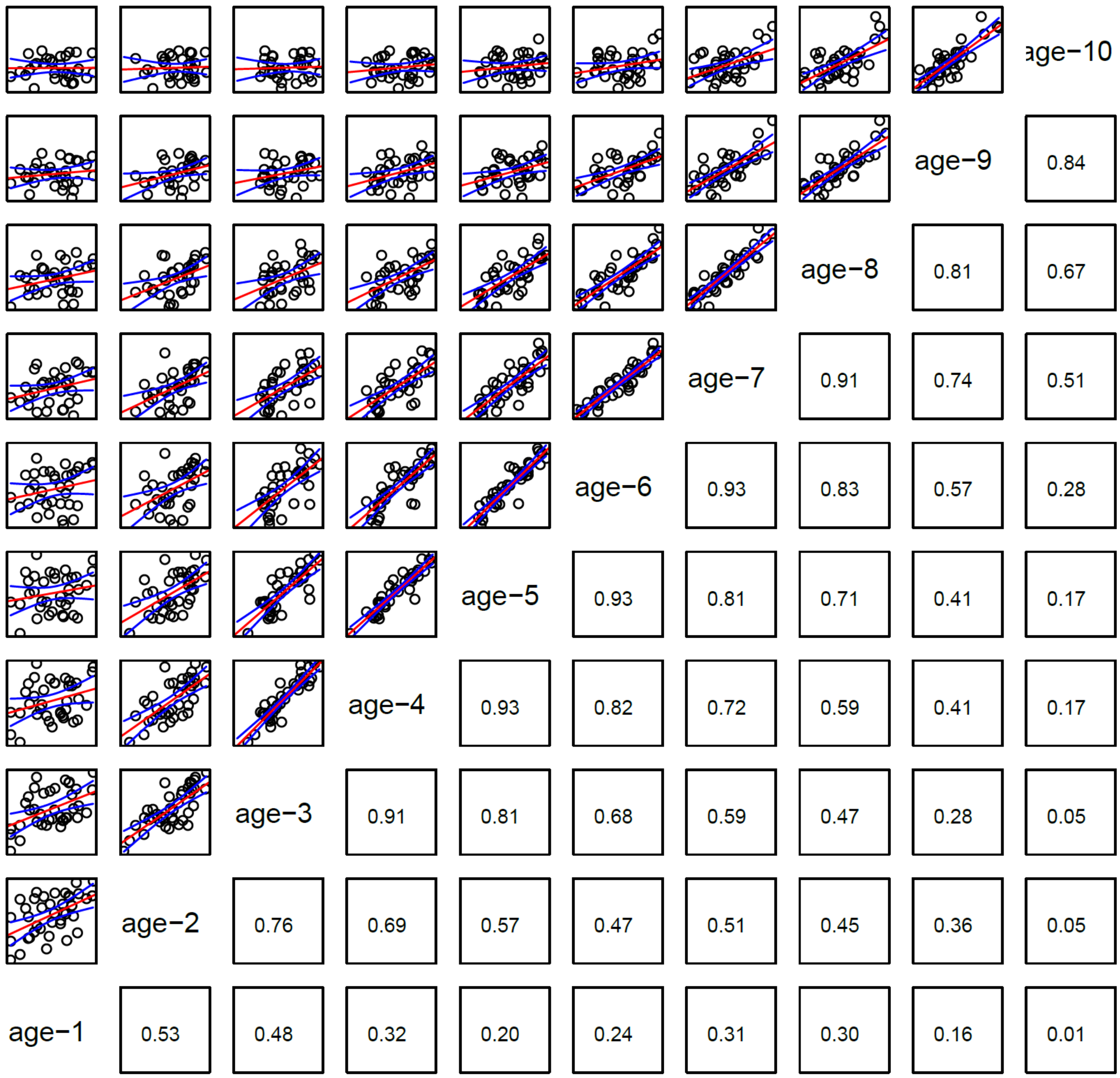


Figure 2.16. Correlations of American plaice catch-at-age by year-class over time.

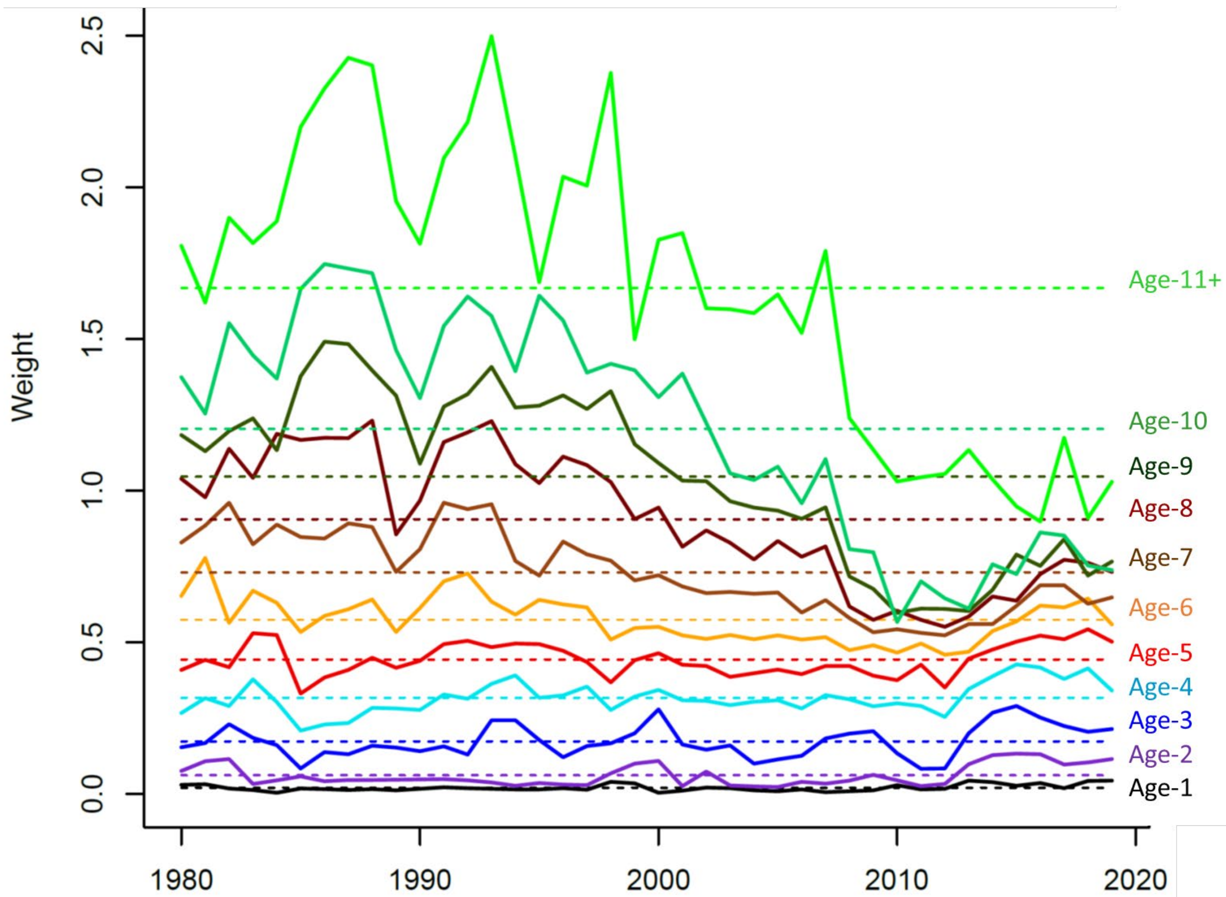


Figure 2.17. Weight at age (kg) of American plaice in the fishery.

Table 2.12. Total catch at age of American plaice in the fishery (in thousands of fish).

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	5	99	1072	2672	3939	3933	3632	1185	1139	850	1380
1981	5	982	2192	5056	5338	3649	2401	1582	645	440	621
1982	10	603	3349	4575	4504	3599	3298	2038	1256	737	718
1983	15	663	1478	5174	4915	3910	2269	1271	701	449	911
1984	3	370	991	2422	6031	3244	1936	580	274	307	769
1985	65	158	1217	1336	2405	2872	2228	1081	438	267	182
1986	59	639	738	2284	1700	1476	1307	631	255	105	100
1987	38	590	1840	1439	2282	1337	895	543	187	62	60
1988	314	786	1840	1833	1597	1444	553	270	177	88	55
1989	27	2618	3216	1991	991	598	765	476	200	103	116
1990	94	1098	5577	3062	1544	578	303	349	193	96	161
1991	3	430	1735	5367	3644	1164	326	167	203	98	105
1992	25	347	1068	2233	5694	2649	833	201	131	118	94
1993	405	243	600	1905	2474	2620	1407	271	294	152	125
1994	124	881	353	1369	2565	1488	1176	682	287	163	132
1995	149	2307	1514	2236	3339	1757	777	697	139	53	51
1996	265	1003	1653	4137	3105	1723	536	272	140	42	46
1997	62	710	492	2290	3441	1677	636	192	91	64	65
1998	6	33	298	981	2824	2970	1193	345	59	50	31
1999	4	57	197	680	1455	1993	1129	397	135	38	32
2000	1	39	376	830	1779	2363	1732	555	149	56	20
2001	14	178	349	1084	2291	2241	1597	881	319	62	37
2002	2	17	89	742	1345	1677	1168	524	274	157	181
2003	40	684	120	471	1223	1194	705	510	263	167	119
2004	10	141	116	391	830	977	498	362	226	77	81
2005	39	360	106	446	824	734	449	205	106	50	51
2006	58	110	121	427	744	564	410	226	106	60	43
2007	136	221	215	697	830	508	230	105	64	30	26
2008	21	56	200	466	843	581	340	187	104	73	66
2009	49	120	207	551	1078	924	535	360	144	53	102
2010	103	248	183	315	716	1109	700	300	184	121	100
2011	44	190	162	209	381	846	887	476	131	65	57
2012	7	475	387	353	524	538	764	794	369	91	100
2013	17	95	167	278	423	636	393	413	349	179	77
2014	9	44	154	402	459	266	553	189	176	139	141
2015	2	77	175	415	448	318	251	339	112	140	138
2016	1	11	256	287	373	305	193	111	162	65	177
2017	1	31	38	565	441	342	244	117	99	106	140
2018	4	8	66	103	722	423	243	194	83	69	207
2019	4	39	19	156	207	730	358	183	82	42	108

Table 2.13. Weight at age (kg) of American plaice in the fishery.

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	0.030	0.076	0.154	0.267	0.409	0.653	0.829	1.039	1.183	1.374	1.807
1981	0.032	0.108	0.168	0.316	0.442	0.778	0.885	0.978	1.130	1.254	1.620
1982	0.018	0.115	0.230	0.290	0.418	0.564	0.960	1.138	1.196	1.552	1.900
1983	0.013	0.033	0.185	0.378	0.530	0.670	0.823	1.042	1.238	1.446	1.816
1984	0.004	0.045	0.161	0.303	0.524	0.630	0.888	1.187	1.133	1.369	1.888
1985	0.018	0.058	0.084	0.209	0.331	0.534	0.847	1.167	1.377	1.665	2.199
1986	0.016	0.042	0.138	0.229	0.384	0.587	0.842	1.174	1.491	1.747	2.326
1987	0.013	0.046	0.131	0.234	0.409	0.609	0.892	1.173	1.483	1.732	2.427
1988	0.016	0.046	0.159	0.284	0.449	0.641	0.880	1.231	1.396	1.717	2.402
1989	0.012	0.047	0.153	0.282	0.416	0.534	0.732	0.856	1.313	1.463	1.954
1990	0.017	0.048	0.141	0.277	0.439	0.613	0.807	0.967	1.089	1.305	1.814
1991	0.022	0.049	0.157	0.328	0.494	0.701	0.960	1.160	1.277	1.543	2.096
1992	0.019	0.045	0.130	0.314	0.505	0.727	0.939	1.192	1.318	1.640	2.215
1993	0.017	0.038	0.243	0.363	0.484	0.634	0.955	1.229	1.408	1.576	2.498
1994	0.015	0.027	0.243	0.391	0.496	0.591	0.768	1.087	1.274	1.394	2.101
1995	0.015	0.036	0.178	0.317	0.494	0.640	0.720	1.025	1.280	1.643	1.687
1996	0.019	0.031	0.121	0.325	0.472	0.625	0.832	1.112	1.314	1.561	2.035
1997	0.014	0.029	0.158	0.354	0.435	0.615	0.790	1.084	1.269	1.389	2.005
1998	0.040	0.068	0.167	0.277	0.368	0.509	0.769	1.028	1.328	1.418	2.377
1999	0.035	0.100	0.200	0.321	0.441	0.547	0.704	0.906	1.153	1.397	1.499
2000	0.004	0.109	0.279	0.343	0.464	0.551	0.721	0.944	1.091	1.308	1.827
2001	0.011	0.026	0.163	0.309	0.426	0.523	0.684	0.815	1.033	1.386	1.849
2002	0.021	0.074	0.146	0.307	0.422	0.511	0.662	0.869	1.031	1.223	1.601
2003	0.019	0.028	0.160	0.293	0.386	0.524	0.666	0.828	0.965	1.057	1.598
2004	0.012	0.025	0.100	0.304	0.398	0.510	0.660	0.773	0.944	1.035	1.585
2005	0.009	0.023	0.114	0.309	0.410	0.523	0.664	0.834	0.934	1.079	1.647
2006	0.015	0.040	0.126	0.282	0.395	0.509	0.598	0.782	0.907	0.958	1.520
2007	0.006	0.034	0.183	0.326	0.422	0.517	0.639	0.816	0.945	1.104	1.790
2008	0.009	0.044	0.199	0.312	0.422	0.474	0.580	0.618	0.717	0.807	1.239
2009	0.012	0.063	0.207	0.289	0.390	0.490	0.533	0.574	0.676	0.797	1.135
2010	0.029	0.045	0.134	0.299	0.375	0.466	0.543	0.604	0.597	0.567	1.030
2011	0.015	0.026	0.083	0.290	0.426	0.496	0.531	0.574	0.611	0.701	1.044
2012	0.017	0.033	0.084	0.254	0.351	0.459	0.523	0.551	0.610	0.645	1.055
2013	0.044	0.098	0.200	0.346	0.446	0.469	0.560	0.585	0.603	0.611	1.134
2014	0.039	0.128	0.268	0.388	0.476	0.537	0.560	0.651	0.673	0.757	1.037
2015	0.027	0.133	0.290	0.427	0.502	0.569	0.620	0.637	0.789	0.725	0.948
2016	0.036	0.131	0.252	0.417	0.522	0.621	0.688	0.725	0.752	0.862	0.898
2017	0.019	0.097	0.224	0.379	0.510	0.615	0.688	0.772	0.840	0.852	1.174
2018	0.043	0.104	0.205	0.414	0.543	0.644	0.627	0.762	0.720	0.753	0.910
2019	0.044	0.115	0.214	0.341	0.502	0.559	0.648	0.732	0.766	0.738	1.029

Fishery Catch Rates

Early stock assessments of American plaice included nominal catch per unit effort (CPUE, Rounsefell 1957, Lange and Lux 1978; NEFC 1986; NEFSC 1992, 1999ab; O'Brien et al. 1999) but fishery catch rates have not been used as an index of abundance in stock assessment models (NEFSC 1992, 1999ab; O'Brien et al. 1999). In a review of Northeast fishery stock assessments, the National Research Council concluded that "fishers have a greater trust in the data that they themselves provide, and therefore an effort should be made to validate and use CPUE data" (NRC 1998). Considering CPUE in the stock assessment process and documentation can be valuable for providing fishery data with greater spatial and temporal resolution than fishery-independent surveys and understanding fishery dynamics (Cadrin et al. 2020).

Several alternative series of standardized LPUE (landings per unit effort) from seafood dealer reports and fishermen logbooks (Figure 2.17; Grezlik et al. 2021, Working Paper 6 Appendix A) and CPUE from at-sea observers and a study fleet were considered from a variety of statistical methods from conventional generalized linear models (GLMs, Gavaris 1980, Hankowsky et al. 2021, Working Paper 7; Terceiro 2021, Working paper 9), generalized additive models (Maunder and Punt 2004; Jones 2021, Working Paper 8) to spatiotemporal models (Anderson et al. 2019; Jones 2021, Working Paper 8). Results from all approaches and sensitivity analyses within approaches produced similar time series that were moderately to strongly correlated with survey biomass indices (r= 0.58 to 0.73), indicating a general decrease in stock size from the 1990s to mid-2000s, a general increase from 2005 to 2017 and a recent decrease (Figure 2.18). Based on model diagnostics, general agreement with survey trends, and robustness of signals to alternative explorations, the Working Group decided to explore stock assessment models that include standardized fishery catch rates as an index of stock size. Criteria for considering abundance indices (Appendix D) indicated that the revised GLM of dealer-logbook LPUE with vessel tonnage class, statistical area, quarter-year, depth and price factors (GLM-2) and the spatio-temporal model were the best candidates to consider as indices of abundance.

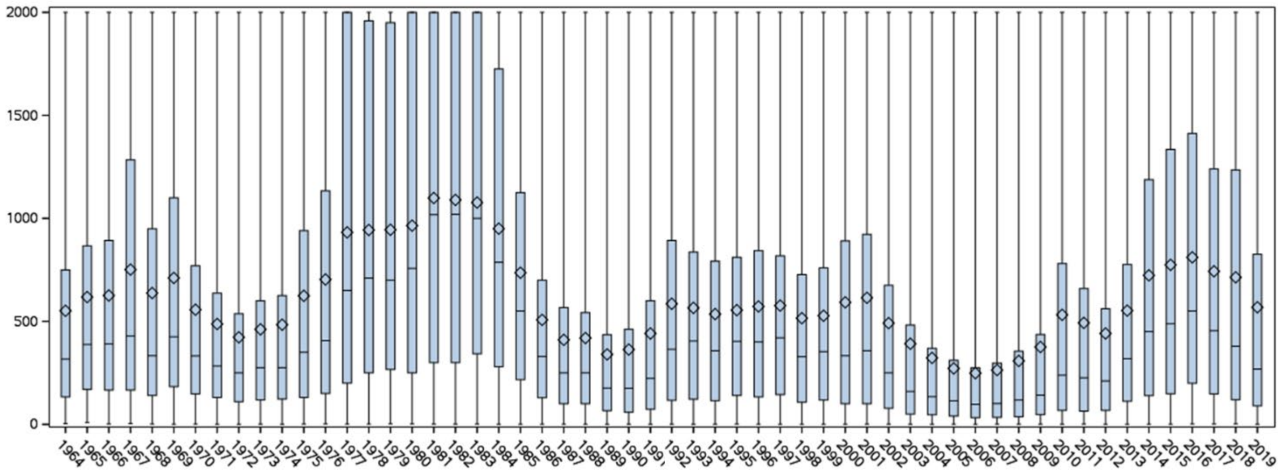


Figure 2.17. LPUE of American plaice from dealer data (diamond: mean; box plot of median, and interquartile range).

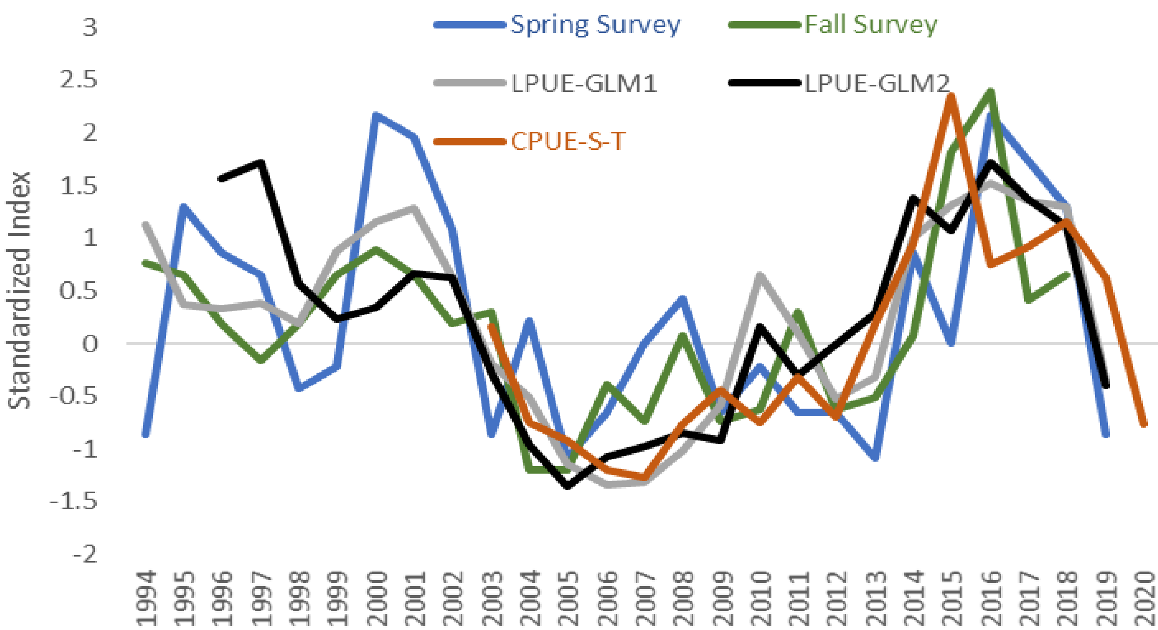


Figure 218. Time series of standardized dealer-logbook LPUE of American plaice from a conventional GLMs, observer-study fleet CPUE from spatiotemporal and spatial standardization models, and trawl survey indices of biomass.

Table 2.14. Standardization parameter estimates and retransformed LPUE index for American plaice from GLM-2.

Parameter	Estimate	Std. Error	Index	CV
(Intercept)	1.323	0.061	3.755	0.061
YEAR1997	0.067	0.019	4.016	0.064
YEAR1998	-0.167	0.019	3.177	0.064
YEAR1999	-0.207	0.019	3.052	0.064
YEAR2000	-0.262	0.019	2.89	0.064
YEAR2001	-0.23	0.022	2.983	0.065
YEAR2002	-0.282	0.026	2.832	0.066
YEAR2003	-0.62	0.029	2.021	0.067
YEAR2004	-0.884	0.027	1.552	0.067
YEAR2005	-1.096	0.026	1.255	0.066
YEAR2006	-0.935	0.029	1.475	0.068
YEAR2007	-0.788	0.029	1.707	0.068
YEAR2008	-0.789	0.029	1.706	0.068
YEAR2009	-0.789	0.031	1.706	0.069
YEAR2010	-0.243	0.034	2.946	0.07
YEAR2011	-0.381	0.029	2.566	0.068
YEAR2012	-0.362	0.027	2.616	0.067
YEAR2013	-0.156	0.028	3.212	0.067
YEAR2014	0.263	0.025	4.883	0.066
YEAR2015	0.305	0.023	5.097	0.065
YEAR2016	0.531	0.023	6.383	0.065
YEAR2017	0.403	0.022	5.618	0.065
YEAR2018	0.302	0.021	5.08	0.065
YEAR2019	-0.152	0.021	3.225	0.065
NEMAREA512	0.145	0.059		
NEMAREA513	0.351	0.055		
NEMAREA514	0.112	0.055		
NEMAREA515	0.084	0.055		
NEMAREA521	-0.081	0.055		
NEMAREA522	-0.405	0.056		
NEMAREA525	-0.572	0.064		
NEMAREA561	-0.632	0.062		
QTR2	0.548	0.011		
QTR3	0.096	0.012		
QTR4	-0.293	0.013		
TON_CLASS2	0.378	0.012		
TON_CLASS3	0.458	0.016		
TON_CLASS4	0.476	0.016		
DEPTH	0.011	0		
price_2019	-0.338	0.005		

Fishermen's Perspectives

The American Plaice Research Track Assessment Working Group and the Northeast Seafood Coalition held a Fishermen's Meeting in Gloucester, MA on September 24, 2021. As a requested follow-up, Working Group members met with fishermen November 3, 2021, in Portland ME. The purpose of the meetings was to solicit input from fishing industry members to consider in support of the American Plaice Research Track Assessment. The meeting included presentations about the American Plaice Research Track Assessment, an overview of the Research Track Assessment process, information specific to the American Plaice assessment, including fishery independent and dependent data sources, biological information, and results from previous American Plaice assessments. An interactive dashboard application was also demonstrated to interactively examine survey and fishery catch data for American plaice. Meeting attendees participated in discussions about American plaice survey and fishery data, management measures, biological traits, and environmental factors influencing American plaice distribution and biology. Working Group members also reached out to plaice fishermen individually to elicit their local ecological knowledge (Pavlovich et al. 2021, Working Paper 5 Appendix A).

Impacts on American Plaice Landings from Management Measures

Industry members emphasized that the declining catch of plaice is not reflective of declining biomass, instead it is representative of the increasing regulatory measures that have prevented targeting the stock in specific areas at specific times of the year. For example, in the 1980s, there were fewer regulations, more vessels and larger vessels, and the size distribution of dabs was larger. The Gulf of Maine cod rolling closures implemented in 2015 in the inshore areas west of the Western Gulf of Maine closure have limited the fleet's ability to target plaice in May and June when plaice are present in the inshore region.

Implementation of the Annual Catch Limit and sector management system in 2010 resulted decreased otter trawling, which allowed increased fixed gear lobster fishing in recent years in areas traditionally trawled, ultimately creating untrawlable areas because of fixed gear, specifically mud bottoms in the inshore region off Cape Ann and west of the Western Gulf of

Maine closure. After Annual Catch Limits decreased in 2013, and lease prices increased in 2014 and 2015, relatively low dab allocations constrained the groundfish fishery.

Increased minimum mesh size regulations have impacted ability to catch plaice. Additionally, many vessels have switched to diamond mesh instead of square to target haddock. Plaice are not retained as well in diamond mesh, and there are few sub-legal sized fish retained. Data from the Electronic Monitoring Maximized Retention program shows low discards of plaice due to mesh size regulations.

American Plaice Distribution Changes and Catch Rates

Fishermen indicated that plaice distribution patterns are currently different than observed in the 1970s and 1980s. Dabs traditionally came inshore in spring, but they are now not caught inshore. All flatfish have been observed moving to deeper water in recent years. Plaice abundance, size, and age have been impacted by changes in water temperature.

Industry members suggested that examining Catch per Unit Effort (CPUE) indices for plaice may be useful to consider in the research track assessment. As indicators of plaice targeting, they suggested criteria based on proportion of plaice in total catch (e.g., >1/3 of total catch), mostly flatfish and monkfish, mesh configuration (i.e., square vs. diamond), gear characteristics (e.g., flat net and sweep vs. rock hoppers), and tow speed (2.5-2.6 knots for flatfish vs. 2.7-2.9 knots for cod and haddock). They suggested examination of CPUE index from the region adjacent to the Cashes Ledge Closure, and efforts that were initiated to determine a plaice fishery footprint.

Survey Catches of Plaice

Industry members noted that fishery-independent surveys have low sampling intensity, lack spatial coverage in areas where plaice occur, and inefficient survey gear to sample plaice. The Northeast Fisheries Science Center Bottom Trawl Survey and MA Division of Marine Fisheries Trawl Survey cannot adequately sample inshore areas due to fixed gear. Lobster gear has prohibited surveys from sampling inshore areas, and it is unclear if plaice are still present in these areas. The NEFSC survey has had low sampling rates in the region west of the Cashes

Ledge Closure where the fishery has had consistent catch rates of ~5,000 pounds/day over the last five years. The NEFSC survey gear is not effective for catching flatfish species and there is low survey catch efficiency for plaice and other flounder species.

Fishermen's Ecological Knowledge

Pavlovich et al. (2021, Working Paper 5 Appendix A) engaged with the fishing industry to use their insights to generate hypotheses about climate impacts on flatfish populations specifically and identify other important issues that may have been unresolved or overlooked by scientists to date. Between 2018 and 2020, they reached out to fishermen known to participate in flatfish fisheries in the northeast, primarily those who operate in the Gulf of Maine. In total, they had personal conversations with twelve fishermen and one fish processor. Ten fishermen and two sector managers attended workshops. Conversations with fishing industry members directed analyses of scientific survey data and landings data. Fishermen presented ideas on five topics (migration, abundance, habitat, distribution, fisherman and fishing fleet decision-making), and two themes emerged: changes over time and subregional differences within the Gulf of Maine/Georges Bank region.

Fishermen spoke regularly about plaice migration as one of the most important factors that influence when and where they fish and how good of a fishing season they have. None of the fishermen interviewed fishes across the entire or even a large proportion of the distribution of plaice, so observations are constrained to the space-time combination in which each fishes. Fishermen shared many insights about the timing and final location of the inshore migration. In general terms, plaice tend to spend the winter in deeper basins, then move closer to shore during late spring and early summer to feed and spawn. In the fall, they move back to the deeper wintering grounds. Not all plaice follow this migration pattern. There are some areas where plaice can be caught all year round, such as around Cashes Ledge (Figure 2), while others only have plaice during certain times of the year, such as Massachusetts Bay. The timing of migration is thought to vary by as much as a month from year to year, perhaps dependent upon the nature of the progression of spring and summer, though not all fishermen had a clear hypothesis of the drivers of the timing. Several fishermen said that the inshore migration has been delayed in the past decade by as much as four to six weeks. Fish used to come inshore as early as March or

April, and now they arrive in May or June. In 2018, the Western Gulf was slow to warm which was blamed for fish arriving later than usual. The timing of the offshore migration has also been delayed, reportedly by as much as two or three months. While they used to head deeper in July and August, now they sometimes stay as late as November. Some fishermen stated that plaice and grey sole probably swim in the water column during migration. This would explain why plaice are difficult to catch prior to arriving to their inshore grounds and why they are not caught in bottom-tending gill nets on Jeffrey's Ledge, after which they are easy to catch. If this hypothesis is true, it would have implications for capture efficiency of the NEFSC spring and fall trawl surveys.

Fishermen agreed that American plaice are highly abundant in relation to abundance in the past. One fisherman commented that his catch per unit effort in 2018 and 2019 was as high as it had been during the best years his father and grandfather fished in the early 1980s. After the 1980s, he saw CPUE decline and remain low until around 2010, after which it has increased continually. Others stated that groundfish in general are much higher than in the past. Several fishermen also reported that the abundance they see is high compared to the quota they are allocated. Fishermen said they must avoid areas they know have a high density of plaice because they will catch their quota too quickly.

Plaice distribution was most often described in terms of the depths at which fishermen caught fish which depends on where fish are in their annual migration cycle. One fisherman said that fish used to come inshore to around 20 fathoms (37 m) in the 1960s and 70s, followed by several decades when they were not found that far inshore and have since returned to those depths since around 2010. Another repeated the same sentiment that plaice are now being found farther inshore than they were in the past. At the same time, several other fishermen said that fish are being found in deeper water, farther from shore. Fishermen said they believe plaice are found in more areas now that their population has increased compared to previous decades. One fisherman commented that the Massachusetts Division of Marine Fisheries bottom-trawl survey does not sample plaice well anymore because they are not migrating inshore to state waters as much as they did in the past.

A major challenge to understanding and describing plaice habitat is the annual migration some portion of the population undertakes for several reasons. Plaice use at least two different types of habitat, wintering and summer grounds, as well as occurring in between those habitats during migration. Furthermore, the transition between winter and summer grounds takes place during the spring and fall, precisely when the Northeast Fisheries Science Center bottom trawl surveys are sampling. Finally, fishermen capture plaice at different times of year, depending on where they fish, making each fisherman's conception of habitat somewhat different from that of other fishermen. The environmental variables fishermen spoke most frequently about were sediment, depth, and temperature. They said plaice prefer muddy, sandy sediments, as opposed to gravely or hardbottom bottom types. They occur over a range of depths and temperature, but that the fish will only be in specific areas at one time. It was stated that half of one degree Fahrenheit or half of a mile makes a big difference in finding fish, but that the specific temperature and location varies. Fishermen also commented that plaice spatially segregate based on size, too, with large fish concentrated at certain depths. Fishermen have started to find fish in deeper water than they have in the past. This goes for both plaice overwintering in deep basins and not moving as far inshore during their spring migrations. This is reportedly also true for other species of flatfish, especially winter flounder, and most groundfish in general. One of the most common concerns fishermen raised was whether the trawl survey adequately sampled quality plaice habitat.

There was consensus among fishermen that some fish stocks have relatively low quotas for the stock size they perceive. This leads to fishermen in the multispecies groundfish fishery to actively avoid many species for which the quota is low, or their allocation of the quota is low, compared to how much fish they could catch if unrestricted. These are called "choke" stocks or species because they constrain where and when fishermen fish or can even shutdown fishing for some. For several fishermen we spoke with, plaice has recently been a choke species. The most frequently mentioned choke species is cod. One sector manager suggested that lease price is a good metric of the disparity between true stock size and the quota allowed for the estimated population size in the stock assessment. In the context of a multispecies fishery where fishermen have some but not perfect control over which species they catch, fishermen often have to lease quota from other fishermen (or lease holding inactive fishermen) if they catch too much of a choke species for which they do not have enough quota. Under this system, the price of the quota can be less reflective of the demand for a species in the seafood market and more reflective of the relationship between true and assessed abundance. Several fishermen gave anecdotes of having particularly large hauls for which they had to purchase additional quota and ended up losing money on the fishing trip. According to fishermen, plaice only became a problem for fishermen in the past two to three years (2016/2017). Abundance was lower before that, but lease prices were lower, too. Now, abundance is higher, but lease prices are also higher.

Fishermen reported that there has been a large expansion of the area of the Gulf of Maine that lobster fishermen fish since the early 2000s. The lobster pots in the water preclude trawl fishermen from areas that were historically important for groundfishing, including for plaice. This has forced fishermen to find other places to fish or, as some have done, switch to lobstering. Furthermore, there is speculation that the NEFSC bottom trawl survey must also be excluded from these areas where plaice have historically been very abundant during the late spring and summer. In addition to losing fishing grounds, one fisherman also said he believes lobster fishing is detrimental to groundfish habitat because the sediment becomes stagnant when not turned over by bottom trawls.

All fishermen recounted a similar trajectory of the groundfish fleet and fishing effort over the past several decades. In the 1960s, 70s, and early 80s there were many more boats that were actively fishing, and plaice catches were good. However, starting in the early to mid-1980s, plaice abundance began to decrease while fishing effort stayed the same, especially in nearshore waters. When the "days-at-sea" management regime was put into effect, an attempt to limit catch by limiting the amount of time that could be fished, fishermen had to focus their efforts by targeting the species that were most valuable in the areas that were most reliable. When the catch shares ("sectors") regime began in 2010, the pressure to catch as much fish as possible in a short amount of time was relieved, but the challenge of negotiating quota, allocation, and leasing came into play. There has been a tremendous decline in the number of boats fishing since the 1980s, and of those that are active, a small number make up the majority of the catch. Whereas there used to be several hundred boats pursuing groundfish, fishing industry representatives stated that 41 boats catch over 60% of the landings, and 60 boats comprise 75% of landings. With the number of boats and fishermen active in the groundfish fishery, fishermen generally believe that any current issues with fish stocks must be caused by environmental changes, pollution, or natural cycles, not overfishing.

Multiple fishermen commented that fishermen generally fish differently when they are accompanied by an observer than when they are not. They suggested that they choose locations with very low chances of accidentally catching a species they do not have enough quota for. Also, the duration of trips is sometimes shorter than it would have been otherwise. This supports other work suggesting that catches are different between observed and unobserved trips. At the same time, fishermen want their observed trips to be analyzed and used in the stock assessment process as some indicator of abundance. At least one fisherman gave strong support for the use of electronic monitoring to facilitate fishing on a small boat, as a deterrent for unallowed discarding, and to collect better data on fish abundance through fishermen's catch.

One fisherman spoke about the importance of prey abundance in determining fish movements and fish condition. He reported that the most common prey item he finds in plaice stomachs are brittle stars (Ophiuroidea). Further, he noted that brittle star abundance was very high several decades ago, then declined for many years, and then began to increase in the early 2000s. As of 2018, he saw brittle star abundance decrease dramatically. Interestingly, there were also reports, first raised by a fish processor and confirmed by fishermen, that in 2019 an unusual proportion of plaice were being landed in poor physical condition, skinny with poor flesh consistency. In fact, the legal minimum size for plaice is 12", but the market does not want anything under 15" because the yield is low. One fisherman said he believes that the higher the concentration of plaice is in an area, the worse their condition becomes due to food limitation. These skinny fish are said to only come from certain areas where fish concentration is particularly high, while condition remains good in other areas. For groundfish in general, dogfish and seals inflict high levels of predation due to their historically high levels of abundance.

Speaking with fishing industry members strongly directed and shaped research questions. Two threads of the research program have been 1) potential subregional differences in plaice population dynamics within the Gulf of Maine, and 2) the importance of understanding habitat and changing oceans when interpreting landings data and trawl survey data.

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*."

Contributors:

Dan Hennen, Larry Alade, Alex Hansell, Tim Miller, Tyler Pavlovich, Steve Cadrin and Rebecca Peters

Several fishery independent surveys are available to index American plaice stock size and age composition: the NEFSC bottom trawl survey, Massachusetts Department of Marine Fisheries (MADMF) bottom trawl survey and the Maine-New Hampshire (ME/NH) Inshore groundfish survey. All three surveys operate in both the spring and fall with seasonal timing differing slightly between surveys. NEFSC bottom trawl surveys sample nearly the entire stock area and have dedicated subsamples of otoliths for age composition. State surveys sample more limited geographic areas and do not have age composition information available. Design-based and model-based indices that integrate data from multiple surveys were considered for the available survey data.

NEFSC Surveys

The Northeast Fisheries Science Center has conducted stratified random bottom trawl surveys of the Northeast US continental shelf, each fall since 1963 and each spring since 1968 (Politis et al. 2014). Surveys have used several vessels, trawl doors, and nets, so calibration factors were applied to account for survey changes. For example, the survey system transitioned from the Albatross survey system to the Bigelow survey in 2009.

Exploratory analyses of survey data by stratum showed that a higher proportion of large plaice were caught in deeper strata, and the size distribution was smaller in inshore strata than offshore strata (Figure 3.1; Pavlovich et al. 2021, Working Paper 1, Appendix A). Based on the large catches of plaice in inshore stratum 66, the Working Group decided to add all inshore strata that were consistently sampled and consistently caught American plaice (strata 61, 65 and 66) to the strata set used for previous assessments (Figure 3.2). Trends in the aggregate abundance or biomass indices were nearly identical between the offshore-only indices used in previous assessments and the combined inshore-offshore indices used in this assessment. In this strata-set, American plaice are distributed throughout the Gulf of Maine and Georges Bank, with the highest concentrations in the western Gulf of Maine, along the Maine coast, and along the northern edge of Georges Bank (Figure 3.3).

Relative efficiency of the Bigelow survey trawl with rockhopper ground gear to a chain sweep trawl was estimated using studies carried out from 2015 to 2017 (Miller 2021, Working Paper 13, Appendix A). In 2015, 108 paired tows were conducted on eastern Georges Bank and off southern New England. In 2016, 117 paired tows were conducted in western Gulf of Maine and northern edge of Georges Bank. In 2017, 103 paired tows were conducted off southern New England. American plaice were caught in 134 paired tows and 19,245 fish were measured for length. The best estimation model includes size effects and variation in relative catch efficiency between each paired tow (Figure 3.4). The estimated efficiency of the rockhopper gear was applied to survey data to estimate spring (2009-2021) and fall (2009-2019) abundance indices for the stock. American plaice biomass was estimated for each spring and fall annual survey assuming 100% efficiency of the chain sweep gear by scaling the survey tow observations by the relative efficiency of the chain sweep and rockhopper sweep gears. Trends in annual biomass estimates for American plaice for the NEFSC spring and fall survey were generally the same (Figure 3.5, Table 3.1). The scale of the biomass estimates is also similar for the spring and fall surveys for most years. The efficiency of the rockhopper gear relative to the chain sweep in terms of biomass changes from year to year due primarily to corresponding changes in the estimated numbers at length. Annual biomass relative efficiency for American plaice varied between 0.66 and 0.72 in the spring and 0.65 and 0.71 in the fall.

The Working Group considered the Albatross (1963-2008) and Bigelow (2009-2019) survey systems as separate indices and alternatively as a single series by applying calibration factors estimated by Miller et al. (2013) which indicate that the Bigelow is substantially more efficient than the Albatross for catching plaice, particularly at small sizes (Figure 3.6). NEFSC survey indices suggest that the stock has fluctuated with peak abundances in the early 1960s, late 1970s-early 1980s, and the late 2010s (Figure 3.7, Tables 3.2-3.4). NEFSC survey indices of abundance at age suggest the same strong year-classes as the fishery catch-at-age (1987, 1993, 2004, 2013; Figure 3.8, Tables 3.1-2). Correlations of survey catch at age by year-class suggest good cohort tracking among adjacent ages and years (*r*=0.25-0.78; Figure 3.9).

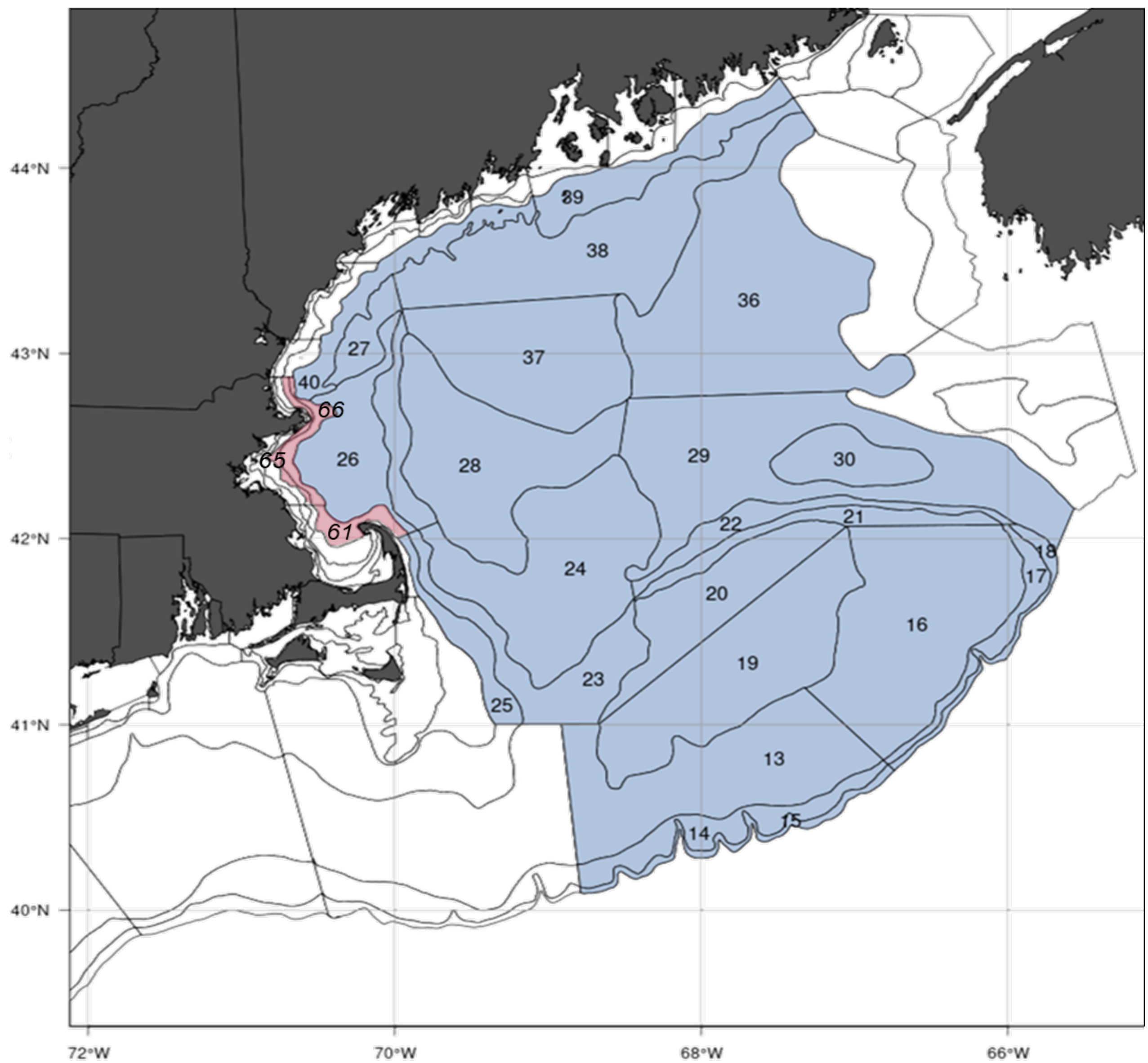


Figure 3.1. NEFSC bottom trawl survey strata used for American plaice indices (selected offshore strata: blue with stratum number; selected inshore strata: red with italic stratum numbers).

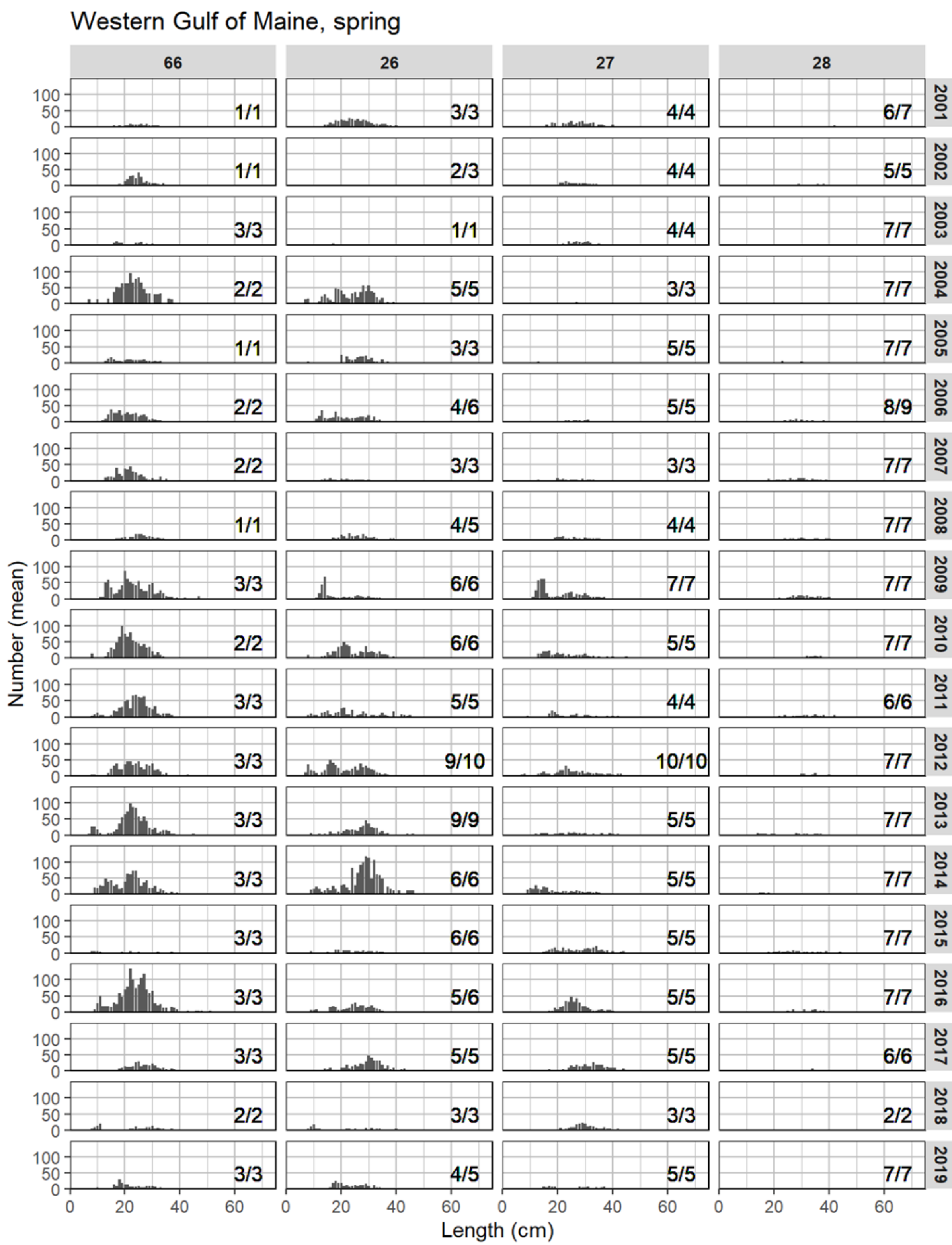


Figure 3.2. Catch of American plaice in the NEFSC spring survey by length, stratum and year in the western Gulf of Maine (numbers in each panel indicate the number of tows that caught plaice in the numerator and total number of tows in the denominator).

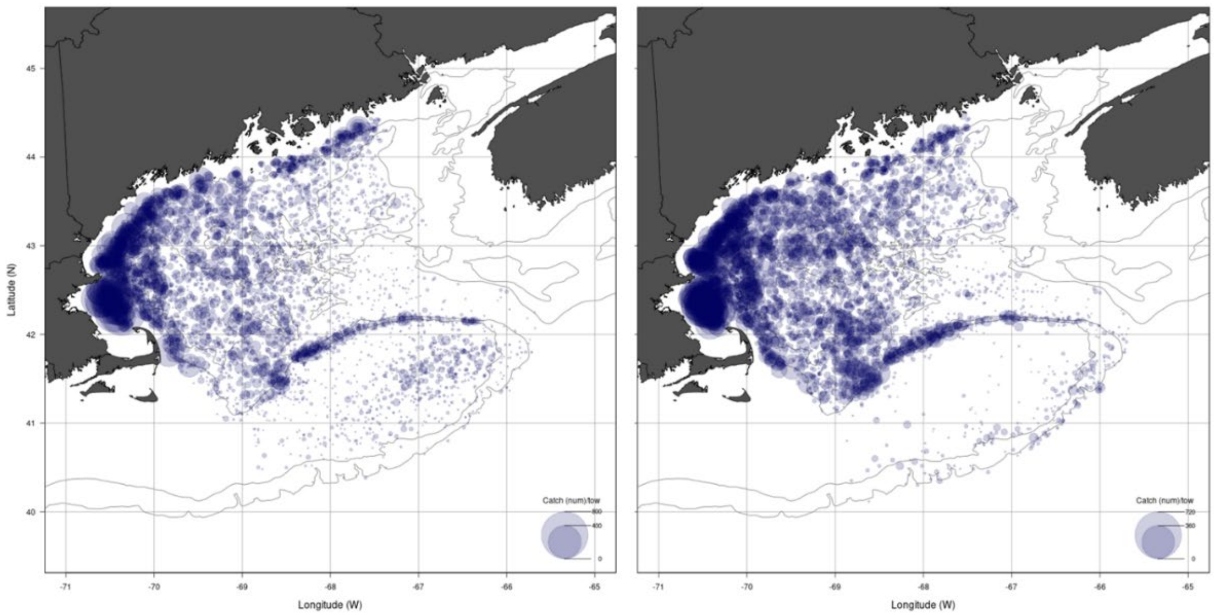


Figure 3.3. Spatial distribution of American plaice from the NEFSC spring survey (left) and fall survey (right).

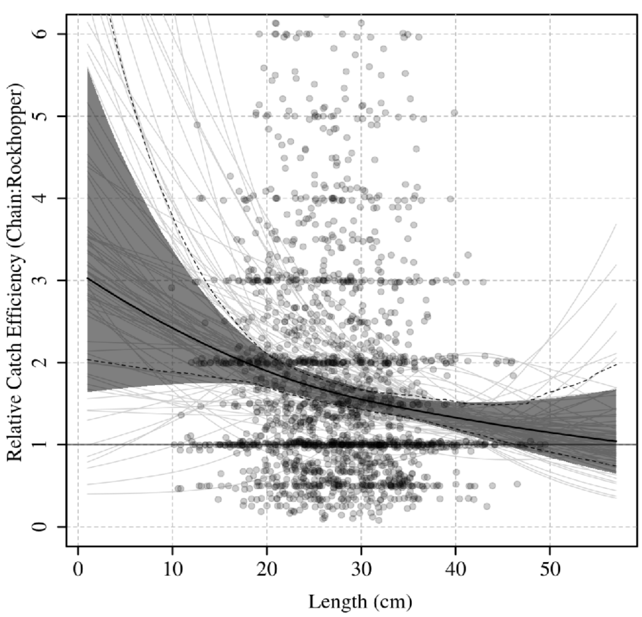


Figure 3.4. Relative efficiency of gears using chain and rockhopper sweeps from the best performing model. Thick and thin lines represent overall and paired-tow specific estimates of relative catch efficiency, respectively. Points represent empirical estimates of relative efficiency for paired observation by length and paired tow. Polygons and dashed lines represent hessian-based and bootstrap-based 95% confidence intervals, respectively.

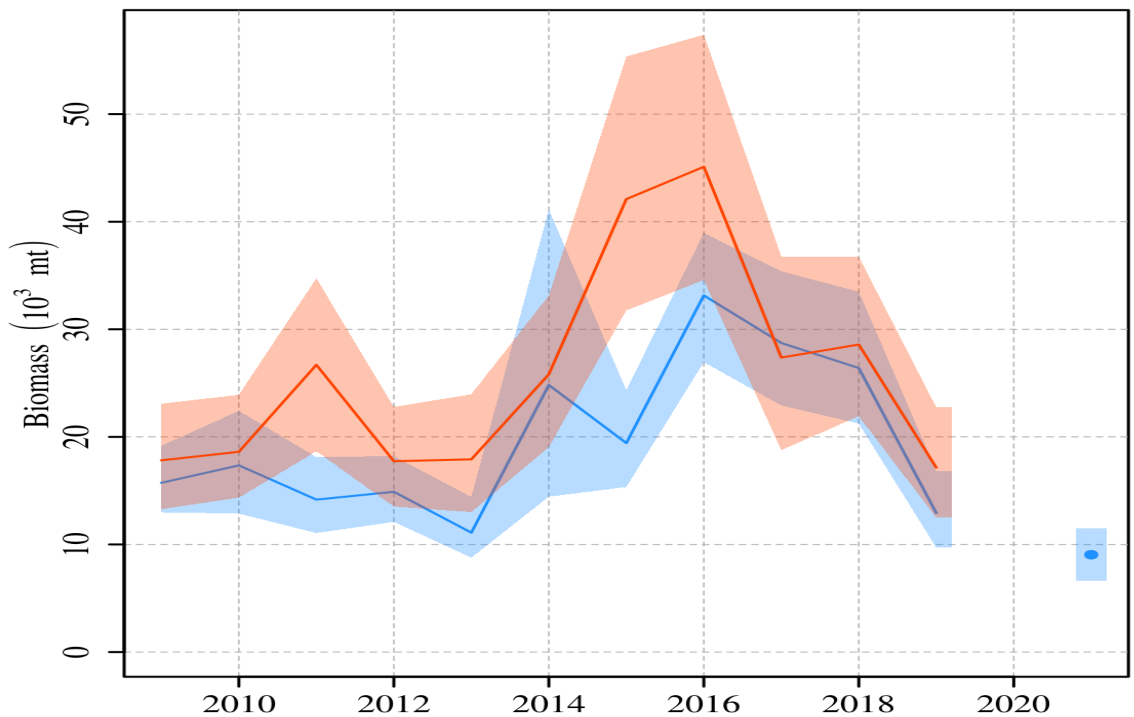


Figure 3.5. Annual spring (blue) and fall (red) biomass estimates for American plaice assuming 100% efficiency for chain sweep gear with shaded polygons representing bootstrap-based 95% confidence intervals. Relative catch efficiency at size estimates and bootstraps are from the best performing model.

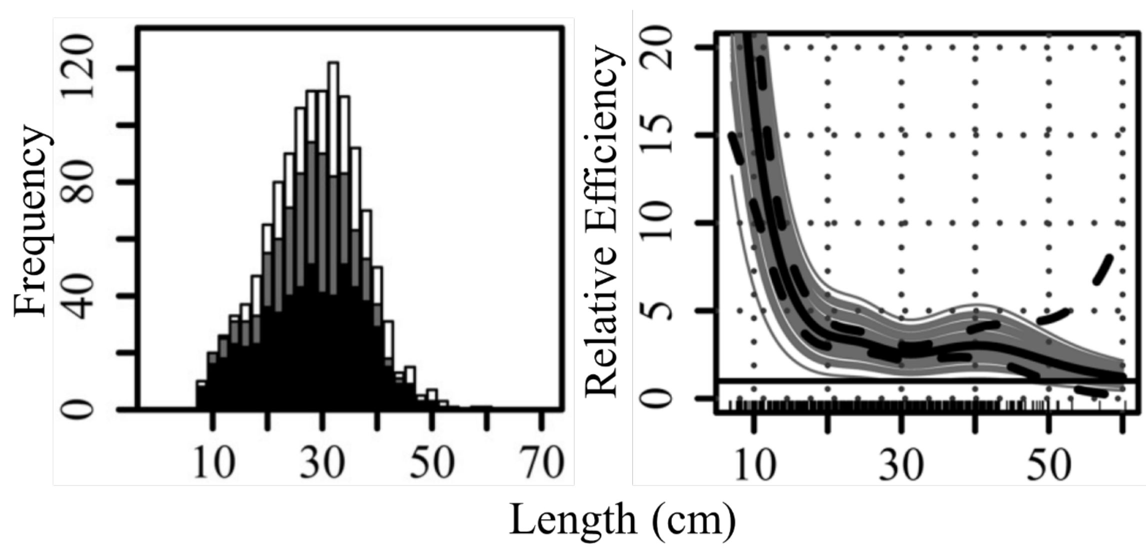


Figure 3.6. Length frequency (left) of paired-tows where American plaice were observed in the Bigelow tow only (black), the Albatross tow only (white), or both (gray); and catch efficiency at length (right) by the Bigelow survey relative to the Albatross survey, with station-specific estimates (gray lines), mean (black line) and 95% confidence intervals (dashed lines). The horizontal black line at 1 indicates equal catch efficiency. Relative efficiency >1 indicates that the Bigelow is more efficient than the Albatross (From Miller 2013).

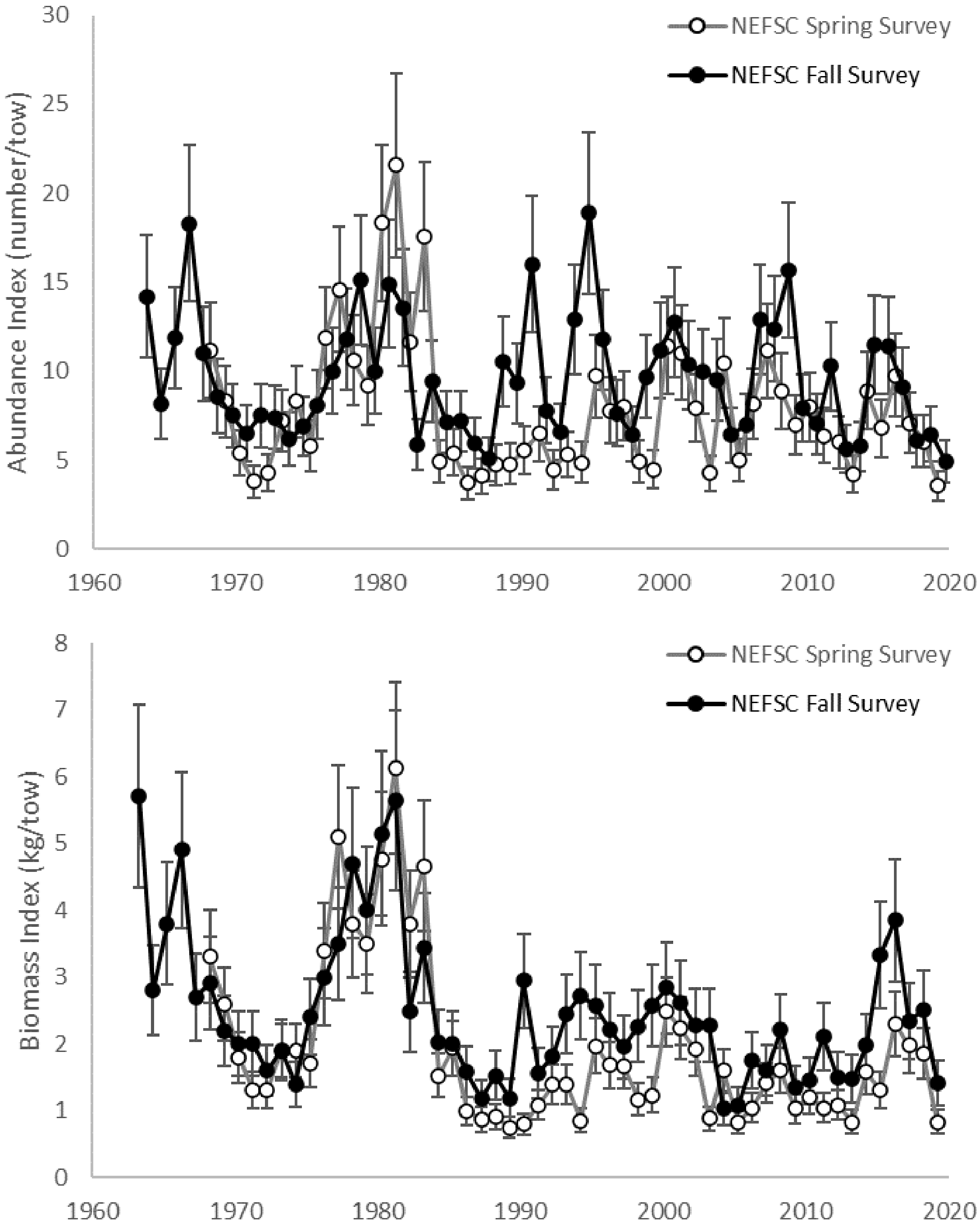


Figure 3.7. NEFSC survey indices of American plaice abundance and biomass, calibrated to Albatross units.

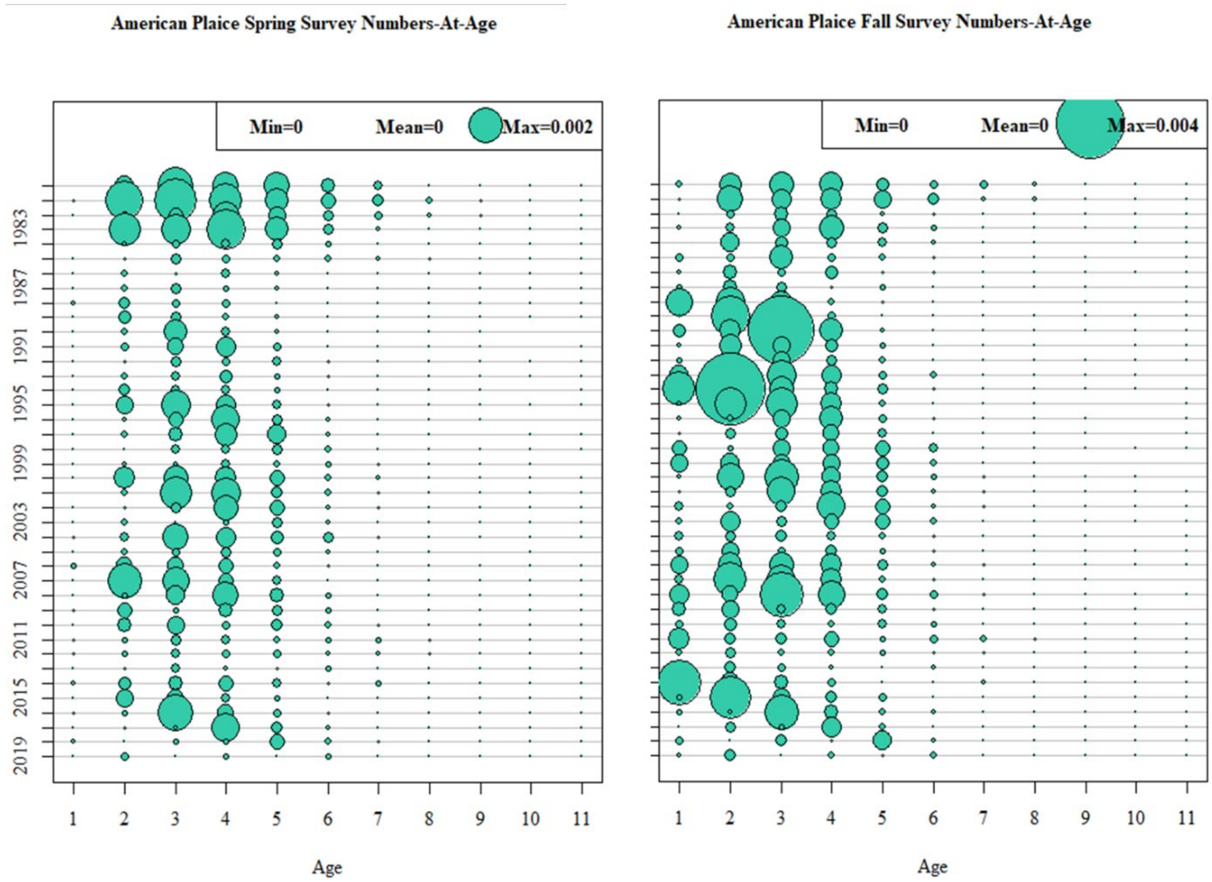


Figure 3.8. NEFSC spring and fall bottom trawl survey abundance at age indices for American plaice.

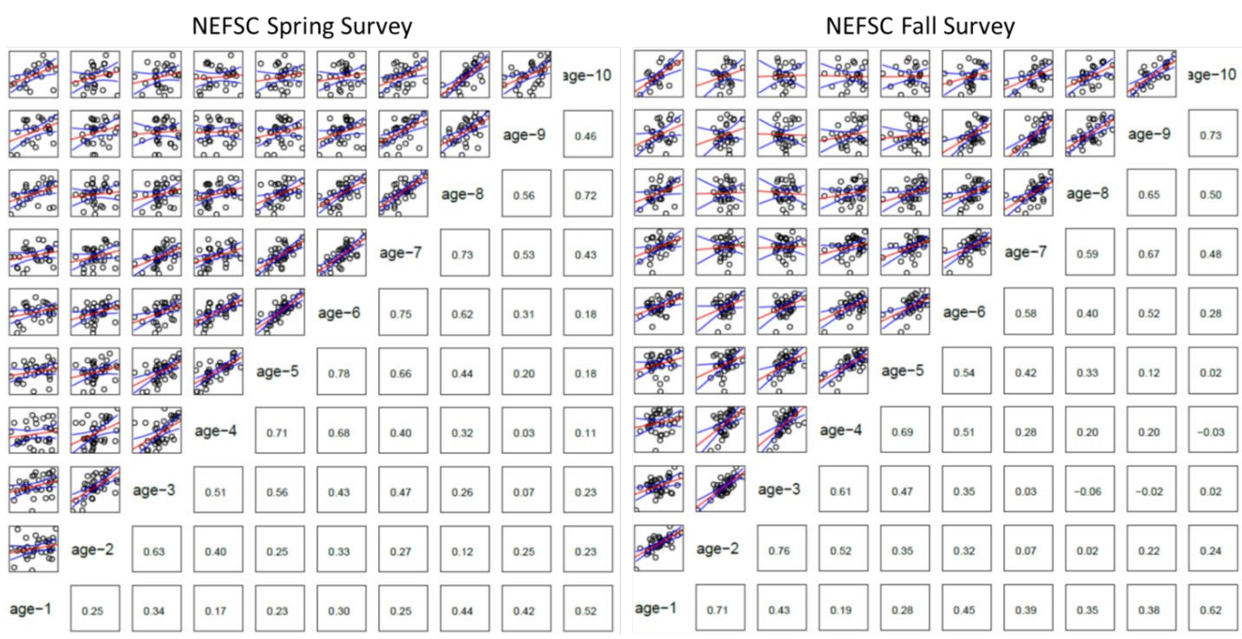


Figure 3.8. Correlation of NEFSC survey indices by year-class over time.

Table 3.1. Estimated chain sweep swept area seasonal survey biomass (mt) for American plaice between 2009 and spring 2021 using relative efficiency estimates for the NEFSC Bigelow survey.

Year	Spring Biomass	CV	CI	Fall Biomass	CV	CI
2009	15,733	0.1	(12,989–19,175)	17,823	0.14	(13,286–23,085)
2010	17,354	0.14	(12,878–22,419)	18,613	0.13	(14,352–23,911)
2011	14,169	0.12	(11,061–18,116)	26,700	0.16	(18,664–34,746)
2012	14,898	0.1	(12,100–18,211)	17,743	0.14	(13,523–22,781)
2013	11,098	0.14	(8,775–14,408)	17,912	0.16	(13,012–23,961)
2014	24,822	0.29	(14,445–41,143)	25,812	0.15	(19,048–33,144)
2015	19,423	0.12	(15,338–24,327)	42,101	0.14	(31,755–55,359)
2016	33,146	0.09	(26,949–38,956)	45,106	0.13	(34,581–57,374)
2017	28,729	0.11	(22,925–35,383)	27,376	0.17	(18,768–36,781)
2018	26,398	0.12	(21,238–33,496)	28,584	0.13	(21,925–36,746)
2019	12,927	0.14	(9,704–16,824)	17,147	0.15	(12,493–22,752)
2020	–	–	–	–	–	–
2021	9,046	0.14	(6,628–11,472)	–	–	–

Table 3.2. Indices of aggregate abundance and biomass of American plaice from NEFSC spring and fall surveys with coefficients of variation (CV), calibrated to Albatross units.

	Spring				Fall			
Year	No/Tow	CV	Wt/Tow	CV	No/Tow	CV	Wt/Tow	CV
1963	-	-	-	-	14.2	13.8	5.7	14.8
1964	-	-	-	-	8.2	13.0	2.8	13.2
1965	-	-	-	-	11.9	12.2	3.8	11.9
1966	-	-	-	-	18.3	14.3	4.9	13.3
1967	-	-	-	-	11.0	20.4	2.7	20.5
1968	11.2	18.5	3.3	16.5	8.6	14.9	2.9	15.9
1969	8.3	13.0	2.6	12.0	7.5	14.9	2.2	13.4
1970	5.4	11.3	1.8	11.6	6.5	13.7	2.0	13.3
1971	3.8	12.8	1.3	14.5	7.5	15.3	2.0	18.9
1972	4.3	9.9	1.3	11.6	7.4	13.9	1.6	12.6
1973	7.2	11.1	1.9	11.0	6.2	9.7	1.9	14.5
1974	8.3	24.5	1.9	16.9	6.9	23.1	1.4	21.1
1975	5.8	14.6	1.7	12.4	8.1	10.9	2.4	10.2
1976	11.9	12.1	3.4	9.7	10.0	17.3	3.0	13.1
1977	14.6	11.2	5.1	18.6	11.8	9.3	3.5	9.4
1978	10.6	9.9	3.8	9.9	15.1	8.6	4.7	10.4
1979	9.2	8.1	3.5	10.2	10.0	8.3	4.0	9.8
1980	18.3	14.8	4.8	9.1	14.9	8.9	5.1	10.5
1981	21.6	9.4	6.1	10.6	13.6	10.2	5.6	11.9
1982	11.6	15.6	3.8	13.2	5.9	14.0	2.5	14.5
1983	17.6	22.2	4.7	12.4	9.4	12.0	3.4	12.6
1984	5.0	14.0	1.5	10.3	7.2	17.4	2.0	11.8
1985	5.4	9.8	1.9	9.8	7.2	12.3	2.0	11.1
1986	3.7	15.3	1.0	13.4	6.0	20.3	1.6	13.8
1987	4.1	11.2	0.9	11.5	5.1	10.9	1.2	10.4
1988	4.7	13.5	0.9	12.2	10.5	27.2	1.5	16.2
1989	4.8	15.0	0.7	13.6	9.3	19.3	1.2	14.6
1990	5.6	17.5	0.8	11.2	16.0	14.9	2.9	11.9
1991	6.5	14.9	1.1	12.7	7.8	11.8	1.6	8.5
1992	4.5	10.5	1.4	10.8	6.6	16.8	1.8	13.9
1993	5.3	13.2	1.4	11.8	12.9	15.5	2.4	13.0
1994	4.9	15.1	0.8	11.8	18.9	19.9	2.7	13.8
1995	9.7	16.8	2.0	11.1	11.8	14.1	2.6	13.7
1996	7.8	13.4	1.7	10.3	7.6	13.6	2.2	17.8
1997	8.0	32.0	1.7	19.9	6.5	13.1	2.0	13.9
1998	4.9	11.4	1.2	9.8	9.7	14.7	2.3	12.1
1999	4.5	12.9	1.2	13.3	11.2	17.0	2.6	17.0
2000	11.4	15.2	2.5	13.9	12.8	21.0	2.8	19.1
2001	11.0	15.4	2.2	11.7	10.4	19.2	2.6	17.3
2002	7.9	13.3	1.9	10.5	10.0	16.7	2.3	17.9
2003	4.3	12.3	0.9	10.4	9.5	20.3	2.3	29.1
2004	10.4	23.2	1.6	19.0	6.4	9.6	1.0	14.9
2005	5.0	31.8	0.8	33.4	7.0	13.4	1.1	17.2
2006	8.2	15.7	1.0	13.9	12.9	18.4	1.7	13.4
2007	11.1	12.5	1.4	12.1	12.4	18.0	1.6	16.0
2008	8.9	12.2	1.6	11.1	15.7	13.7	2.2	12.5
2009	7.0	12.4	1.0	8.0	7.9	15.2	1.3	12.4
2010	8.0	12.8	1.2	11.6	7.0	12.2	1.4	11.5
2011	6.4	10.7	1.0	11.0	10.3	21.9	2.1	15.2
2012	6.0	13.3	1.1	9.9	5.7	16.3	1.5	13.4
2013	4.2	11.9	0.8	10.9	5.8	10.7	1.5	17.3
2014	8.9	21.0	1.6	28.0	11.5	13.1	2.0	15.2
2015	6.8	12.9	1.3	11.8	11.4	12.0	3.3	15.3
2016	9.8	9.4	2.3	9.0	9.1	11.8	3.8	13.3
2017	7.1	10.9	2.0	10.4	6.1	15.7	2.3	18.7
2018	6.0	11.3	1.9	11.0	6.5	11.5	2.5	12.3
2019	3.6	16.4	0.8	12.7	5.0	13.9	1.4	14.5

Table 3.3. Northeast Fisheries Science Center (NEFSC) spring survey abundance indices-at-age (numbers/tow) from 1970 to 2012 for American plaice, calibrated to Albatross units.

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	0.000	2.238	4.092	3.036	2.980	1.608	1.073	0.244	0.298	0.231	0.085
1981	0.378	4.505	4.994	3.854	2.800	1.748	1.453	0.782	0.399	0.323	0.317
1982	0.017	0.944	1.856	3.249	2.079	1.286	0.908	0.487	0.349	0.188	0.121
1983	0.000	3.745	3.448	4.645	2.718	1.203	0.587	0.322	0.146	0.147	0.217
1984	0.000	0.522	0.917	1.097	1.228	0.736	0.255	0.082	0.009	0.022	0.056
1985	0.022	0.397	1.197	1.012	0.766	0.827	0.485	0.408	0.110	0.073	0.093
1986	0.009	0.787	0.439	1.159	0.603	0.322	0.189	0.128	0.038	0.013	0.004
1987	0.107	0.773	1.296	0.747	0.558	0.267	0.163	0.074	0.029	0.031	0.021
1988	0.584	1.398	1.033	0.933	0.306	0.229	0.097	0.027	0.081	0.037	0.010
1989	0.014	1.580	1.272	0.864	0.492	0.279	0.145	0.028	0.071	0.007	0.005
1990	0.005	0.875	2.787	1.055	0.520	0.183	0.067	0.048	0.048	0.000	0.000
1991	0.034	0.905	1.947	2.378	0.923	0.147	0.065	0.020	0.018	0.000	0.027
1992	0.090	0.412	1.295	0.922	1.126	0.433	0.099	0.038	0.024	0.013	0.005
1993	0.329	0.869	1.161	1.567	0.631	0.462	0.166	0.083	0.015	0.007	0.040
1994	0.029	1.445	1.138	1.116	0.739	0.229	0.098	0.035	0.007	0.000	0.028
1995	0.032	2.039	3.411	2.325	1.123	0.446	0.215	0.025	0.032	0.027	0.047
1996	0.022	0.512	1.832	3.289	1.305	0.508	0.215	0.044	0.018	0.000	0.000
1997	0.011	0.847	1.636	2.587	2.293	0.548	0.088	0.012	0.000	0.005	0.015
1998	0.058	0.228	1.161	1.125	1.290	0.804	0.186	0.063	0.004	0.009	0.012
1999	0.085	0.536	0.540	1.173	0.800	0.701	0.429	0.169	0.023	0.015	0.000
2000	0.048	2.456	2.913	2.454	1.758	0.904	0.628	0.154	0.074	0.022	0.016
2001	0.000	0.790	3.754	3.432	1.418	0.861	0.408	0.190	0.100	0.050	0.019
2002	0.068	0.361	1.319	2.871	1.807	0.597	0.389	0.236	0.172	0.065	0.008
2003	0.036	0.808	0.265	0.716	1.262	0.613	0.221	0.126	0.089	0.041	0.072
2004	0.443	1.058	3.095	2.425	1.521	1.234	0.351	0.099	0.183	0.012	0.028
2005	0.193	0.773	1.026	1.257	0.918	0.499	0.227	0.114	0.000	0.016	0.000
2006	0.761	1.964	1.910	1.805	0.817	0.376	0.347	0.119	0.019	0.018	0.023
2007	0.255	4.094	3.221	1.840	1.125	0.335	0.139	0.068	0.010	0.043	0.005
2008	0.118	0.720	2.201	3.015	1.647	0.713	0.269	0.086	0.060	0.038	0.030
2009	0.362	1.787	0.630	1.685	1.314	0.675	0.326	0.118	0.040	0.010	0.021
2010	0.157	1.745	2.157	1.027	1.384	0.909	0.351	0.124	0.083	0.030	0.018
2011	0.399	0.701	1.191	1.084	0.910	0.671	0.716	0.394	0.184	0.098	0.037
2012	0.361	0.765	1.103	0.978	0.956	0.524	0.556	0.437	0.191	0.051	0.072
2013	0.253	0.437	1.149	0.611	0.369	0.678	0.192	0.247	0.173	0.089	0.049
2014	0.553	1.560	1.663	1.880	1.055	0.458	0.768	0.261	0.261	0.263	0.194
2015	0.065	2.092	1.995	1.111	0.644	0.271	0.166	0.250	0.076	0.062	0.075
2016	0.486	0.755	4.238	1.929	0.948	0.451	0.165	0.237	0.250	0.124	0.185
2017	0.056	0.425	0.603	3.346	1.362	0.527	0.348	0.114	0.068	0.117	0.129
2018	0.511	0.181	0.645	0.648	1.872	0.887	0.408	0.270	0.203	0.116	0.310
2019	0.151	0.957	0.259	0.665	0.317	0.653	0.251	0.121	0.047	0.031	0.099

Table 3.4. Northeast Fisheries Science Center (NEFSC) spring survey abundance indices-at-age (numbers/tow) from 1970 to 2012 for American plaice, calibrated to Albatross units.

Year	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10	Age11+
1980	0.887	2.626	2.911	2.757	1.524	0.980	0.907	0.560	0.296	0.191	0.479
1981	0.397	3.115	2.309	2.464	2.088	1.417	0.574	0.513	0.070	0.157	0.462
1982	0.219	0.914	1.661	1.256	0.563	0.480	0.296	0.165	0.187	0.073	0.134
1983	0.550	1.171	2.015	2.886	1.308	0.664	0.331	0.171	0.094	0.028	0.178
1984	0.000	2.200	1.573	1.210	1.058	0.511	0.122	0.101	0.000	0.028	0.065
1985	1.035	0.956	2.719	0.981	0.778	0.406	0.188	0.050	0.031	0.000	0.013
1986	0.508	1.633	0.977	1.476	0.467	0.419	0.156	0.097	0.036	0.013	0.040
1987	0.634	1.678	1.222	0.415	0.657	0.223	0.103	0.045	0.039	0.000	0.023
1988	3.144	3.411	2.508	0.783	0.469	0.102	0.035	0.000	0.028	0.000	0.024
1989	0.460	4.637	2.819	0.968	0.170	0.072	0.014	0.023	0.019	0.017	0.048
1990	1.572	2.555	7.767	2.847	0.574	0.244	0.114	0.069	0.021	0.020	0.040
1991	0.493	2.614	2.022	1.551	0.701	0.283	0.044	0.063	0.000	0.010	0.020
1992	0.695	1.216	2.063	1.297	0.775	0.298	0.067	0.049	0.030	0.022	0.037
1993	2.170	2.781	3.533	2.341	1.034	0.793	0.109	0.039	0.037	0.043	0.043
1994	3.868	8.218	2.952	1.687	1.277	0.391	0.248	0.131	0.013	0.031	0.046
1995	0.526	3.769	3.788	2.476	0.894	0.211	0.026	0.026	0.000	0.000	0.015
1996	0.554	0.798	2.026	2.781	0.917	0.382	0.071	0.042	0.032	0.000	0.030
1997	0.367	1.232	1.500	1.895	1.065	0.313	0.042	0.010	0.012	0.000	0.015
1998	1.882	0.657	2.023	1.975	1.851	1.086	0.113	0.046	0.010	0.011	0.023
1999	2.065	2.243	2.090	2.119	1.569	0.845	0.196	0.026	0.000	0.000	0.000
2000	0.569	3.180	4.080	2.211	1.404	0.709	0.387	0.132	0.034	0.031	0.031
2001	0.411	1.267	3.290	2.582	1.424	0.529	0.396	0.255	0.162	0.024	0.037
2002	1.122	0.808	1.359	3.394	1.780	0.531	0.400	0.280	0.164	0.061	0.102
2003	0.822	2.378	1.246	1.749	1.749	0.882	0.338	0.131	0.056	0.083	0.109
2004	1.062	1.210	1.329	1.175	0.808	0.485	0.189	0.091	0.030	0.000	0.023
2005	0.985	2.062	0.965	1.388	0.764	0.441	0.155	0.112	0.081	0.035	0.057
2006	2.051	2.824	2.983	2.463	1.322	0.557	0.434	0.176	0.032	0.041	0.050
2007	1.131	3.862	3.310	2.499	0.883	0.462	0.150	0.038	0.019	0.000	0.000
2008	2.354	1.964	5.184	3.192	1.425	0.930	0.416	0.047	0.068	0.015	0.045
2009	1.683	2.125	1.116	1.342	1.083	0.355	0.089	0.051	0.024	0.027	0.027
2010	1.015	1.569	1.132	0.812	1.152	0.703	0.339	0.082	0.046	0.044	0.058
2011	2.551	1.365	1.343	1.790	0.738	1.001	0.806	0.374	0.172	0.032	0.132
2012	0.848	1.212	0.865	0.851	0.674	0.373	0.386	0.245	0.115	0.027	0.048
2013	1.130	1.359	1.034	0.577	0.339	0.539	0.166	0.105	0.174	0.031	0.035
2014	5.195	2.214	1.730	0.965	0.268	0.205	0.522	0.086	0.161	0.093	0.156
2015	0.723	4.888	2.172	1.440	0.997	0.349	0.178	0.278	0.103	0.070	0.258
2016	0.730	0.521	4.017	1.619	0.792	0.570	0.189	0.115	0.236	0.119	0.334
2017	0.159	1.322	0.638	2.406	0.773	0.271	0.135	0.107	0.033	0.095	0.214
2018	0.983	0.172	1.387	0.354	2.180	0.629	0.272	0.226	0.052	0.064	0.188
2019	0.576	1.447	0.295	0.901	0.379	0.865	0.138	0.128	0.051	0.017	0.130

State Surveys

MADMF Inshore Bottom Trawl Survey

The Massachusetts Division of Marine Fisheries (MADMF) has conducted research bottom trawl surveys during the spring and fall since 1978 (King et al. 2010). The survey strata included in the MADMF survey index of plaice primarily includes habitat within the Massachusetts state waters in the southwestern Gulf of Maine (Figure 3.9). The abundance indices of MADMF surveys have different trends from the NEFSC surveys, with both the fall and spring indices currently at or near record low (Figure 3.10). Previous stock assessments hypothesized that difference between the NEFSC offshore and MADMF inshore survey is likely due to plaice shifting to deeper waters and decreasing availability of the resource to the survey (NEFSC 2017), and MADMF indices were excluded from the assessment model in the 2019 assessment update (NEFSC 2022b).

Application of age-length keys from NEFSC surveys suggest that MADMF surveys catch mostly age 1 and 2 plaice. The age-based indices had moderate to poor year class tracking (r=0.0–0.8), which could result from seasonal movement of plaice between inshore and offshore or using age-length keys form offshore to characterize the inshore survey.

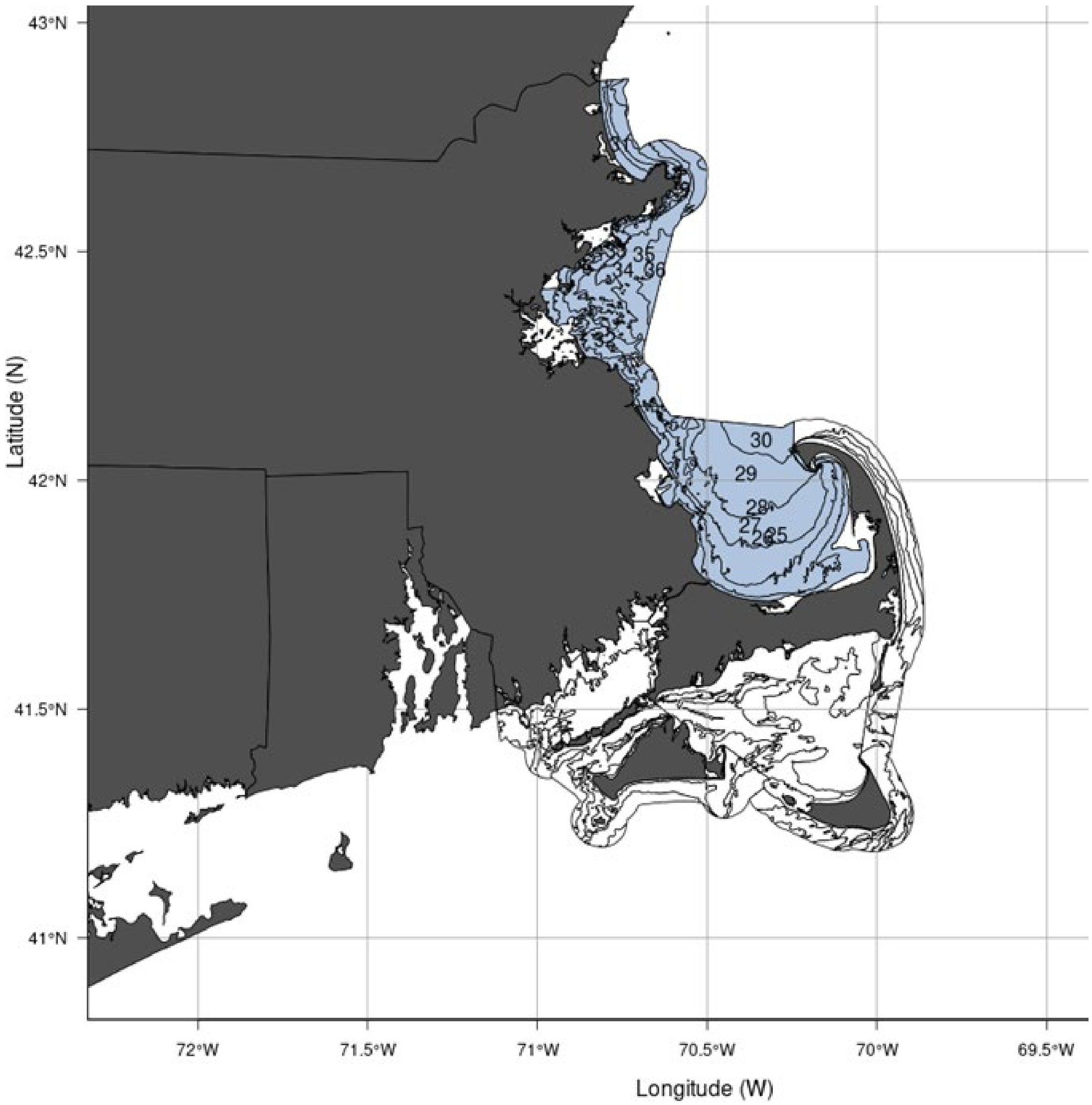


Figure 3.9. MADMF bottom trawl survey strata included in the American plaice assessment (shaded blue).

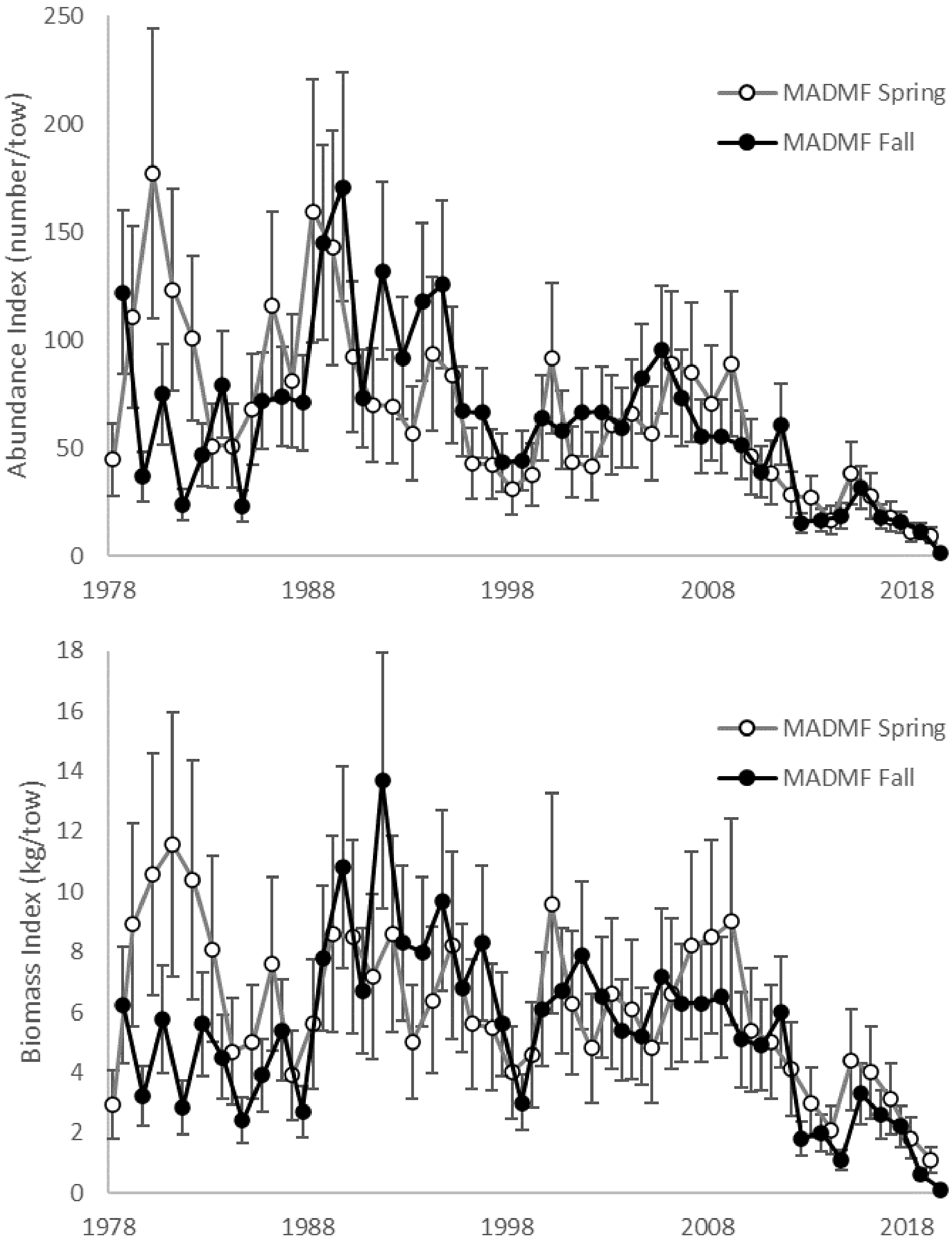


Figure 3.10. MADMF survey indices of American plaice abundance and biomass.

ME-NH inshore groundfish trawl Survey

The Maine Department of Marine Resources and New Hampshire and Game Department (ME-NH) have conducted spring and bottom trawl surveys along the New Hampshire and Maine Coast since fall 2000. The survey has a stratified random sampling design with a fixed component. Sampling occurs in region 5 and 4 depth zones that range from 5 fathoms to greater than 55 fathoms (Sherman et al. 2005; Figure 3.11).

The ME-NH inshore groundfish trawl survey has not been included in previous assessments. Indices of plaice abundance and biomass from the ME-NH survey generally increased in the 2000s to peaks in 2008 or 2009 and generally decreased since then. Age structures have been collected on the survey, but they have not been processed. Application of age-length keys from NEFSC surveys suggest that the ME-NH survey catches mostly ages 1 and 2.

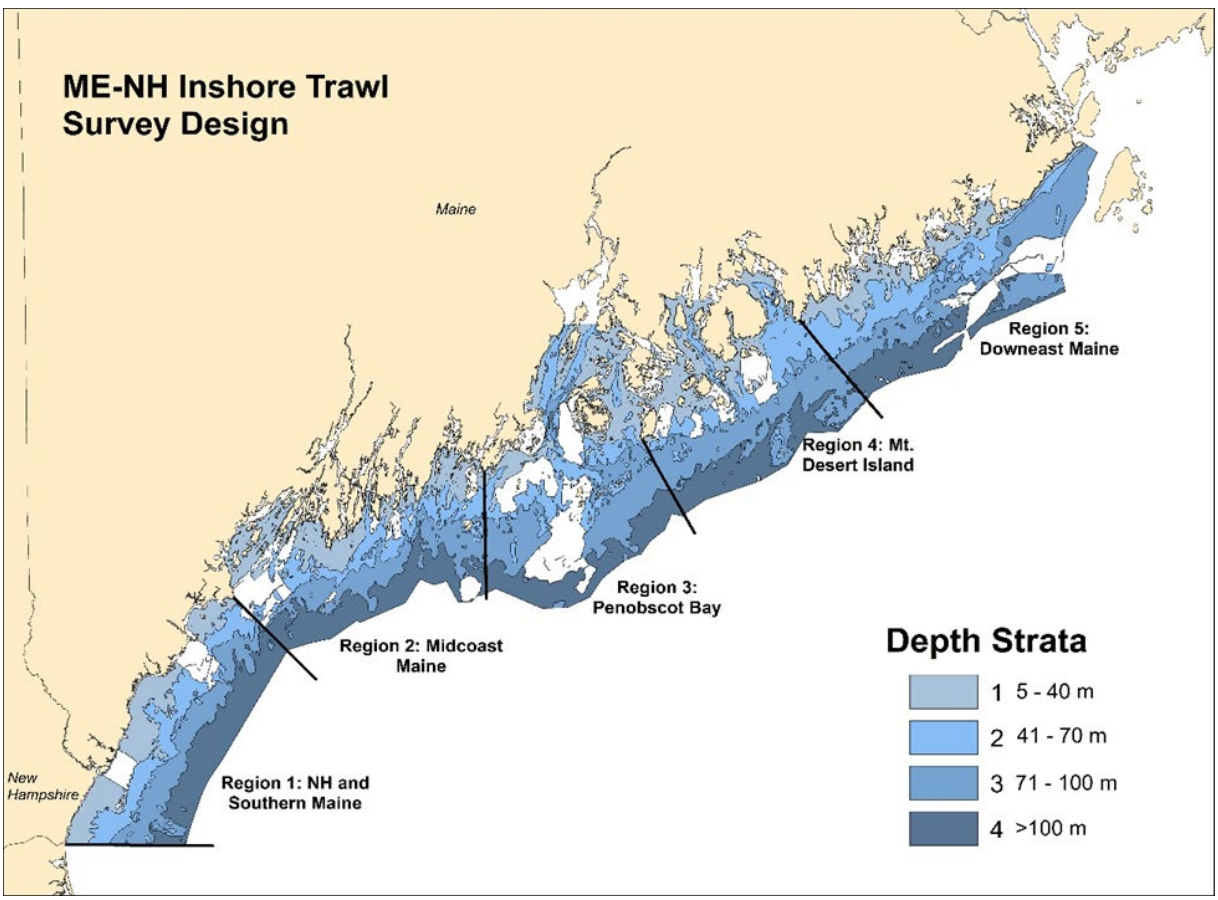


Figure 3.11. ME-NH groundfish trawl survey strata.

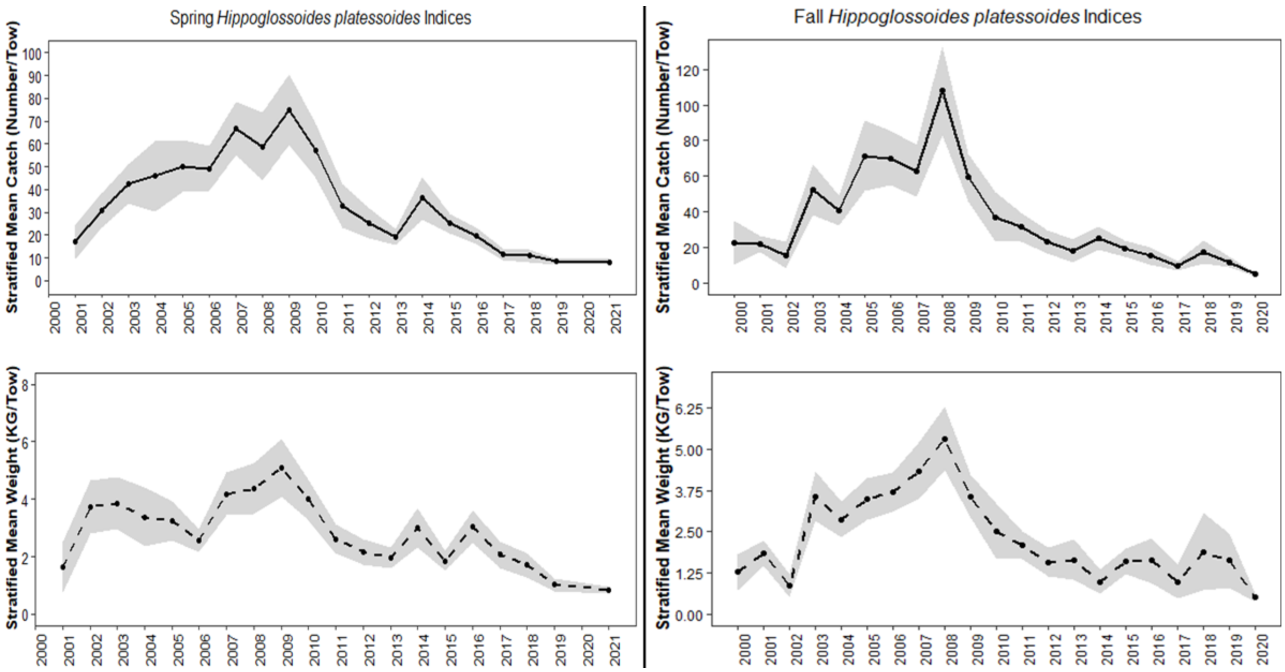


Figure 3.12. ME-NH survey indices of American plaice abundance and biomass.

Integrated Survey Indices

Data was compiled from the three semi-annual trawl surveys (Figure 3.12): NEFSC (1963-2019), (MADMF; 1981-2019) and ME-NH (2005-2019) using a vector auto-regressive spatiotemporal model (VAST, Thorson 2015) to estimate changes in spatial distribution and develop standardized indices of abundance (Hansell et al. 2021, Working Paper 12, Appendix A). 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).

VAST is 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 gamma distribution was assumed for the distribution of positive catch. A factor was used to account for vessel effects between the three surveys. Local covariates vary across space, while regional covariates are univariate and represent temporal changes across the entire stock. Bottom temperature and depth associated with each tow were explored as local covariates. Models with AMO or NAO as regional covariates did not converge.

Over the time series, the center of gravity has been variable in both the spring and fall with periods of northeast and southwest movement (Figure 3.13). Since the 1960's, the effective area occupied has decreased in the spring and fall by an average rate of 177.9 (SD =9240.4) and 80.6 (SD = 7608.8) km^2/year (Figure 3.14). Counterfactual analyses predicted that depth was the primary driver of both spring and fall distribution changes.

The combined spatiotemporal index of abundance mostly reflects the NEFSC survey because of its larger spatial extent. VAST estimates of relative abundance for the NEFSC and MADMF trawl surveys are similar to design-based estimates, but VAST estimates are different than design-based estimates for the MENH survey (Figure 3.15). The Working Group decided to explore stock assessment models that include the VAST index of stock size.

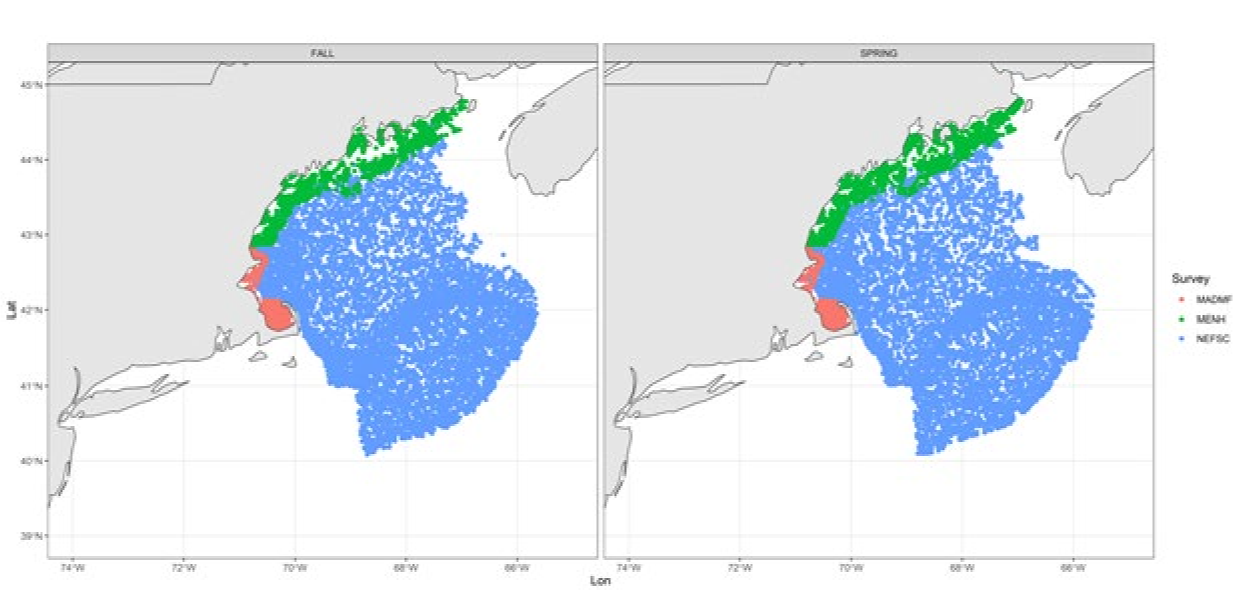


Figure 3.12. Tow locations for the three surveys used in VAST model for American plaice

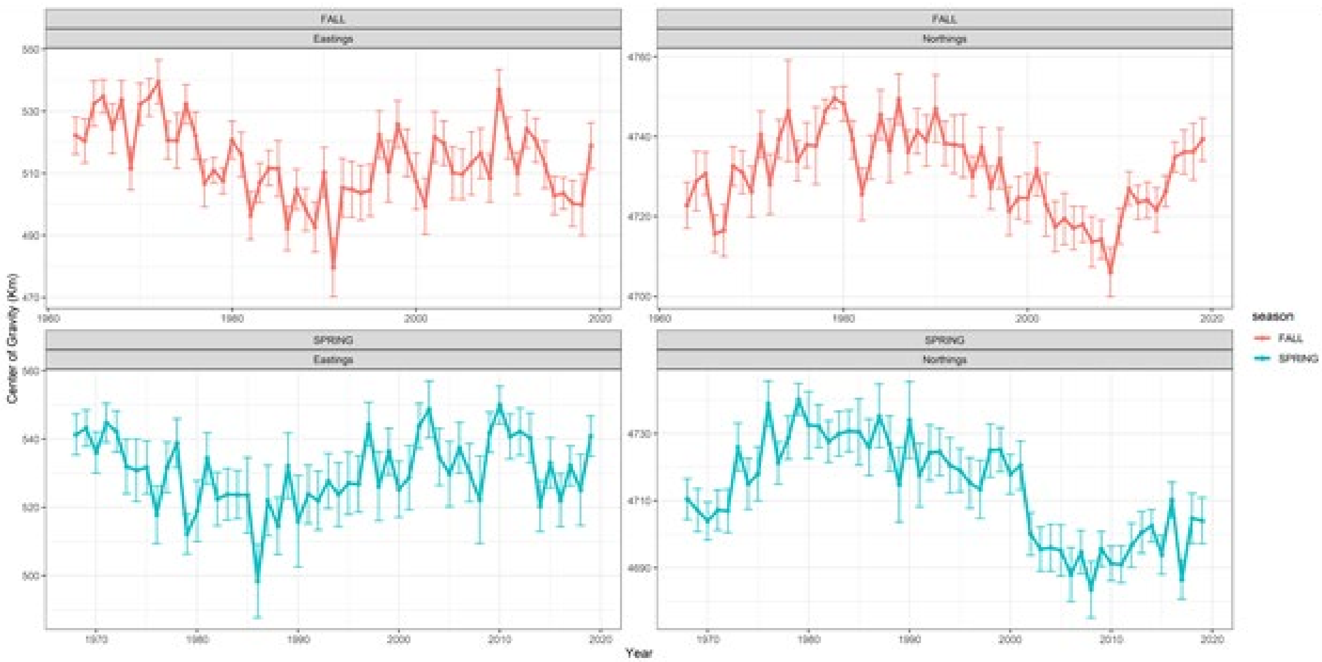


Figure 3.13. Center of gravity estimates produced by VAST for American plaice.

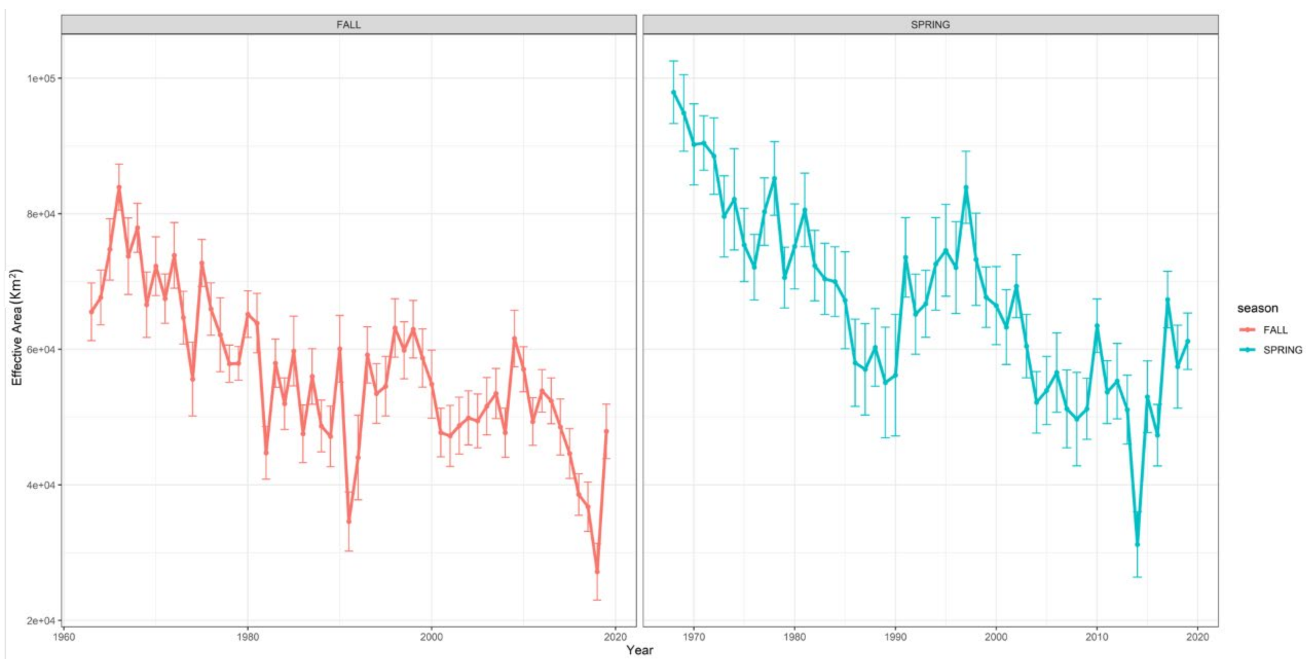


Figure 3.14. Effective area estimates produced by VAST for American plaice.

Figure 3.15. Comparison of design-based and VAST indices for NEFSC, MADMF, ME-NH and a single integrated index (all) for American plaice.

Table 3.5. VAST indices of NEFSC, MADMF, ME-NH and combined spring surveys.

Year	All	CV	NEFSC	NEFSC CV	MADMF	MADMF CV	MENH	MENH CV
1968	1.350	0.849	1.367	0.835				
1969	1.519	0.831	1.527	0.813				
1970	1.500	0.822	1.488	0.798				
1971	1.379	0.838	1.384	0.825				
1972	1.276	0.881	1.274	0.861				
1973	1.273	1.022	1.264	1.002				
1974	1.428	1.265	1.409	1.229				
1975	1.068	1.107	1.080	1.094				
1976	1.086	0.896	1.090	0.888				
1977	1.500	0.809	1.481	0.804				
1978	1.453	0.846	1.479	0.839				
1979	1.463	0.672	1.438	0.660				
1980	1.628	0.981	1.624	0.985				
1981	2.431	0.796	2.419	0.773	2.328	1.848		
1982	1.698	0.961	1.691	0.974	1.340	0.794		
1983	1.526	0.864	1.527	0.869	1.648	0.905		
1984	0.952	1.032	0.959	1.033	0.778	0.798		
1985	1.089	1.452	1.093	1.455	1.014	1.210		
1986	1.060	1.356	1.042	1.364	1.285	1.163		
1987	0.757	1.607	0.768	1.601	0.625	1.502		
1988	0.696	1.362	0.687	1.371	0.703	1.270		
1989	0.809	1.599	0.806	1.620	1.002	1.041		
1990	0.916	1.692	0.899	1.732	1.489	1.080		
1991	1.061	1.161	1.073	1.152	1.142	0.814		
1992	1.089	1.127	1.059	1.155	1.207	0.931		
1993	0.944	1.010	0.934	1.018	0.814	0.883		
1994	0.805	1.210	0.789	1.228	0.818	1.077		
1995	1.126	1.016	1.099	1.041	1.164	1.068		
1996	0.990	1.081	0.976	1.106	1.038	0.991		
1997	1.145	0.977	1.141	0.983	0.981	1.028		
1998	1.027	1.100	1.005	1.126	0.825	1.107		
1999	1.052	0.971	1.041	0.989	0.972	0.834		
2000	1.195	0.920	1.185	0.933	1.313	0.953	1.223	1.291
2001	0.961	1.038	0.964	1.046	1.147	0.912	1.276	1.351
2002	1.060	0.896	1.096	0.889	0.972	0.823	1.012	1.534
2003	0.916	1.018	0.936	1.022	0.994	0.871	1.067	1.388
2004	0.746	0.952	0.752	0.967	1.180	0.726	0.856	1.034
2005	0.552	1.101	0.559	1.113	0.829	1.107	0.701	0.858
2006	0.656	0.957	0.659	0.976	0.980	0.686	0.686	0.947
2007	0.714	1.009	0.725	1.019	1.182	0.887	0.685	1.013
2008	0.733	1.356	0.728	1.399	1.278	0.851	0.803	1.013
2009	0.558	0.680	0.558	0.697	0.980	0.646	0.902	0.927
2010	0.569	0.757	0.586	0.753	0.824	0.638	0.980	0.803
2011	0.559	0.751	0.568	0.756	0.923	0.824	1.172	0.844
2012	0.531	0.821	0.543	0.823	0.797	0.931	0.997	0.695
2013	0.384	0.852	0.389	0.861	0.565	0.884	0.850	0.852
2014	0.419	0.746	0.428	0.749	0.955	1.094	0.863	1.014
2015	0.446	0.784	0.449	0.797	0.641	0.911	1.248	0.816
2016	0.515	0.770	0.528	0.770	0.880	0.942	1.619	0.753
2017	0.576	0.828	0.591	0.829	0.632	0.888	1.238	0.969
2018	0.433	1.241	0.445	1.240	0.445	1.174	0.901	0.881
2019	0.382	0.843	0.397	0.833	0.309	1.324	0.919	0.959

Table 3.6. VAST indices of NEFSC, MADMF, ME-NH and combined fall surveys.

Year	All	All CV	NEFSC	NEFSC CV	MADMF	MADMF CV	MENH	MENH CV
1963	1.203	0.991	1.205	0.972				
1964	1.163	1.223	1.180	1.221				
1965	1.382	1.340	1.379	1.330				
1966	1.588	1.257	1.597	1.257				
1967	1.193	1.255	1.179	1.238				
1968	1.352	1.229	1.343	1.189				
1969	1.150	1.270	1.132	1.227				
1970	1.121	1.144	1.117	1.108				
1971	0.988	0.926	0.987	0.915				
1972	1.029	1.114	1.038	1.117				
1973	0.954	0.866	0.962	0.859				
1974	0.754	1.005	0.754	1.001				
1975	1.051	0.960	1.060	0.957				
1976	1.203	1.267	1.181	1.249				
1977	1.320	1.367	1.313	1.347				
1978	1.478	0.772	1.479	0.769				
1979	1.276	0.702	1.293	0.708				
1980	1.360	1.153	1.349	1.146				
1981	1.435	1.264	1.433	1.258	1.248	2.306		
1982	0.956	1.255	0.955	1.274	0.855	1.044		
1983	1.085	1.119	1.098	1.126	0.687	0.488		
1984	0.912	1.011	0.920	1.014	0.657	1.292		
1985	1.038	1.232	1.038	1.226	0.890	0.754		
1986	0.819	1.023	0.823	1.029	0.907	0.582		
1987	0.798	0.911	0.801	0.917	0.791	0.648		
1988	0.788	0.804	0.779	0.805	1.291	0.977		
1989	0.851	0.899	0.827	0.886	1.614	1.268		
1990	1.211	1.376	1.197	1.378	1.372	1.163		
1991	1.141	1.292	1.100	1.287	2.664	3.161		
1992	0.836	1.128	0.828	1.129	0.996	1.493		
1993	1.126	1.360	1.112	1.351	1.182	0.910		
1994	0.898	0.906	0.891	0.908	1.224	0.896		
1995	0.943	0.930	0.935	0.930	1.317	1.334		
1996	0.965	0.898	0.965	0.899	0.892	0.575		
1997	0.896	1.061	0.886	1.053	0.742	0.726		
1998	1.009	1.109	0.999	1.106	0.879	0.916		
1999	0.971	0.912	0.961	0.919	1.134	0.764		
2000	1.019	1.318	1.029	1.333	1.121	0.834	1.223	1.579
2001	1.064	1.160	1.043	1.153	1.598	1.392	1.276	1.724
2002	0.877	1.209	0.883	1.229	1.051	0.746	1.012	1.553
2003	0.984	1.193	1.000	1.210	0.752	1.079	1.067	1.481
2004	0.733	0.779	0.727	0.791	0.875	0.673	0.856	0.886
2005	0.688	0.679	0.676	0.679	1.086	0.790	0.701	0.601
2006	0.703	0.709	0.703	0.722	0.942	0.686	0.686	0.650
2007	0.684	0.647	0.679	0.657	1.063	0.797	0.685	0.694
2008	0.762	0.665	0.768	0.676	1.072	0.805	0.803	0.813
2009	0.892	0.786	0.901	0.798	0.724	0.800	0.902	0.836
2010	0.827	0.712	0.839	0.725	0.659	0.927	0.980	0.787
2011	0.967	0.816	0.960	0.816	1.312	0.832	1.172	0.990
2012	0.794	0.641	0.812	0.652	0.841	0.865	0.997	0.693
2013	0.691	0.581	0.700	0.591	0.471	0.850	0.850	0.724
2014	0.777	0.740	0.783	0.753	0.819	1.066	0.863	0.875
2015	1.098	0.859	1.115	0.873	0.918	1.096	1.248	1.019
2016	1.164	0.920	1.196	0.937	1.035	1.392	1.619	1.220
2017	0.822	0.919	0.844	0.935	0.621	0.676	1.238	1.200
2018	0.578	0.697	0.593	0.709	0.272	0.564	0.901	0.793
2019	0.634	0.641	0.653	0.652	0.424	0.829	0.919	0.881

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*."

Contributors:

Amada Hart, Tim Miller, Lisa Kerr, Steve Cadrin, Larry Alade, Cole Carrano, Dan Hennen, and Alex Hansell

A wide range of approaches to stock assessment modeling were explored for this research track assessment. Assumed biological parameters were reconsidered, including growth, maturity, and natural mortality. Based on life history information and analyses, the assumed rate of natural mortality was increased, and the assumed maturity at age was revised.

In addition to updating the previously developed VPA, several forms of statistical catch-at-age model applications were developed for this research track assessment including conventional statistical catch at age (Age Structured Assessment Program, ASAP, Legault and Restrepo 1998), statistical catch at age with length-based selectivity and discard estimation (Stock Synthesis, SS, Methot and Wetzel 2013), and a state-space model with environmental covariates (Working Group proposed, WHAM, Stock and Miller 2021). Each modeling approach was developed to take advantage of their relative strengths, and alternative models were developed in parallel so that model improvements were shared.

Among these alternative approaches, The Working Group proposes WHAM as the basis for status determination and fishery management advice. WHAM allows for process error in selectivity to fit age composition data. General results were supported by all WHAM candidate runs as well as ASAP, SS and VPA. Among models and runs with comparable data and assumptions, results were similar.

Assumed Biological Parameters

Growth and Maturity

Explorations of spatial and temporal patterns supported a Working Group decision to account for growth empirically, using annual stock-wide age-length keys and a single length-weight relationship (see ToR2, Age Composition). Maturity at length and age observations from NEFSC surveys were explored to investigate spatial and temporal patterns (Goffe et al. 2021, Working Paper 4, Appendix A). Maturity classification is described by Burnett et al. (1989). Observed proportion of fish mature at length and age were derived for the entire stock for all years, for the Georges Bank and Gulf of Maine regions, annually, and in multi-annual periods. For the entire stock area, female length at 50% maturity was 26 cm and median age at maturity was 3.3 years, which are similar to the 26 cm and 3.0 years estimated by O'Brien et al (1993). Female median age at maturity was 3.4 years for the Gulf of Maine and 3.3 years for Georges Bank was and, respectively. Maturity at age was relatively stable over time (Figure 4.1) and among regions, with no significant differences. The Working Group decided to assume a stock-wide, constant maturity schedule based on observed proportion mature at age.

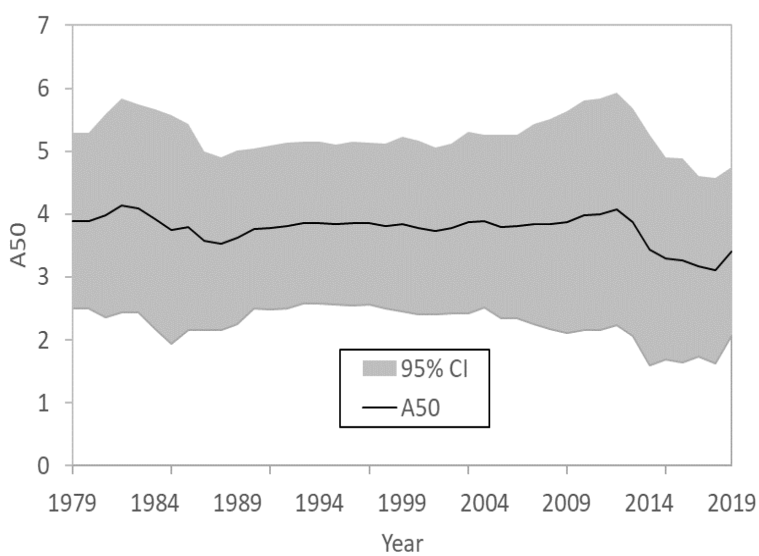


Figure 4.1. Estimates of age at 50% maturity.

Table 4.1. Estimates of maturity at age from NEFSC spring survey observations.

Age	Proportion Mature
1	0.00
2	0.14
3	0.44
4	0.75
5	0.89
6	0.96
7	0.98
8	0.99
9	0.99
10	1.00
11+	1.00

Natural Mortality

Previous stock assessments of American plaice in US waters (NEFSC 1992 … 2022b), assumed a lifetime constant natural mortality rate (M=0.2) based on relative abundance of ages 9+ from an unexploited plaice population in the Gulf of St. Lawrence (Pitt 1972). A direct estimate of M=0.18 is also available from another stock of American plaice on the Grand Banks that has had a moratorium on directed fishing and strict bycatch limits since the 1990s (Kenchington 2013). However, American plaice in US waters have significantly different life history than those in Canadian waters. Growth is faster and maturity is earlier in US waters than in Canadian waters (Lux 1970, Lange and Lux 1978, Neilson and Hurley 1986, Walsh 1994). Based on life history theory, M should also be different for American plaice in US waters than those in Canadian waters. As alternatives to previous assumptions, several life-history based M estimators were applied to the available information for American plaice in US waters (Cadrin 2021, Working Paper 15, Appendix A).

Maximum observed age of American plaice in NEFSC surveys was 24 years, sampled in spring 1983 from 719 age samples. Von Bertalanffy growth parameters were estimated from NEFSC survey data 1980-2019 for combined males and females (L_{∞}=51.46 cm, k=0.196, t_0=0.091, W_{∞}= 916 g, k_W=0.234, t_{0W}=0.000, b=3.60). Mean length at age and year was derived from the annual averages of spring and fall surveys. Age at 50% maturity was 3.3 years, and length at 50% maturity was 24 cm for combined males and females, estimated from NEFSC spring survey data. Age at 50% selectivity was 4.75 years, derived from the most recent stock assessment (NEFSC 2022b). Fishery length at age and year was derived from weight at age and the length-weight relationship by Silver et al. (2021 Working Paper 3, Appendix A). Mean length at partially selected ages (1-5) were from surveys and length at fully-selected ages (6+) were from fishery samples, because there were more samples from young fish from surveys and more samples for old fish from the fishery. The average bottom temperature for the Gulf of Maine, 8°C, was from NEFSC (2021).

These life history parameters were used to derive alternative estimates of M for plaice in US waters using several methods developed for demersal finfish based on maximum age, growth in length, growth in weight, and maturity (Table 4.2, Kenchington 2013; Cope 2021, http://barefootecologist.com.au/). Estimates of lifetime M varied by covariates. Estimates based entirely on maximum age averaged M=0.25 (range 0.22 to 0.31), and those based entirely on growth averaged M=0.32 (range 0.29 to 0.34). Estimates based on maturity (maturity or growth and maturity) were considerably greater (average M=0.58, range 0.35 to 0.80). Age-specific estimates also varied by method, but the estimates of M>1.0 for ages 1-2 American plaice in US waters from Chen and Watanabe (1989), Gislason et al. (2010) and Charnov et al. (2013) are not considered to be realistic for American plaice. By comparison, estimates of M are less than 1.0 for age-1 Atlantic herring, which is a forage species (NEFSC 2012).

In a critical review of M estimators, Kenchington (2013) compared estimates based on life history to direct estimates from age composition or tagging of unfished or lightly fished populations, including estimates for American plaice on the Grand Bank. He concluded that none of the methods provide accurate estimates for every species, but growth-based estimators performed well for many species when k was well estimated, and his estimator based on maximum age was the most consistent. Those general conclusions apply to his results for American plaice on the Grand Bank. Estimators based on growth performed well for recovering the direct estimate for Grand Banks plaice (-3% to +4%). Estimators based on maximum age or maximum age and growth had intermediate performance for Grand Banks plaice (-8% to +5%). The Charnov method performed well for Grand Banks plaice (-1%), but other estimators based on maturity or growth + maturity performed poorly (-10% to +144%).

The maximum age of American plaice in US waters was sampled after the directed fishery developed in the 1970s, so the observed maximum age could underestimate unfished longevity and overestimate M. However, estimates of M based entirely on maximum age were slightly less than those derived entirely from growth. Considering the relative consistency of estimates based on maximum age (M=0.25) and growth (M=0.32) and the poor performance of maturity-based estimates for Grand Banks plaice, the Working Group decided to assume the approximate average of estimates based on maximum age or growth (M=0.3), which is the estimated value from a commonly used method based on growth (Pauly 1980).

Table 4.2. Estimates of natural mortality for American plaice in US waters from various life-history estimators, ordered by value of estimate.

Method	Natural Mortality	Basis	Reference
ZM_CA_dem	0.09	MaxAge&Growth	Zhang & Megrey (2006)
Chen-Wat	0.20	MaxAge&Growth	Chen & Watanabe 1989
User input	0.20	GSLsurvival	Pitt (1972)
Then_lm	0.23	MaxAge	Then et al. (2015)
Hamel_Amax	0.23	MaxAge	Hamel (2015
Then_nls	0.27	Growth&Temp	Then et al. (2015)
Jensen_k 1	0.29	Growth	Jensen (1997)
Pauly_lt	0.30	Growth	Pauly (1980)
Kenchington	0.31	MaxAge	Kenchington (2013)
Jensen_k 2	0.31	Growth	Jensen (1997)
Gislason	0.32	Growth	Gislason et al. (2010)
Then_VBGF	0.34	Growth	Then et al. (2015)
Hamel_k	0.34	Growth	Hamel (2015)
Pauly_wt	0.34	Weight	Pauly (1980)
Charnov	0.37	Growth&Maturity	Charnov et al. (2013)
Ri_Ef_Amat	0.50	Growth&Maturity	Rikhter & Efanov (1976)
Jensen_Amat	0.52	Maturity	Jensen (1997)
Roff	0.67	Maturity	Roff (1984)

Stock Assessment Models

Woods Hole Assessment Model

WHAM (Stock and Miller 2021; https://github.com/timjmiller/wham), is a state-space age-structured stock assessment model. It can be configured in a similar manner to ASAP (Legault and Restrepo 1998), with fits to aggregated catch, index, and age composition data. WHAM also provides several alternative models of age composition, and can include process errors as random effects and environmental covariates. A review of the essential features for next-generation stock assessments concluded that only WHAM and a similar State-Space Assessment Method (SAM, Nielsen and Berg 2014) model random effects correctly (Punt et al. 2020).

Development of a WHAM model for American plaice began with 'bridge' runs to explore the implications of switching from a VPA model to WHAM with updated model assumptions. Runs with alternative configurations of selectivity, age composition likelihoods, and abundance at age were explored with updated data for the fleet and the survey indices previously included in the VPA (NEFSC spring and fall bottom trawl). These base runs were expanded to explore an alternative catch model that was fit to an extended catch time series beginning in 1960, and the inclusion of Massachusetts Division of Marine Fisheries (MADMF) and Maine New Hampshire (ME-NH) inshore trawl surveys, a landings per unit effort (LPUE) index, Vector-Autoregressive Spatio-Temporal (VAST) model-based indices, and a split in the NEFSC indices between Bigelow and Albatross vessel years. Finally, these runs were extended to explore environmental covariate links to recruitment and catchability. Model results and diagnostics for all runs (including exploratory runs not discussed in detail here) are available on GitHub (https://github.com/ahart1/PlaiceWG2021). This section provides an overview of modeling decisions in WHAM then describes candidate models in detail at the end of this subsection, and further details are in Hart et al. (2022, Working Paper 19, Appendix A).

All models were assessed for first and second order convergence. Models passed first order convergence criteria when their final gradient was smaller than $1e^{-10}$ and passed second order convergence if the Hessian was invertible. Akaike's Information Criteria (AIC; Akaike

1974) was used to compare models with the same likelihood structure (e.g., models fit to the same data with the same distributional assumptions). Smaller AIC scores indicated model improvement and scores within +/-2 of each other were considered equivalent, in which case the more parsimonious of the two models were selected. Mohn's rho values (Mohn 1999) were used to identify retrospective patterns in recruitment (R), spawning stock biomass (SSB) and fully recruited fishing mortality (Fbar), and smaller absolute values were preferred.

The WG preferred model framework (WHAM) underwent additional diagnostic analysis. Five principal diagnostics (convergence, AIC, retrospective consistency, residuals, and prediction skill) were used to compare model fit and performance to identify the candidate WHAM models. Model fit was examined using conventional observed minus predicted residuals as well as one-step-ahead (OSA) residuals, which are more informative for state-space models (Thygesen et al. 2017). OSA residuals should be uncorrelated and normally distributed for models that appropriately describe the system. Mean absolute scaled error (MASE) scores were calculated for possible candidate WHAM models to compare differences in model and index prediction skill (Carvalho et al. 2021; Kell et al. 2021). MASE<1 indicated models with predictive skill that is better than a naïve approach, and MASE=0.5 has twice the predictive skill as a naïve approach. Finally, simulation self-tests were performed for the three candidate WHAM runs. Each model run was used to generate 100 datasets with parameters fixed at their estimated values. Simulations then refit the model to these generated datasets to evaluate relative error in F, SSB, R, and catch estimates and model convergence rates.

Preliminary WHAM Runs

Two 'bridge' runs were conducted to configure WHAM as closely as possible to the 2019 VPA model (NEFSC 2022b). Data from the VPA (1980-2018) was imported into an ASAP data file for use in both the ASAP and WHAM bridge runs. The first WHAM bridge run used identical input data and model configuration as the 2019 VPA and ASAP bridge run and included age-disaggregated indices for all ages in the NEFSC spring and fall bottom trawl surveys. The second WHAM bridge run used the same data but implemented age-aggregated indices and age compositions for NEFSC spring and fall surveys, which is more consistent with the survey sampling designs and their precision in an integrated statistical model.

Three additional WHAM runs explored to evaluate the consequences of updating input data and changing natural mortality (M=0.3) and maturity assumptions. An additional year of catch (1980-2019), NEFSC spring (1980-2019) and fall survey (1981-2019), and weight-at-age (1980-2019) data were added to existing time series from the VPA. In addition, as described in the ToR2 section, fleet discards were estimated rather than imputed and the age classes for NEFSC indices were expanded to include an 11+ group rather than the 9+ used in the VPA.

One model run (9) was conducted with all updated data, M, and maturity assumptions. This was compared to a second model run (10) with updated data and maturity assumptions but reverted to the VPA M assumption (0.2) and a third model run (11) with updated data but reverted to assuming M=0.2 and the annually varying maturity expectations used in the VPA. Reverting to the M=0.2 scaled SSB, F, and R estimates downward compared to the run with all data updated, but changes in the maturity assumption did not scale model estimates as dramatically and this third model run performed similarly to the second run where only M was reverted to the VPA setting (Figure 4.2).

Selectivity runs began by exploring selectivity models for the fleet and indices. Fleet selectivity was initially assumed to be logistic for all runs, but both age-specific and logistic selectivity was explored for the NEFSC spring and fall indices. For runs that assumed age-specific selectivity of indices, NEFSC spring ages 4 and 5 and NEFSC fall age 4 selectivity were fixed at 1 based on a preliminary run that freely estimated selectivity for all ages (run 12). Two fleet selectivity blocks (1980-1999 and 2000-2019) were implemented in run 16 to account for differences between historic and contemporary fishing patterns associated with a series of mesh size increases in the late 1990s and early 2000s (see Appendix B) and to break the time series roughly in half to avoid assuming constant selectivity for the entire time series. These selectivity blocks improved the fit to catch data and lowered the AIC score. Age-specific selectivity estimates were dome-shaped except for the NEFSC fall survey which exhibited a higher estimated selectivity for the 11+ group than for age-10. Run 16A explored the alternative use of logistic selectivity for both indices, but this resulted in a higher AIC and likelihood contributions with few changes to model expectations (Figure 4.2). Time-varying selectivity was also explored by implementing independent and identically distributed (iid) random effects for the fleet and both indices (run 23), resulting in a better fit to the data and smaller age composition residuals.

Limiting the selectivity random effect to only the fleet (run 23A) resulted in a slight reduction in the Mohn's rho values for SSB and Fbar, but a higher AIC with little difference in the fit to catch data. Indices were not fit as well and had higher variability in selectivity for ages 2-3 when random effects for index selectivity were not included. Assuming an autocorrelated (AR1) process for selectivity random effects (run 24) did not improve model fit, with larger fleet age composition residuals and retrospective patterns.

Abundance-at-age Random Effects Runs

Model runs treating recruitment deviations as independent random effects built on run 23 with selectivity deviations treated as iid random effects. Assuming iid random effects for recruitment deviations (run 25) reduced the CVs of recruitment estimates, particularly towards the end of the time series, but otherwise performed similarly to runs without recruitment random effects. Assuming recruitment deviations are autocorrelated by year (AR1 random effects, run 26) resulted in a slightly smaller AIC and Mohn's rho for SSB and Fbar but otherwise very similar performance. Moving to a full state-space model (run 27) with random effects for survival at all ages resulted in a much smaller AIC and rho for R, SSB, and Fbar than for runs with only a recruitment random effect, but CVs for recruitment estimates were much higher.

General trends in F and selectivity were similar to runs 25-26, but the scale differed in some years and selectivity for the NEFSC fall index was less variable than in previous runs. Catch residuals for this run were smaller in magnitude for some years and slightly more evenly distributed around zero. Run 27 had improved fits to age composition data (smaller OSA residuals), but composition of older fish tended to have more negative residuals than composition of younger fish.

WHAM Runs with Extended Catch Time Series

Catch data for American plaice is available as early as 1960, but no age composition data is available prior to 1980, so a run was conducted to examine the impact of including the extended catch time series and several additional years of NEFSC spring (1968-2019 with age composition available beginning in 1980) and fall (1963-2019 with age composition available beginning in 1980) index data (run 28). The full state-space model did not converge, but a model with recruitment iid random effects (as in run 25) converged and exhibited similar model performance for years in which both models had data (1980-2019). Residuals for fits to catch and index data prior to 1980 tended to be positive, but after 1980 had similar trends and magnitudes to those from run 25. Runs 28 (1960-2019) and 25 (1980-2019) had similar fits to data and estimates of 1980-2019 SSB, F, and recruitment, but historical estimates of R and SSB were highly uncertain (Figure 4.3).

WHAM Runs with Alternative Stock Indices

Modeling NEFSC Albatross and Bigelow Surveys as Separate Series

A full state-space model (run 29) with iid selectivity random effects was fit to examine the impact of splitting the NEFSC spring and fall indices into separate Albatross (1980-2008) and Bigelow (2009-2019) indices, as recommended in the previous stock assessment (NEFSC 2022b). The fit to the fleet age composition data was slightly worse (i.e., larger likelihood contribution) and fleet OSA residuals for fit to aggregate catch data were less normally distributed than the run with the combined NEFSC indices (run 27). Age composition residuals calculated as observed minus predicted values showed slight improvement for age 1, particularly in the NEFSC fall index, but OSA residuals for fit to aggregate indices indicated similar residual patterns and magnitudes for fits to combined (run 27) and split (run 29) surveys. OSA residuals for fit to age composition data were fairly normally distributed for the fleet and aggregated spring and fall indices in run 27, with some differences in run 29 due to the split between Albatross and Bigelow years. In run 29 these residuals were reasonably normally distributed for the fleet and all indices, but indices had age-specific residual patterns with generally positive residuals for ages 5-11+ particularly later in the Albatross time series and a better mix of positive and negative for younger ages. Bigelow fall residuals were also a bit less normally distributed than for other age composition data. AIC for run 29 was smaller as were Mohn's rho estimates for SSB, Fbar, and recruitment so model selection favors this model over run 27 (Figure 4.4).

Run 29 fit to age composition in numbers at age and Albatross calibrated data during both Albatross and Bigelow years (i.e. split the combined NEFSC spring and fall indices without removing the calibration to Albatross units from 2009-2019). Three further runs (29A-29C) explored the consequences of using biomass age composition data and uncalibrated Bigelow time series (i.e., no calibration to Albatross units was used) from 2009-2019. Run 29A implemented a full state-space model with iid selectivity random effects for the fleet and Albatross spring index, but fit to biomass age composition data and used data in Bigelow units from 2009-2019. The selectivity pattern for the Bigelow fall index was more dome-shaped than in run 29, and did not have an increase in estimated selectivity from age 10 to 11+ as was previously estimated. Selectivity random effects for the Albatross spring index were less variable than in run 29, indicating that they could potentially be removed (explored in run 29B). Catchability estimates for Bigelow spring and fall and the Albatross fall indices were more similar than in run 29, although the Albatross spring catchability estimate was much lower than for the other three indices. Catch residuals were larger in magnitude for run 29A but were a more even mix of positive and negative values than in run 29. The fleet and spring indices had similar OSA age composition residual patterns but had larger maximum residuals with most differences for ages 1-3, while the fall indices had smaller maximum residuals. OSA residuals for fit to the fleet and Albatross fall index were more normally distributed (i.e. a better fit) than in run 29, with other indices showing a similar distribution in OSA residuals.

Run 29B implemented the same model as in run 29A but excluded selectivity random effects for the survey indices. Model diagnostics were very similar to run 29A. Mohn's rho values for recruitment and Fbar were slightly smaller (improved) while the value for SSB was slightly larger (worse performance). However, the delta AIC values were < 2 so these models should be considered equivalent, and the simpler model (run 29B) with selectivity random effects only for the fleet was the preferred run for further development. Run 29B-1 implemented the same model but made slightly different assumptions for index effective sample size based on those used in ASAP run 51a. This adjustment had very little impact on model estimates or fit.

Run 29C implemented the same model as in run 29A but fit to biomass age composition data and used Albatross units for both the Albatross and Bigelow years to identify whether the improvements in run 29A were attributed to the switch to biomass age composition data, or to the switch from Albatross to Bigelow units for 2009-2019. Selectivity estimates for age 11+ in both fall indices were higher than age 10 so the improvement in the Bigelow fall index in run 29A and 29B appears to be attributable to the switch to Bigelow units. The selectivity random effects for the Albatross spring index were less variable in runs 29A and 29C than in run 29, indicating that this random effect may be less informative when biomass units were used. Run 29B did not include this selectivity random effect and still resulted in improved Mohn's rho values over run 29, supporting the decision to exclude selectivity random effects for the indices in models with logistic-normal age composition likelihoods. Catch residuals were more evenly distributed around zero early in the time series for run 29C as was also seen in runs 29A and 29B but not in run 29. This suggests that this improvement is due to the switch to biomass units rather than the switch to Bigelow units. OSA age composition residuals for the fleet and spring indices had similar patterns but larger maximum values compared to run 29. The fall indices followed similar patterns but had slightly smaller maximum values. The patterns and magnitude of these residuals were very similar to those in run 29B. The Working Group decided to pursue models using NEFSC survey indices in Albatross and Bigelow units rather than Albatross calibrated units for the Bigelow survey.

Incorporating State Inshore Survey Indices

Models were fit to combinations of NEFSC, inshore state trawl (MADMF and ME-NH), LPUE, and VAST indices built upon run 27 to explore other indices of abundance available for American plaice. A full state-space model was fit to NEFSC and MADMF spring and fall indices (run 30) but did not include selectivity random effects for the indices because the model failed to converge (i.e., Hessian was not invertible) when they were included. The exclusion of selectivity random effects was expected to have the biggest impact for ages 1-3 of the NEFSC spring index, and these ages tended to have the largest OSA age composition residuals. The fit to fleet and NEFSC spring age composition was worse, with only minor improvement in fit (i.e., smaller maximum residuals) to the NEFSC fall index. Residuals for fit to the NEFSC indices were consistently positive at the end of the time series while residuals for fit to the MADMF indices were consistently negative. This pattern was also observed for runs that included ME-NH indices. AIC was not comparable to run 27, but the Mohn's rho value for F was smaller than run 27, while values for SSB and Fbar were slightly larger.

A full state-space model with selectivity random effects converged when fit to both NEFSC and ME-NH spring and fall indices (run 31). The ME-NH inshore trawl survey fully selected much younger ages (ages 2 and 1 for the spring and fall respectively) compared to the NEFSC indices (ages 4-5 in the spring and age 4 in the fall) but had little impact on residual patterns for ages 1-2 in run 31. Fleet age composition OSA residuals were slightly smaller in run 31 than in run 27 but followed a similar pattern, while the NEFSC spring and fall indices had larger residuals, particularly for ages 1-3. Fits to the ME-NH age composition followed a similar pattern with large residuals observed in many years for ages 1-3 and generally smaller residuals for older ages. Mohn's rho values for R and SSB were smaller than in run 27, but the value for Fbar was slightly larger. State inshore surveys data were also considered in the development of integrated indices from multiple surveys (below).

Incorporating Fishery Dependent Indices (LPUE)

A model with recruitment random effects converged when fit to NEFSC and LPUE indices (run 32). The selectivity pattern for the LPUE index was specified to mirror the fleet's estimated selectivity-at-age. Fit to catch data was worse than the run (run 25) that only fit to NEFSC indices (i.e., larger extreme catch residuals, less normally distributed OSA residuals), but the fit to NEFSC indices was similar. OSA residuals for fit to age composition data were similar between run 25 and run 32 for both the fleet and the NEFSC indices. Mohn's rho values for SSB and Fbar were smaller in run 32 than in run 25, but the rho value for R was larger. An additional sensitivity run (32A) was conducted to fit to the same indices but implemented a catchability random effect for the LPUE index. This resulted in slightly more normally distributed OSA residuals for fit to the fleet and aggregate indices, and much more normally distributed OSA residuals for the LPUE index. Despite these improvements the inner-quartile range of MASE scores for this run was much broader than for runs 27 and 29-29H and at a prediction horizon of 1-year the mean is greater than 1 (on average less accurate than a mean approach). The MASE score describing the prediction skill of the LPUE index in run 32A also had a mean above 1 suggesting that this index does not improve the prediction skill of the model. Further development of models incorporating LPUE without inshore surveys was not pursued.

Incorporating Multiple Inshore Surveys and Fishery Dependent Indices

Three runs explored the consequences of including multiple state indices in the assessment model. Run 33 fit a full state-space model to NEFSC, MADMF, and ME-NH spring and fall indices but did not include selectivity random effects. Fit to the fleet and NEFSC spring index were similar to run 30 that fit only the NEFSC and MADMF indices, but the fit to the NEFSC fall index was worse, particularly for ages 1-2. Fit to the MADMF spring index was worse with very large OSA residuals for ages 1-2 in some years, and fit to the MADMF fall index had slightly smaller maximum OSA residuals. The ME-NH indices had very large residuals for ages 1-3. Despite these differences, the magnitude and pattern of SSB and F estimates were very similar to run 30, with more variability in recruitment estimates between runs (Figure 4.6). Mohn's rho values for R, SSB, and Fbar in run 31 were all larger than for run 30.

Two runs (34A and 35A) explored models fit to a combination of NEFSC, state, and LPUE indices. Similar to run 32, which fit to both the NEFSC and LPUE indices, the LPUE index selectivity was specified to mirror the fleet selectivity estimates for these runs. Both runs had a poorer fit to catch data (i.e. larger catch residuals and less normally distributed aggregate catch OSA residuals) compared to runs that fit to only the NEFSC and state indices (runs 31 and 30 compared, respectively, to 34A and 35A). OSA residuals for fit to the aggregate LPUE index were not very normally distributed for either run, indicating that these models are not entirely appropriate for this data. Age composition OSA residuals for the NEFSC spring and fall indices were both slightly more normally distributed than in runs 31 and 30, indicating a slight improvement in fit to these indices. Mohn's rho values for SSB and Fbar were smaller in run 34A than in run 31 but larger for R. In contrast, all rho values for run 35A were smaller than in run 30. These runs converged when a recruitment random effect was specified but did not converge when random effects for all numbers-at-age were implemented, thus they were not considered further in this analysis.

Incorporating Model-Based Indices (VAST)

Run 37E replaced the design-based indices with fits to several iterations of model-based Vector-Autoregressive Spatio-Temporal (VAST) indices. Earlier iterations (37-37D) were developed using preliminary VAST index data and are thus not discussed in detail here. VAST analysis was used to generate both spring and fall indices based on raw NEFSC, MADMF, and MENH trawl data in uncalibrated units (i.e., no survey units calibrated to Albatross units), and covered numbers-at-age for 1-11+. Run 37E fit to these updated VAST indices and assumed logistic selectivity for the fleet and both spring and fall indices based on a preliminary run that freely estimated selectivity-at-age. OSA residuals for fit to the aggregate catch were less normally distributed than in run 29B, but OSA residuals for fit to fleet age composition data were generally more normally distributed with the exception of some very large residuals that were attributed to age 10 and were generally large and positive. OSA residuals for fit to aggregate indices were more normally distributed in the fall than in the spring and OSA residuals for fit to age composition data showed a similar trend as the fleet, with age 10 residuals often very large and positive, but otherwise a fairly normal distribution. Observed-predicted residuals for the indices showed a better fit than prior VAST runs (e.g. 37B), indicating that poor fit in those runs could be a data issue. This run could not be directly compared to other runs via AIC due to differences in the input data, but Mohn's rho values for R, SSB, and F were larger than those in runs 29B and 29F.

Based on model diagnostics from all explorations of alternative stock indices (single vs. separate NEFSC Albatross and Bigelow series; original vs. calibrated units; aggregate biomass vs. aggregate abundance indices; inshore state surveys; separate NEFSC and state trawl survey series vs. a combined-survey VAST index; fishery dependent LPUE; and combinations of these alternatives), The Working Group decided that the preferred WHAM runs fit tol NEFSC Albatross and Bigelow surveys as separate aggregate biomass indices in their uncalibrated units (i.e., Bigelow not calibrated to Albatross units).

WHAM Runs with Alternative Age Composition Models

Two additional sets of runs explored alternative age composition models (logistic-normal, Dirichlet-multinomial) that more explicitly weight indices (Fisch et al. 2021) with the aim of improving model performance. These runs were explored to determine if OSA residual patterns for fit to age composition data in runs 29-29C could be resolved. Runs 29F-29F5 implemented a logistic normal age composition model that estimates an additional weighting parameter for each index and treated zero observations as missing. Run 29H was an exploratory run that implemented a Dirichlet-multinomial age composition model to directly estimate effective sample size. These runs were fit to split Albatross and Bigelow indices (i.e., Bigelow years not calibrated to Albatross units) with aggregate index data in biomass units as in runs 29A-29C.

Run 29F only converged when a wider range of selectivity parameters were fixed at 1 compared to run 29B, and included age 11+ for the Albatross fall index. F estimates tended to be lower than estimates from run 27 and other split runs (29-29C) and had a slightly different trajectory but peaks and valleys were similar. R estimates tracked the estimates from other runs fairly well and tended to be on the higher side of the range. SSB estimates were similar or slightly higher prior to 2000 but fell between the estimates for run 29 (on the higher end) and run 27 (on the lower end). CVs around estimates of F were more variable over time and were higher for estimates of R. OSA residuals for the fleet and aggregate spring indices were similarly or slightly more normally distributed and fall indices were slightly less normally distributed than in run 29F compared to 29B. Age composition OSA residuals were more normally distributed for the fleet and all indices, although their magnitude was slightly larger. Some years had residual patterns where residuals for all or most ages were consistently positively or negatively biased, but there was not a pattern of generally positive residuals in ages 5-11+ as was observed in run 29-29C.

Runs 29F1-29F5 varied starting selectivity estimates and run 29F2 explored the inclusion of a selectivity random effect to try to improve the estimation of age 11+ selectivity for the Albatross fall index. Runs 29F1-29F5 had larger AIC values than run 29F so they were not considered an improvement on this prior run based on this statistic alone. However, runs 29F2, 29F4, and 29F5 had other model improvements that qualified them for further consideration.

Run 29F1 used the selectivity estimates from run 29B as starting estimates for the run, and initially tried to estimate selectivity at older ages freely in a preliminary run. However, the final version of this run required that age 11+ selectivity for the Albatross fall index was fixed at 1 in order to converge. This run had minor differences in selectivity estimates for both spring and the Bigelow fall index but had little effect on estimates for the Albatross fall index.

Run 29F2 used the selectivity estimates from run 29B as starting estimates for the run and implemented a selectivity random effect for the Albatross fall index. When this initial run failed to converge, the Albatross fall age 11+ selectivity was fixed at the starting estimate (0.58) and the revised run converged. Although the AIC value was slightly larger than in run 29F, the OSA residuals for fit to the aggregate fleet, and both spring and fall Albatross indices were more normally distributed, indicating that this model more appropriately fits to this data. This run was one of three candidate runs.

Run 29F3 used the selectivity estimates from run 29B and fixed age 11+ selectivity for the Albatross fall index at this starting estimate (0.58). Estimates of selectivity-at-age followed similar patterns to run 29F but were generally smaller in magnitude for ages that were not fully selected in run 29F-3. Other diagnostics were similar to those in run 29F.

Run 29F4 reverted to using 0.5 as starting estimates for selectivity-at-age and used an initial run to identify a single age for each index to fix at full selectivity. This resulted in a final run that fixed Albatross spring age 6, Bigelow spring age 5, Albatross fall age 4, and Bigelow fall age 3 selectivity at 1. This run successfully converged while freely estimating the selectivity for Albatross fall age 11+, but this estimate was near 1 and had wide confidence bounds that spanned from 0 to 1. Although the AIC value was larger than for run 29F (and thus less preferred based on this metric), the free estimation of selectivity for the Albatross fall plus group was preferable as it more appropriately represents the uncertainty in this estimate than in runs that fixed this parameter. Uncertainty in selectivity estimates for Albatross spring age 5, Bigelow spring age 4, and Bigelow fall age 4 were also highlighted in this run with confidence bounds around estimates spanning from near 0 to near 1. This run was one of three candidate runs.

Run 29F5 was specified as in run 29F4 but explored two approaches to freely estimate the Albatross fall age 11+ selectivity at value farther from 1 (full selectivity). A preliminary run freely estimated Albatross fall selectivity-at-age but forced the estimate for age 10 and 11+ to match, resulting in an estimate near full selectivity (0.98). Because this did not lower the estimate of selectivity for Albatross fall age 11+, the full run instead implemented an AR1 random effect on age for this index (age-varying rather than time-varying). This change resulted in slightly lower selectivity estimates than in run 29F4 for the Albatross fall index except for age 4 which was fixed at full selectivity. The OSA residuals for fit to the aggregate fleet and index data were similarly or slightly more normally distributed than in run 29F4, and the Albatross fall residuals in particular were more normally distributed. Mohn's rho and AIC values were larger than those for run 29F4. This run was one of three candidate runs.

A single run (29H) explored a Dirichlet-multinomial age composition likelihood model. This run had difficulty calculating the OSA residuals so they were not used to compare with other models. The AIC value was not comparable to other runs due to the change in likelihood structure, but Mohn's rho values were comparable, with SSB and Fbar rho values smaller than in runs 29F, 29F2 and 29F4, and R rho value larger than in these three runs. Mohn's rho values for run 29 were always smaller than those for run 29F5.

The Working Group decided to proceed with candidate models that implemented a logistic normal age composition based on the improvement in index age composition residuals which were positively biased for ages 5-11+ in prior runs and more normally distributed fleet OSA residuals. The Working Group requested runs 29F1-29F5 to explore options to improve the estimation of the Albatross fall age 11+ selectivity. Run 29F2 fixed the Albatross fall age 11+ selectivity at the estimate from run 28B but was selected as a candidate because OSA residuals for fit to aggregate fleet and both spring and fall Albatross indices were more normally distributed. Runs 29F4 and 29F5 were selected because they successfully estimated the Albatross fall age 11+ selectivity freely, but each had trade-offs that made it difficult to select as a single candidate run. Run 29F4 was equivalent to run 29F2 according to AIC, fixed fewer ages at full selectivity (more estimates informed by data) and freely estimated age 11+ selectivity, but several ages had large confidence bounds around estimates, including the Albatross fall age 11+ selectivity estimate. Run 29F5 had more normally distributed OSA residuals for the Albatross fall index, freely estimated age 11+ selectivity for this index at a lower value than run 29F4, and had the smallest AIC value but the largest Mohn's rho values of the candidate runs.

WHAM Runs with Environmental Covariates

Results from ToR1 identified potential drivers of recruitment and survey catchability, so WHAM runs were developed to link those parameter estimates to sea surface temperature anomalies (SST anomaly), bottom temperature anomalies (BT anomaly) and the North Atlantic (NAO), and Atlantic Multidecadal Oscillations (AMO). A preliminary run of each environmentally linked model was conducted to fit to the environmental covariate data. These model runs were fit without a specified effect on stock dynamics to allow comparisons via AIC to models with the covariate effect specified (runs 39, 41, 42, 43, 44, 46, 47, 48, 49, and 50). OSA residuals for fit to the environmental covariates were under-dispersed for two of these runs when no effect was specified (42 fit to NAO and 48 fit to AMO) and estimated environmental covariates had much narrower distributions than the observed data (Figure 4.7). These results indicate a modeling error that may also impact AIC calculations and thus make comparisons of these models with AIC inappropriate. Although runs 42 and 48 were most strongly impacted by the modeling error, its existence lowers our confidence that AIC and Mohn's rho values are accurate for other runs that fit to environmental covariate data without specifying a covariate effect. Within the timeframe of this research track assessment, we were unable to resolve the underlying modeling problem and could not make confident comparisons between runs that did and did not specify an environment covariate effect, so our advice is to exclude environmental covariates from candidate models at this time. However, we outline some general conclusions drawn from the runs that did specify an environmental covariate effect that may warrant further consideration in future analyses.

Both random walk and AR1 environmental processes were explored for covariates that affected either recruitment or catchability. In most cases there were few or no differences between models based on what process was implemented, but the choice of environmental process did impact the sign and magnitude of bottom temperature effects on catchability (runs 41A and 47A). Catchability generally had a positive relationship with bottom temperature (except for the Bigelow spring index when an AR1 process was assumed). This conclusion aligns with the results of preliminary analyses from ToR1 (see ToR1 section) which found decreasing catchability as plaice moved into deeper, colder water and highlights these runs as key models to revisit in future analyses. Changes in the magnitude or sign of the environmental effect could indicate a mis-specified environmental process in one of these models but could also be attributed to a mis-specified effect on one or more of the indices. We recommend exploring the consequences of specifying an environmental effect on catchability for a subset of the available indices in future analyses to evaluate the later source of misspecification. Furthermore, where runs do not have variable performance based on the environmental process model, there was a working group recommendation to implement an AR1 process because the variance of projections asymptotes for this process rather than going to infinity as for random walk processes. This is not expected to influence model results but has consequences for model projections.

In contrast to the models with environmental effects on catchability, runs that affected recruitment generally had a negative relationship (e.g. run 46 estimated larger recruitments as sea surface temperature anomalies became more negative). This result contradicted the preliminary analyses from ToR1 (section 1) that suggested increasing recruitment as sea surface temperature increased. This conflict further justified the selection of candidate models without environmental effects on recruitment in this research track, but this conclusion should be reassessed in future analyses (after addressing the underlying modeling errors) since the exploratory analyses from ToR1 strongly suggested an environmental relationship with recruitment.

To ensure that future model explorations with environmental covariates are comparable via AIC, we also recommend including all available covariates in each model but only specifying links for those that are being actively evaluated (i.e., similar to the current approach of fitting to the covariates without specifying an effect, but fit to multiple covariate data sets at once).

Candidate WHAM Runs

Three models emerged as candidates (29F2, 29F4, and 29F5) and were compared in greater detail here. All three candidate models fit to aggregate (kg/tow) and age composition (abundance) data for four indices and a single fleet. NEFSC bottom trawl survey data was split into separate Albatross and Bigelow indices for both the spring and fall, and Bigelow data was uncalibrated (i.e. no calibration to Albatross units was used). Natural mortality was fixed at 0.3 following a revision from 0.2 in the prior VAST assessment and a constant maturity-at-age schedule was implemented in all three runs. No stock-recruit relationship was estimated, so recruitment was assumed to be random about an estimated mean. Random effects were implemented for all numbers-at-age, to allow variable survival for each age class and year. Candidate models estimated a random effect for fishery selectivity, allowing for time-varying selectivity. All candidate runs assumed logistic selectivity for the fleet and estimated age-specific selectivity for the four indices. A logistic-normal age composition model was implemented for each of the candidate runs.

Selectivity of Candidate WHAM Runs

All three candidate runs have a similar model configuration but sought to address uncertainty in age 11+ selectivity for the Albatross fall index in different ways, and consequently had minor differences in selectivity estimates (Figure 4.8). Run 29F2 implemented a selectivity random effect for only the Albatross fall index and fixed age 11+ selectivity at the estimated value from run 29B (0.58). Selectivity was fixed at 1 for ages 5 and 6 in the Albatross spring index, age 5 in the Bigelow spring index, age 4 in the Albatross fall index, and ages 3 and 4 in the Bigelow fall index, with starting estimates set to estimated values from run 29B except for Albatross spring age 4 and Bigelow fall age 5 selectivity which had starting estimates set to 0.5.

Run 29F4 only fixed a single age at full selectivity (age 6 in Albatross spring, age 5 in Bigelow spring, age 4 in Albatross fall, and age 3 in Bigelow fall), and set index selectivity starting estimates for all other ages to 0.5. No selectivity random effects were included for the indices. Selectivity for Albatross fall age 11+ was freely estimated in this run but had an estimate near 1 with large confidence bounds (Figure 4.8).

Run 29F5 was specified identically to run 29F4 but modeled age-specific selectivity for the Albatross IV fall index as an AR1 random effect (not time varying, just age varying), aside from the age fixed at full selection (Figure 4.8).

Diagnostics of Candidate WHAM Runs

All three candidate runs met first and second order convergence criteria. AIC scores for run 29F2 and 29F4 were within +/-2 of each other so these runs should be considered to have equal goodness-of-fit and parsimony, so other diagnostics should inform the selection between the two. AIC for run 29F5 was not comparable to the other candidate runs because of the difference in likelihood structure (AR1 random effect), so other diagnostics should also be used to select between this model and the other candidates. None of the candidate runs had strong retrospective patterns, and consequently there were only minor differences in Mohn's rho values between these runs (Figure 4.9). Run 29F5 had the largest rho values out of the three runs, and run 29F4 had the smallest values for R and SSB while run 29F2 had the smallest rho value for Fbar. The distribution of OSA residuals varied slightly between candidate runs for fit to both aggregate and age composition data, but residuals were generally normally distributed (Figure 4.10). Runs 29F4 had a slightly different distribution of OSA residuals for fit to the Albatross fall index than run 29F2 and 29F5 which respectively fixed or estimated age 11+ selectivity at a lower value.

On average the Bigelow fall index was more accurately predicted than the spring index according to MASE scores, but confidence bounds around the fall MASE scores were larger (Figure 4.11). MASE scores for all three candidate runs decreased as the prediction horizon increased from 1-3 years (typical management horizons) and started to increase for Bigelow fall horizons greater than 4 years (Figure 4.12). Prediction skill should decrease as prediction horizon increases, but it is possible to see the opposite trend for models that appropriately describe the data or when there is a trend in the index itself. In such cases longer horizons provide more data for use in predictions and are thus improved (smaller MASE scores) over shorter horizons that use less data for predictions. On average runs 29F4 and 29F5 had slightly lower MASE scores than run 29F2 over a typical management horizon (1-3 years), indicating slightly improved predictive skill for these two candidates (Figure 4.13). However, the differences are sufficiently small that this statistic alone does not strongly point to the selection of one candidate over another.

Self-tests for Candidate WHAM runs

Run 29F2 had the lowest convergence rate (67%) of the three candidates, with all simulations for runs 29F4 and 29F5 converging. Median relative errors for SSB and recruitment were greater than 1, indicating the tendency to overestimate these values (Figures 4.14- 4.15). In particular, there was more variability in the SSB relative error for runs 29F4 and 29F5 towards the end of the time series than was observed in run 29F2. The relative error for recruitment was larger in magnitude and had larger interannual differences than other relative errors. Runs 29F4 and 29F5 had median relative recruitment errors closer to 1 than run 29F2, but run 29F4 had a wider spread than run 29F5.

Catch relative errors were slightly larger than 1 so there is a minor tendency to overestimate catch, but the scale of this overestimation was much smaller than for SSB or recruitment (Figure 4.16). Of the four relative error metrics, catch relative error had the smallest magnitude, very low interannual variation, and had median values closest to 1. Relative errors in the last 10-15 years of the time series had very little differences between the middle 50% and 80% of catch relative errors across simulations and were very close to 1. Relative error for fishing mortality was the only metric with median relative errors smaller than 1, indicating the tendency to underestimate F in self-tests (Figure 4.17). This pattern was particularly apparent over the last 10 years of the time series for all three candidate models, where median values were mostly less than 1.

Results from Candidate WHAM Runs

All three candidate models estimated similar trends in R, Fbar and SSB and had similar trends in CVs around these estimates but the scale varied slightly between models (Figure 4.18). Runs 29F4 and 29F5 had similar estimates and trends in CVs, but had generally higher F estimates, slightly higher R estimates, and slightly lower SSB estimates compared to run 29F2. This corresponded to similar CVs around F estimates in all runs, but run 29F2 had slightly higher CVs around R estimates early in the time series and consistently higher CVs around SSB estimates compared to runs 29F4 and 29F5. Considering the close similarity in results among candidate runs, and slightly better model performance (retrospective consistency of R and SSB, prediction skill, high convergence rate and low error rate of predicted SSB in self-tests), the Working Group proposes WHAM run 29F4 as the basis for status determination, projection and fishery management advice.

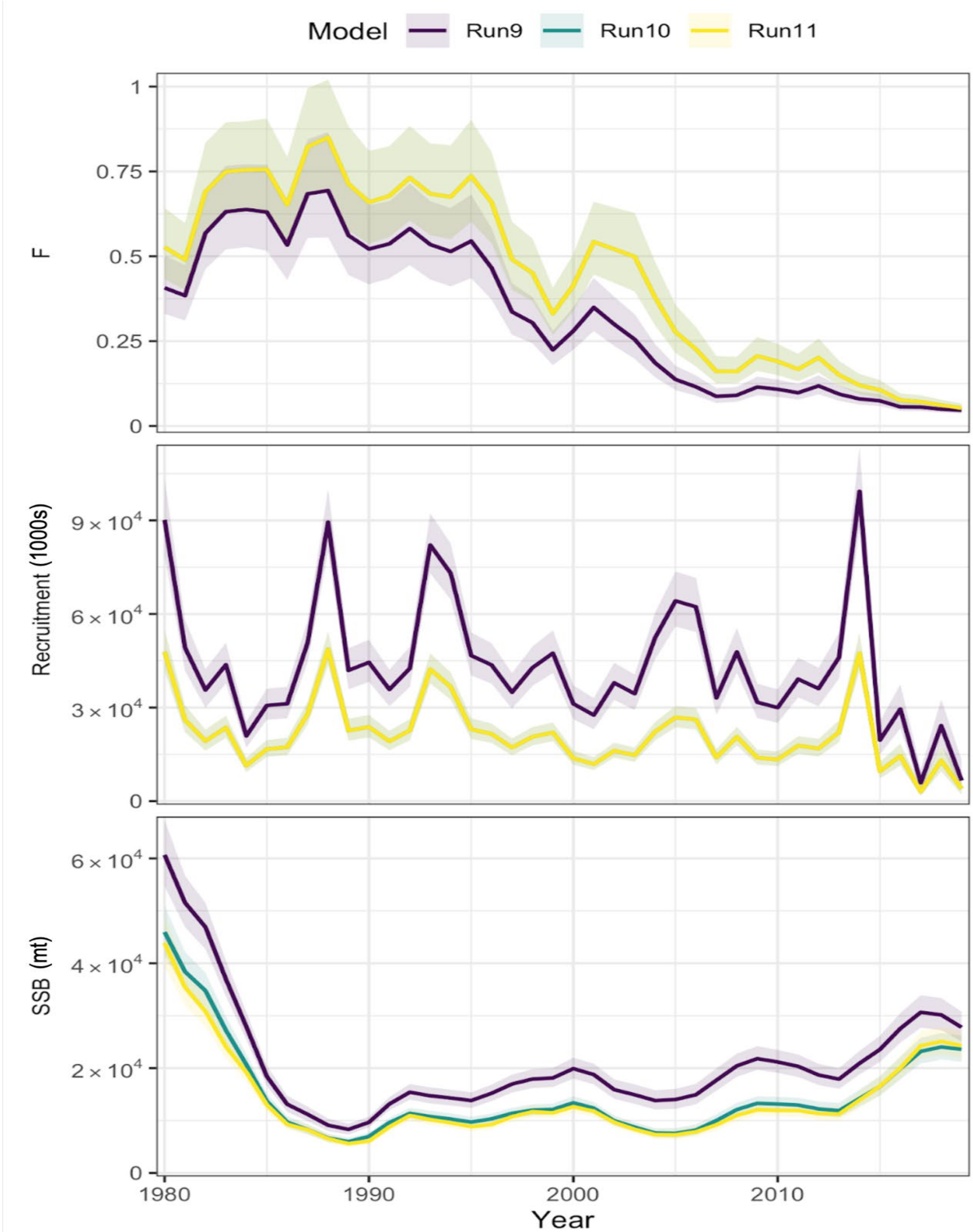


Figure 4.2: Model estimates of spawning stock biomass (SSB), fully-selected fishing mortality (F), and recruitment for runs where all data, natural mortality (M), and maturity assumptions were updated (run 9, purple line), only data and maturity were updated (M remained at 0.2; run 10, green line), and a run where only the data were updated but M and maturity assumptions remained identical to those in the VPA (run 11, yellow line).

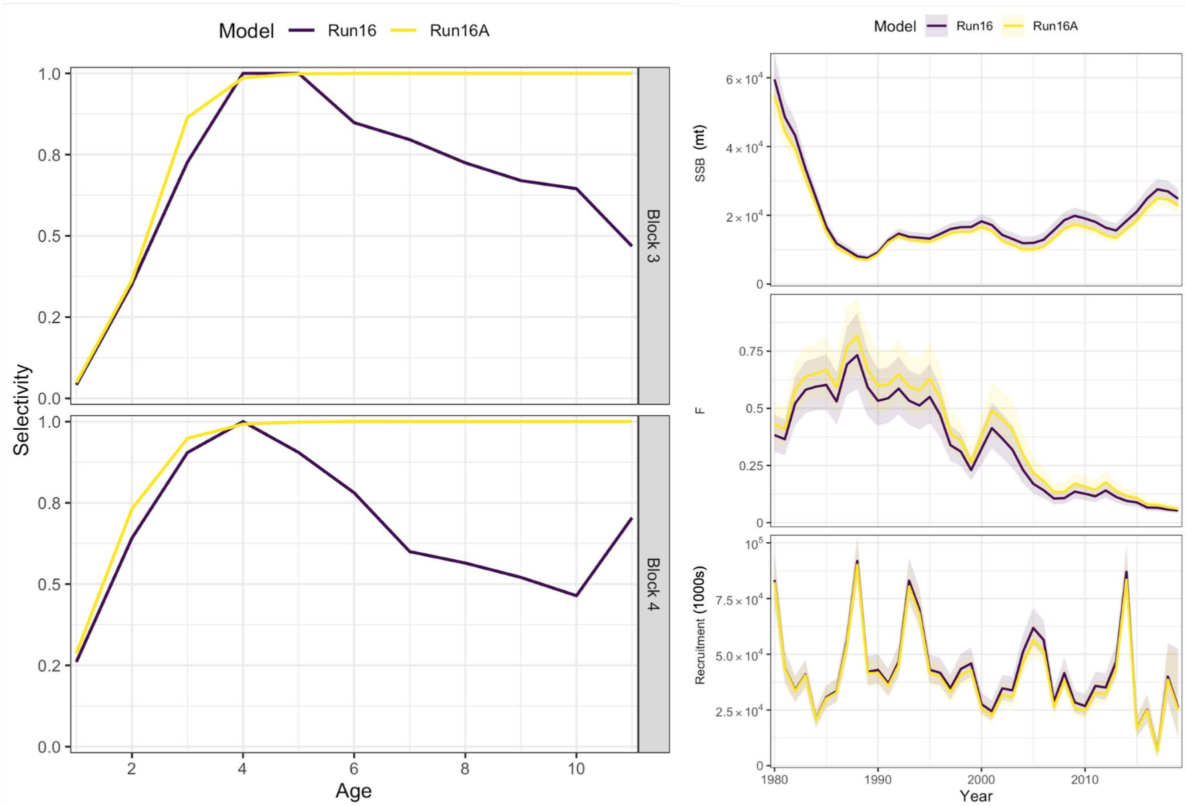


Figure 4.3: Selectivity (left) for NEFSC spring (block 3) and fall (block 4) indices and model predictions of SSB, F, and recruitment (right). Run 16 assumed age-specific selectivity for both spring and fall NEFSC indices while run 16A assumed logistic selectivity for these indices.

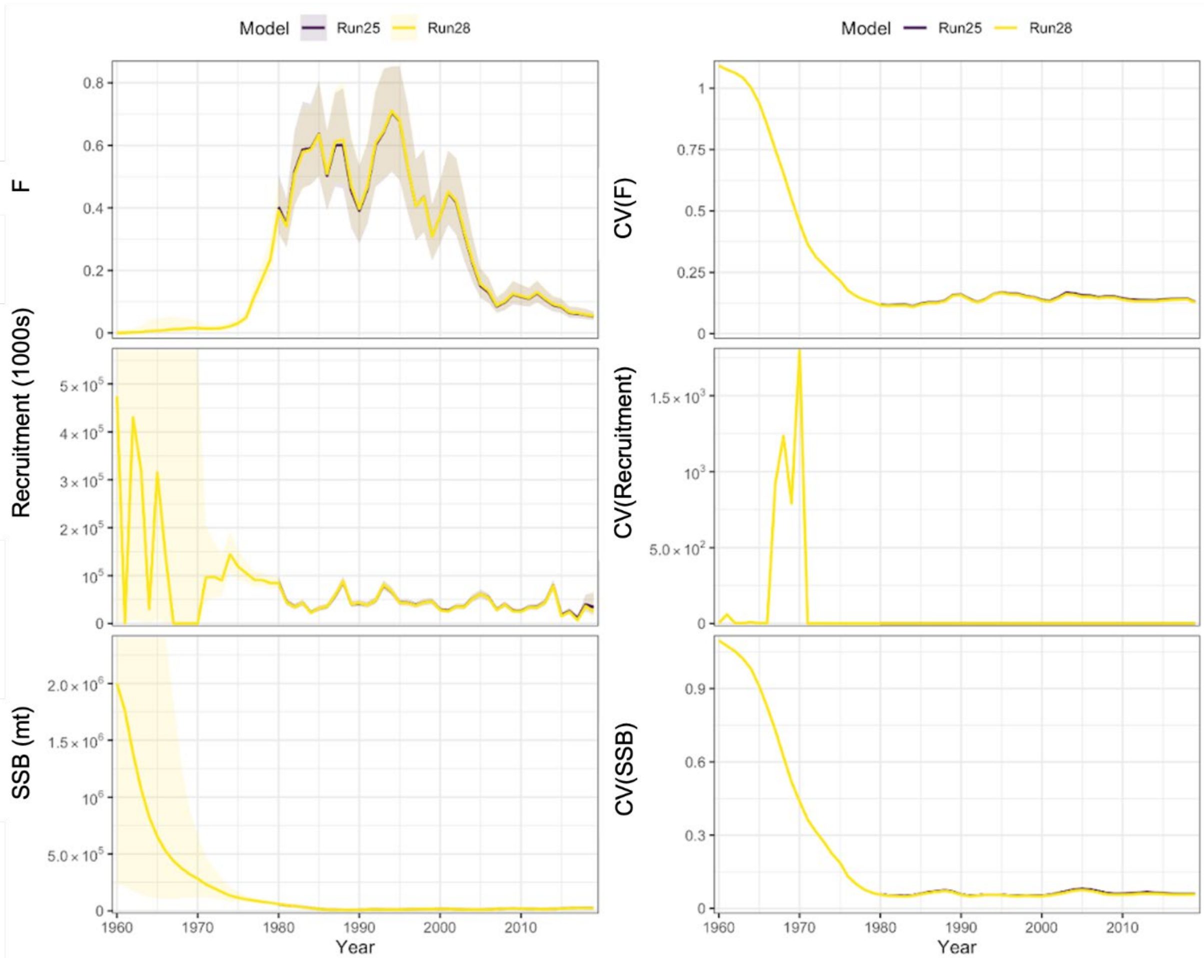


Figure 4.4: Model estimates of spawning stock biomass (SSB), fishing mortality (F) and recruitment (left), corresponding CVs around these estimates (right). Both run 25 and 28 implemented recruitment and selectivity random effects, with run 28 fit to catch data beginning in 1960 rather than in 1980 as in run 25.

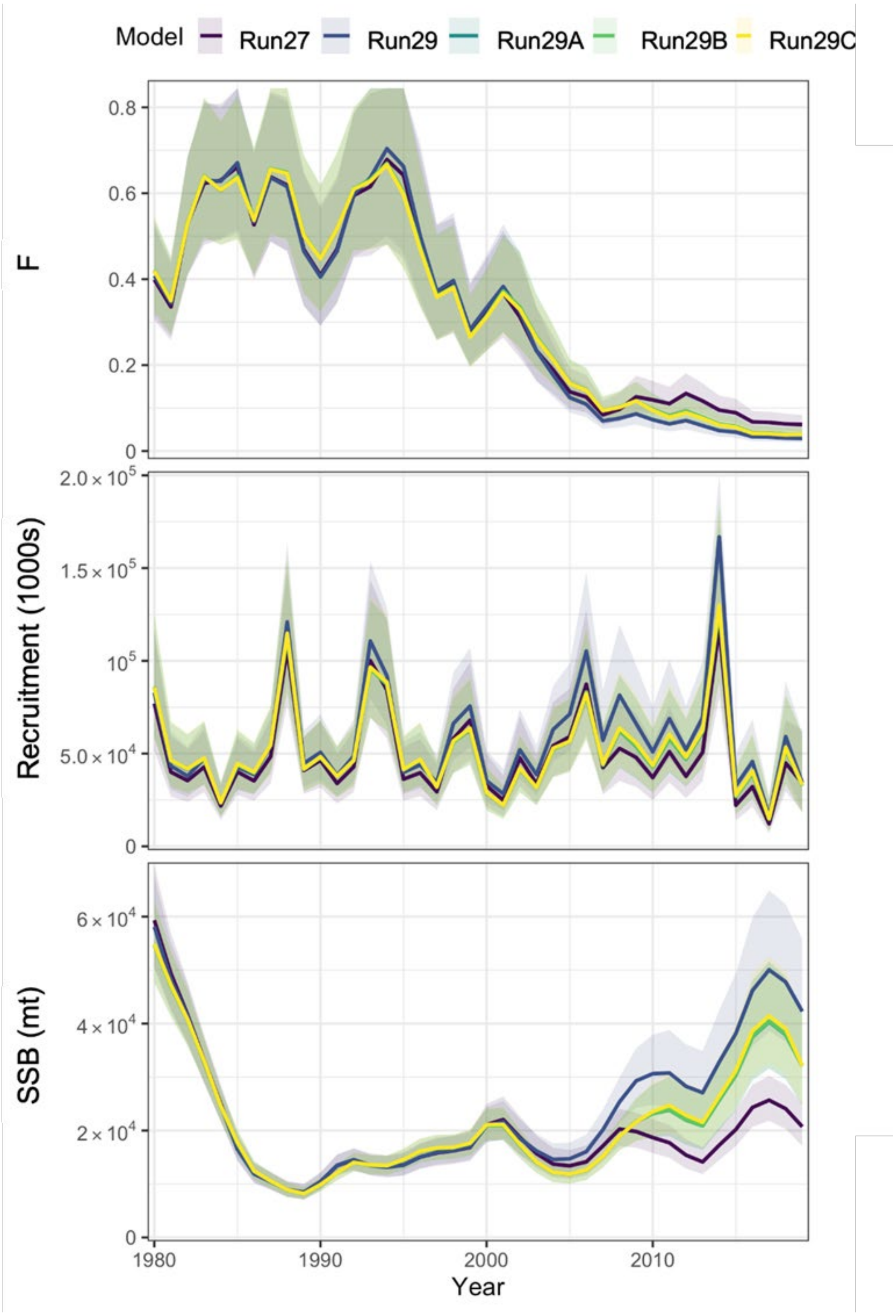


Figure 4.5: Model estimates of spawning stock biomass (SSB), fishing mortality (F) and recruitment. Run 27 implemented a full state-space model with iid selectivity random effects fit to the full NEFSC spring and fall indices (1980-2019), while runs 29-29C were fit to split Albatross (1980-2008) and Bigelow (2009-2019) spring and fall indices. Run 29 fit to age composition in numbers-at-age and Albatross survey units, runs 29A-B fit to biomass age composition and Bigelow survey units from 2009-2019 but 29B excluded index selectivity random effects, and run 29C fit to biomass age composition but used Albatross survey units from 1980-2008 and 2009-2019.

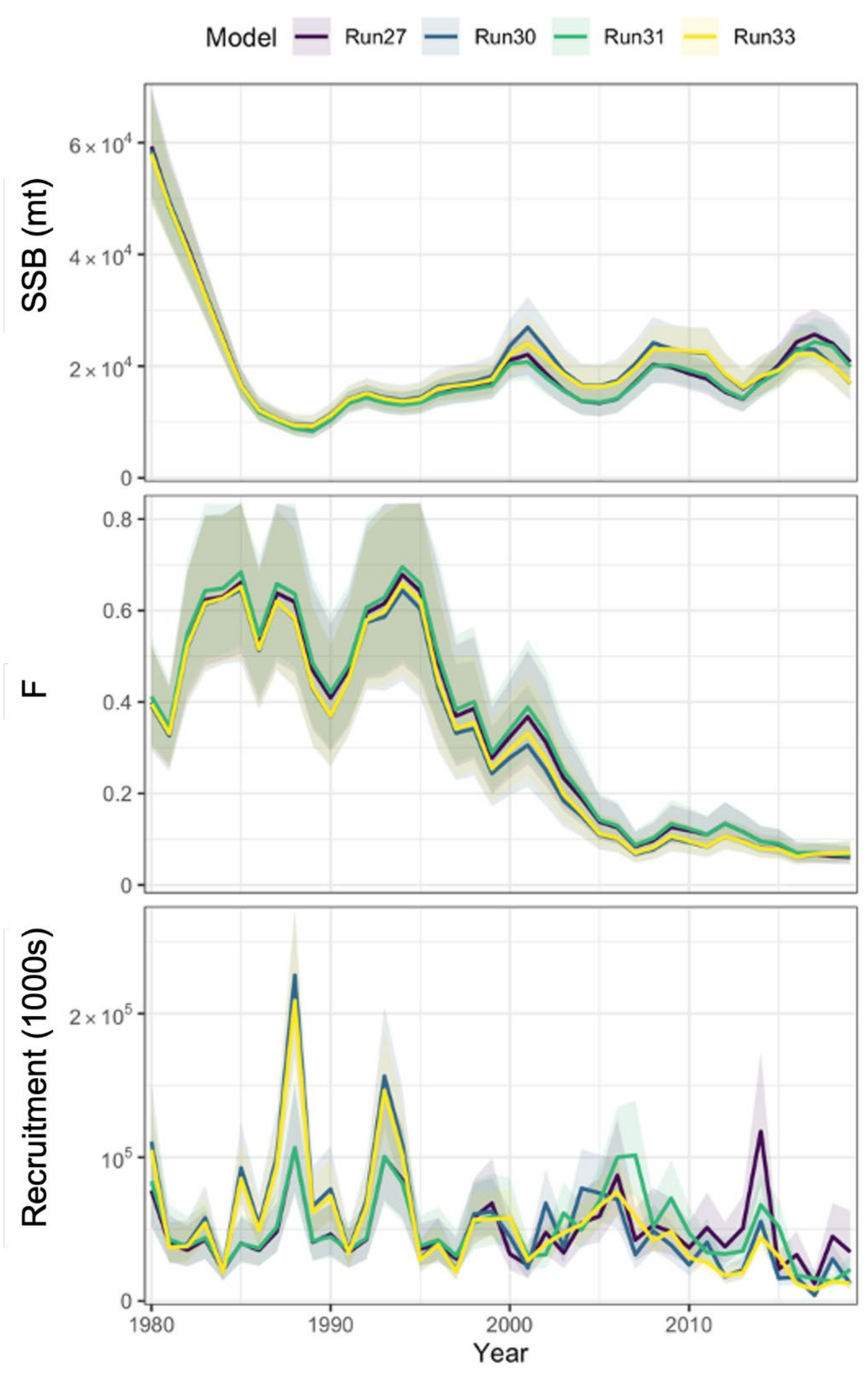


Figure 4.6: Model estimates of spawning stock biomass (SSB), fishing mortality (F) and recruitment. Run 27 implemented a full state-space model with iid selectivity random effects fit to the full NEFSC spring and fall indices, run 30 fit a full state-space model to NEFSC and MADMF indices without selectivity random effects, run 31 fit a full state-space model to NEFSC and ME-NH indices with iid selectivity random effects, and run 33 fit a full state-space model to NEFSC, MADMF, and ME-NH indices without selectivity random effects.

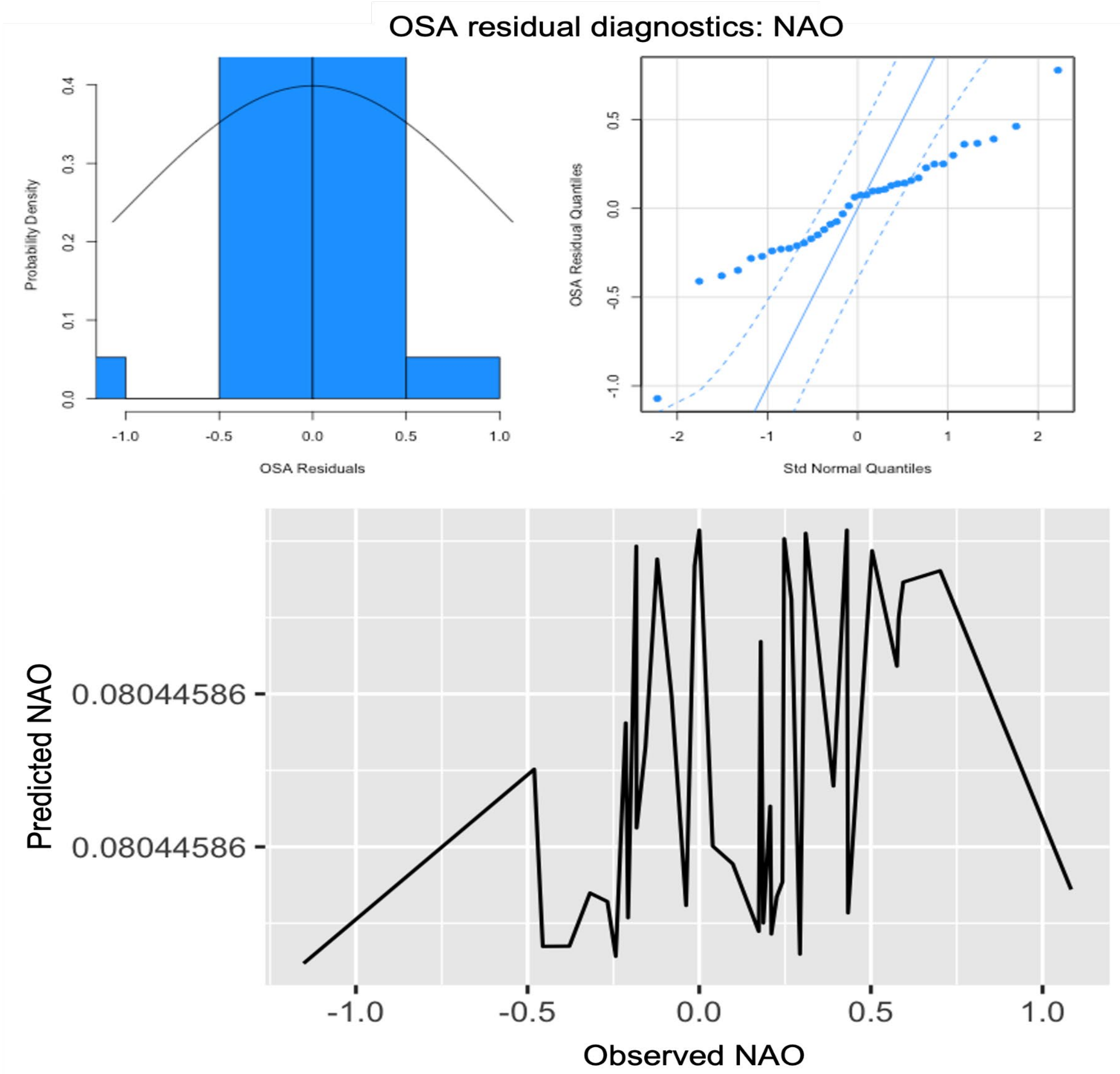


Figure 4.7: Example of over-dispersed one-step ahead (OSA) residuals for fit to a North Atlantic Oscillation (NAO) covariate (top) and comparison of model predicted values for this covariate vs. observed values for this covariate (bottom). Both results indicate a modeling error in WHAM when an model is fit to an environmental covariate without an effect on stock dynamics specified (i.e. to establish a base model that is comparable via AIC to runs with an effect specified) that was not resolved in this research track due to time constraints.

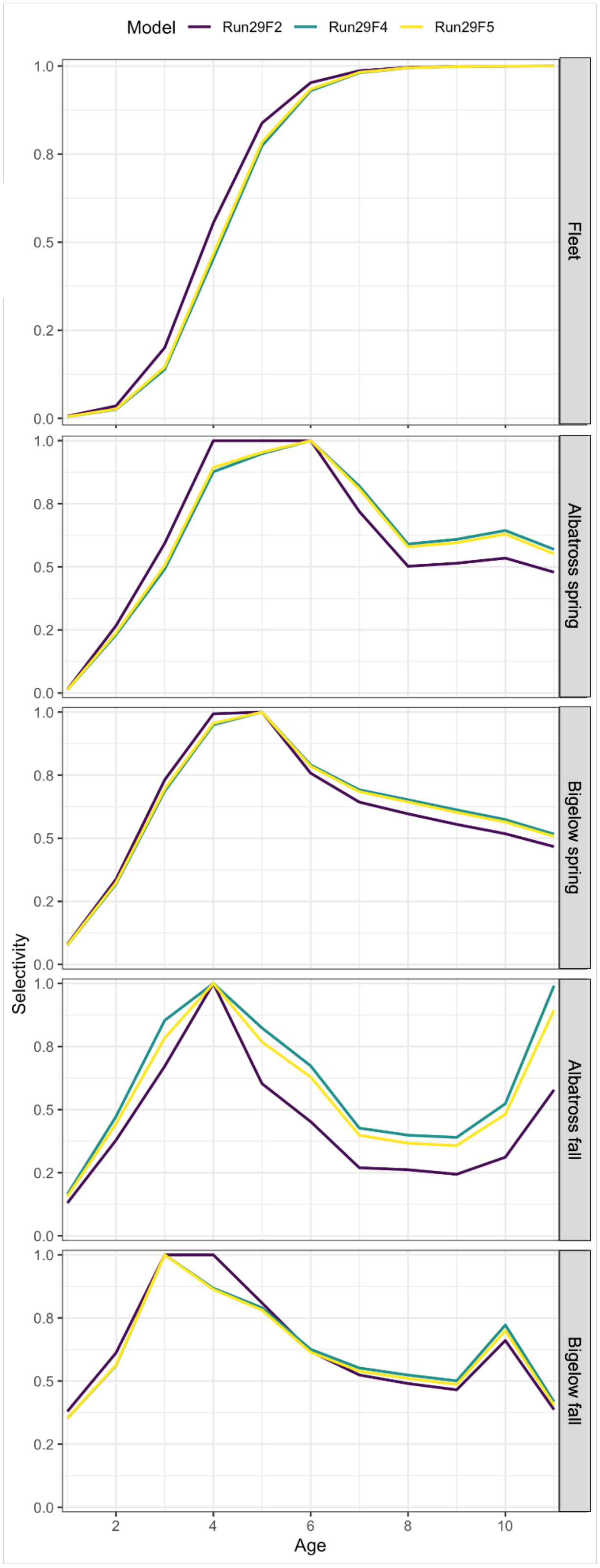


Figure 4.8: Selectivity estimates for candidate runs 29F2 (purple), 29F4 (green) and 29F5 (yellow) for the fleet (logistic selectivity) and four indices (selectivity-at-age).

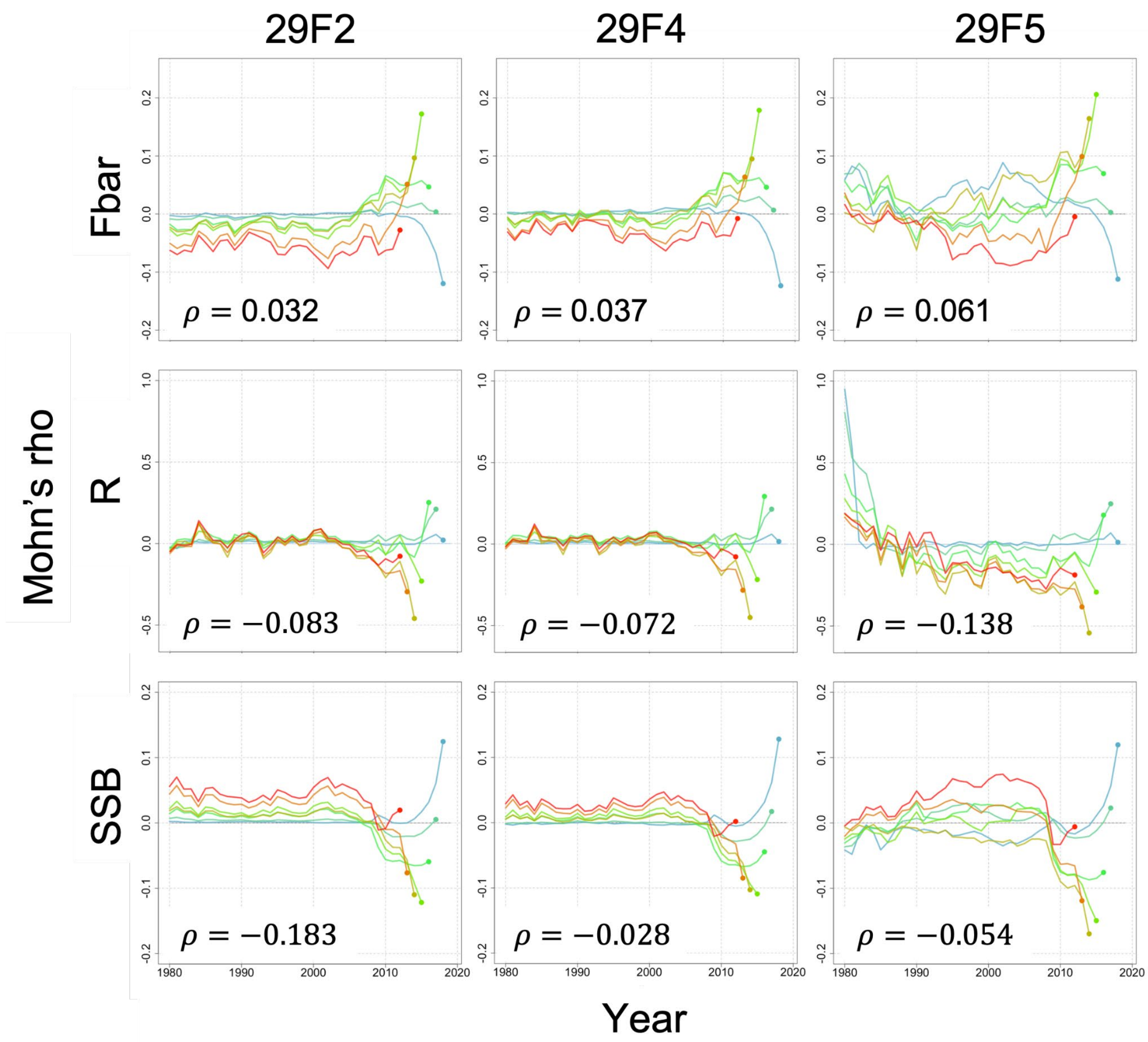


Figure 4.9: Seven-year retrospective peels on relative scale for F, R, and SSB for each of the candidate model runs (29F2, 29F4, 29F5).

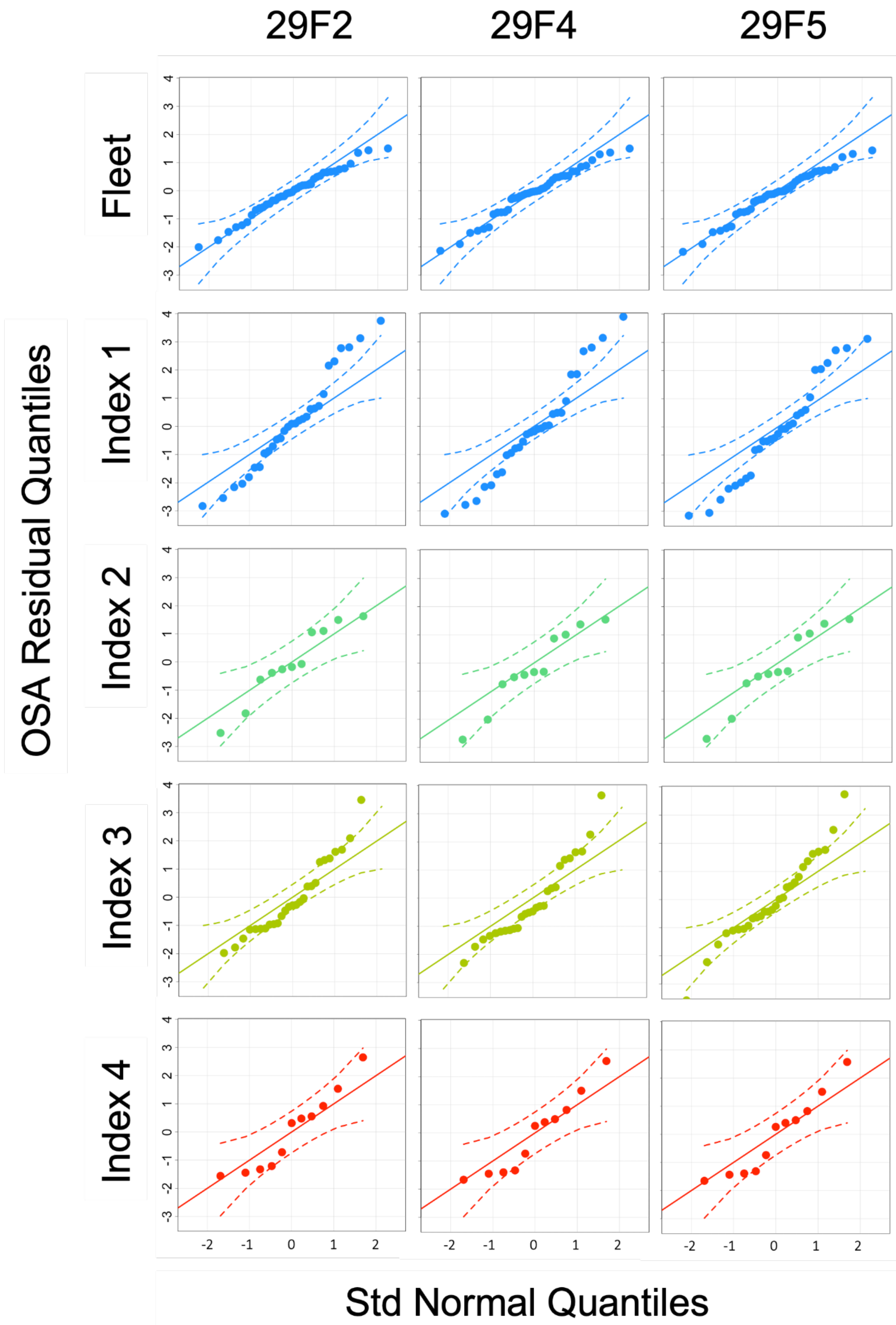


Figure 4.10: QQ plots reflecting the normalcy of OSA residual distributions for fit to aggregate fleet and index data.

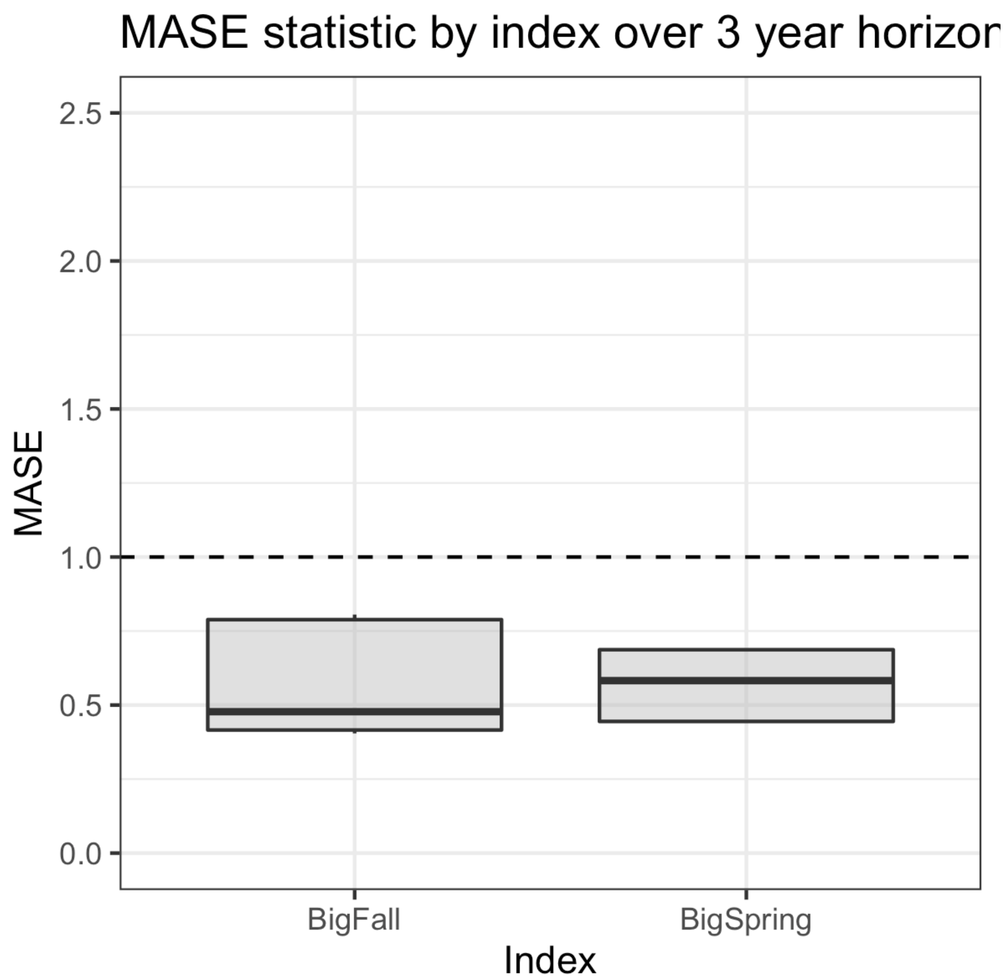


Figure 4.11: Mean absolute scaled error (MASE) calculated over a three year horizon to describe the accuracy with which spring and fall Bigelow indices are predicted by candidate runs. MASE scores less than 1 indicate indices are predicted with more accuracy than a naive approach, with smaller scores reflecting increased accuracy.

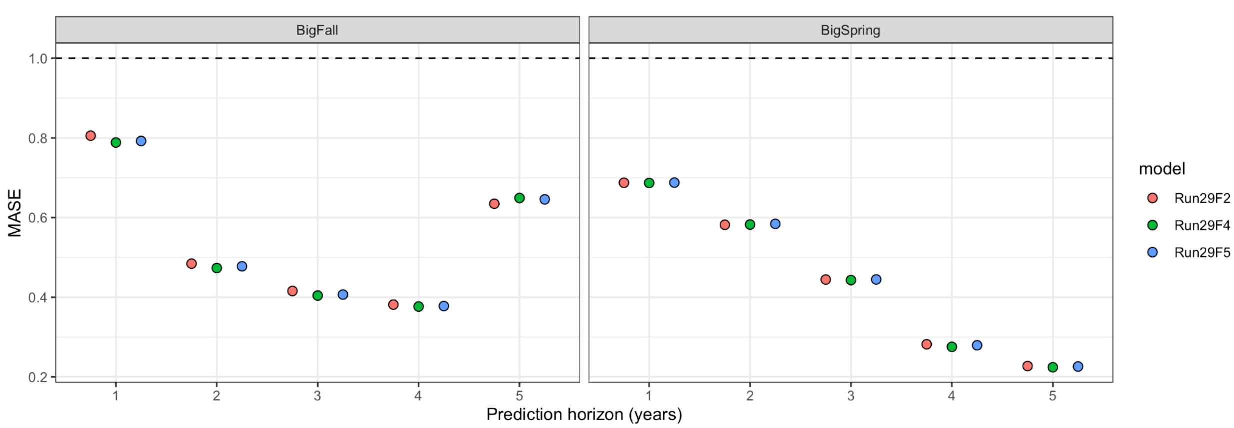


Figure 4.12: Mean absolute scaled error (MASE) for both the spring and fall Bigelow indices over prediction horizons of 1-5 years. MASE scores less than 1 indicate indices are predicted with more accuracy than a naive approach, with smaller scores reflecting increased accuracy.

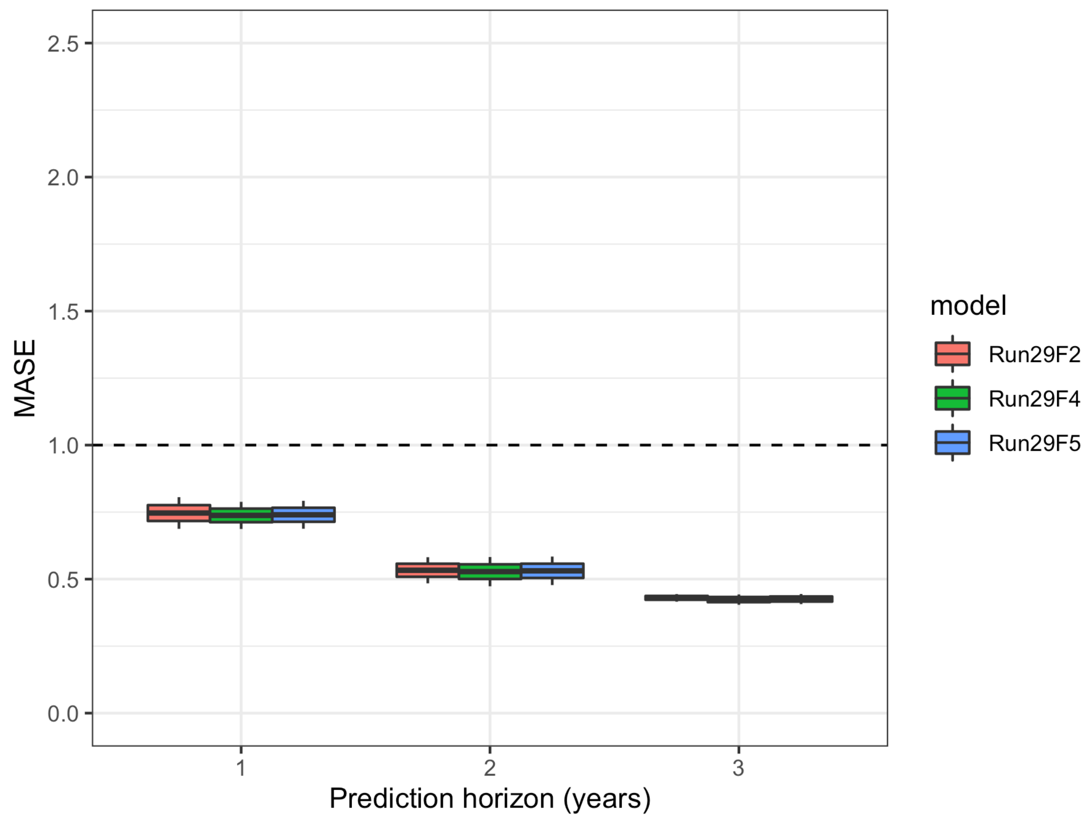


Figure 4.13: Mean absolute scaled error (MASE) calculated over 1-3 year prediction horizons for each candidate model by averaging MASE scores for spring and fall Bigelow indices for each model and horizon. MASE scores less than 1 indicate indices are predicted with more accuracy than a naive approach, with smaller scores reflecting increased accuracy.

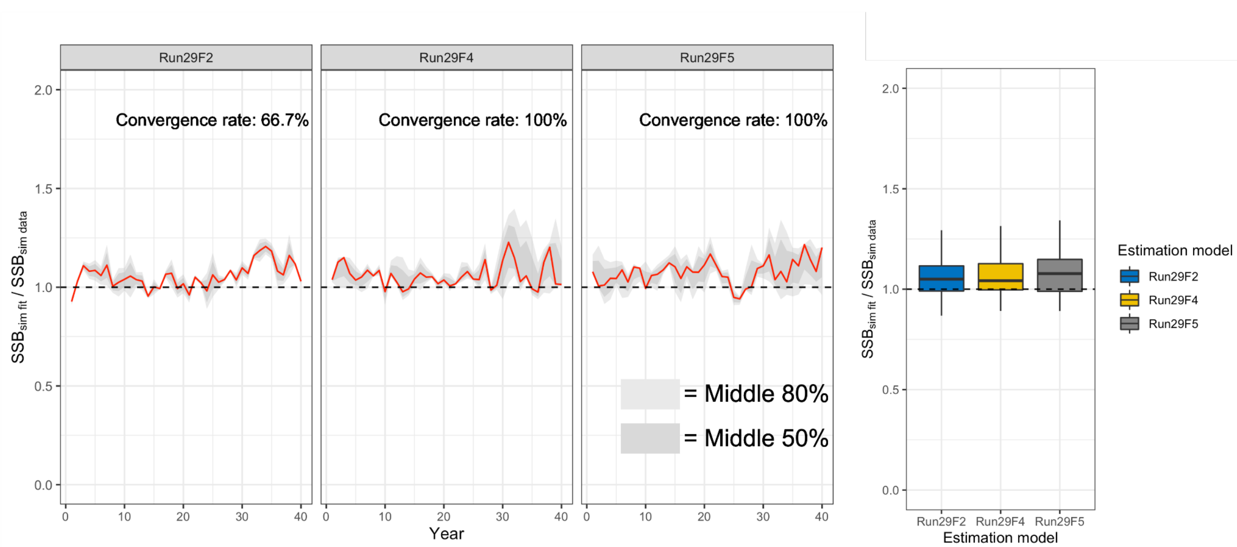


Figure 4.14: Spawning stock biomass (SSB) relative error for candidate model self-tests over time (3 leftmost panels, median in red, middle 50% and 80% of simulations in dark gray and gray respectively). Boxplot of relative error aggregated across all simulations and years for each of the candidate models (rightmost panel).

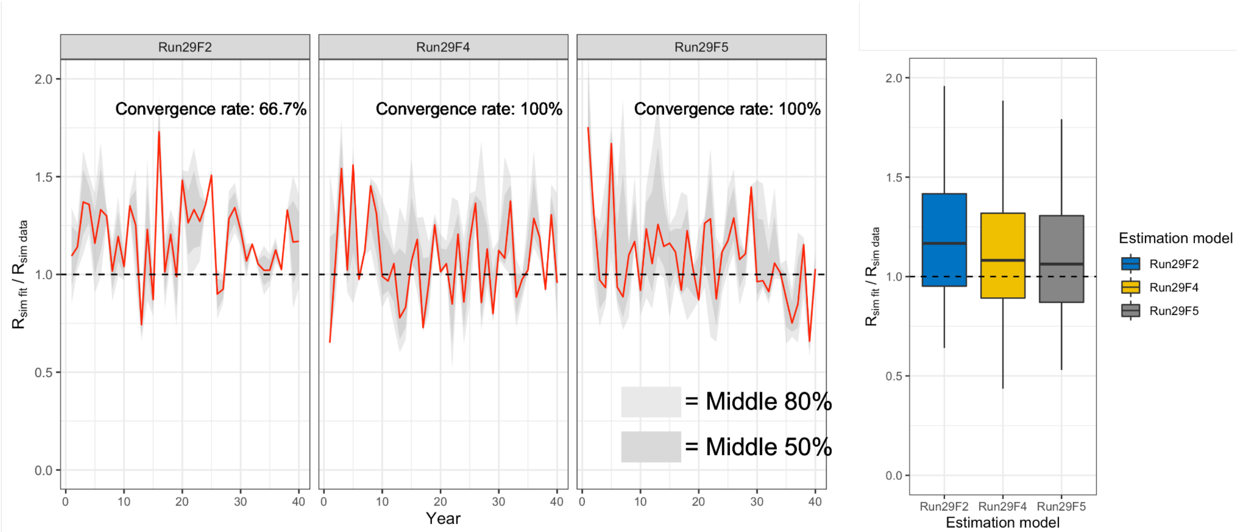


Figure 4.15: Recruitment (R) relative error for candidate model self-tests over time (3 leftmost panels, median in red, middle 50% and 80% of simulations in dark gray and gray respectively). Boxplot of relative error aggregated across all simulations and years for each of the candidate models (rightmost panel).

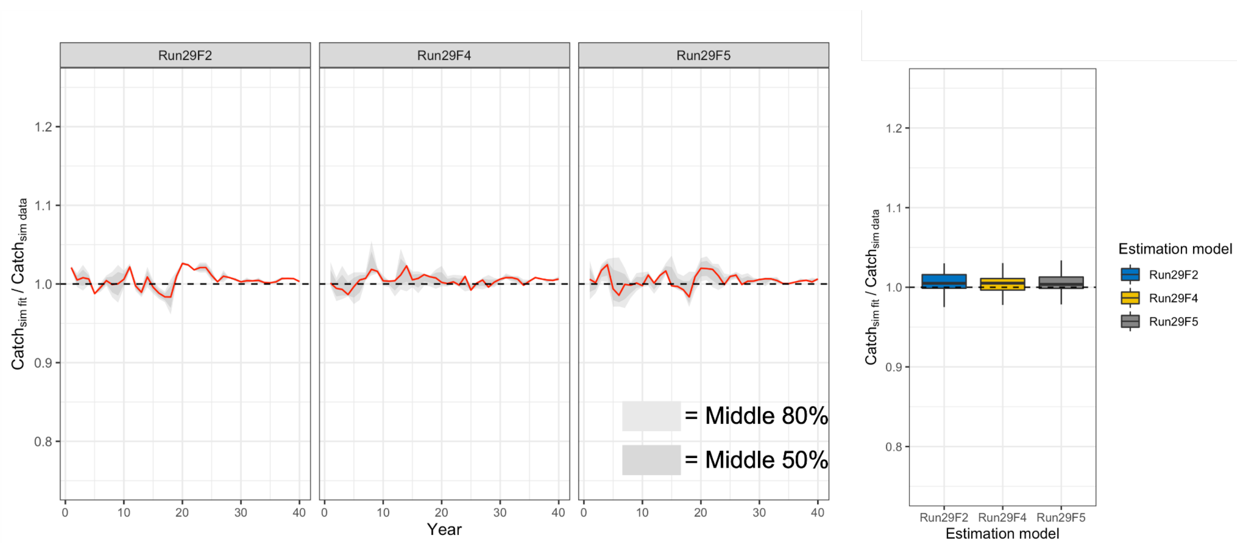


Figure 4.16: Catch relative error for candidate model self-tests over time (3 leftmost panels, median in red, middle 50% and 80% of simulations in dark gray and gray respectively). Boxplot of relative error aggregated across all simulations and years for each of the candidate models (rightmost panel).

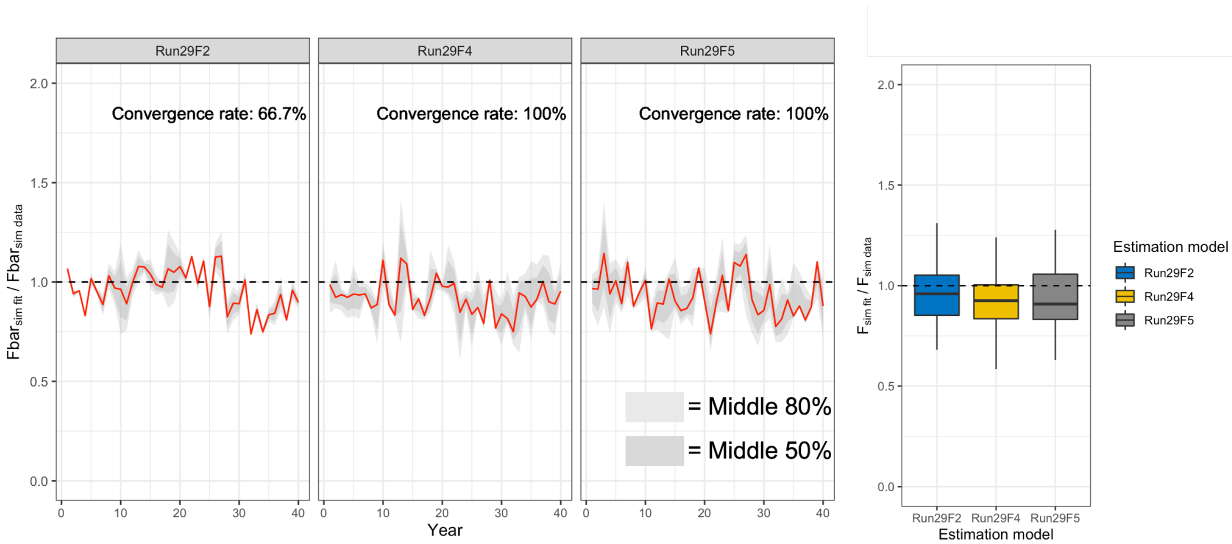


Figure 4.17: Fishing mortality (Fbar) relative error for candidate model self-tests over time (3 leftmost panels, median in red, middle 50% and 80% of simulations in dark gray and gray respectively). Boxplot of relative error aggregated across all simulations and years for each of the candidate models for each of the candidate models (rightmost panel).

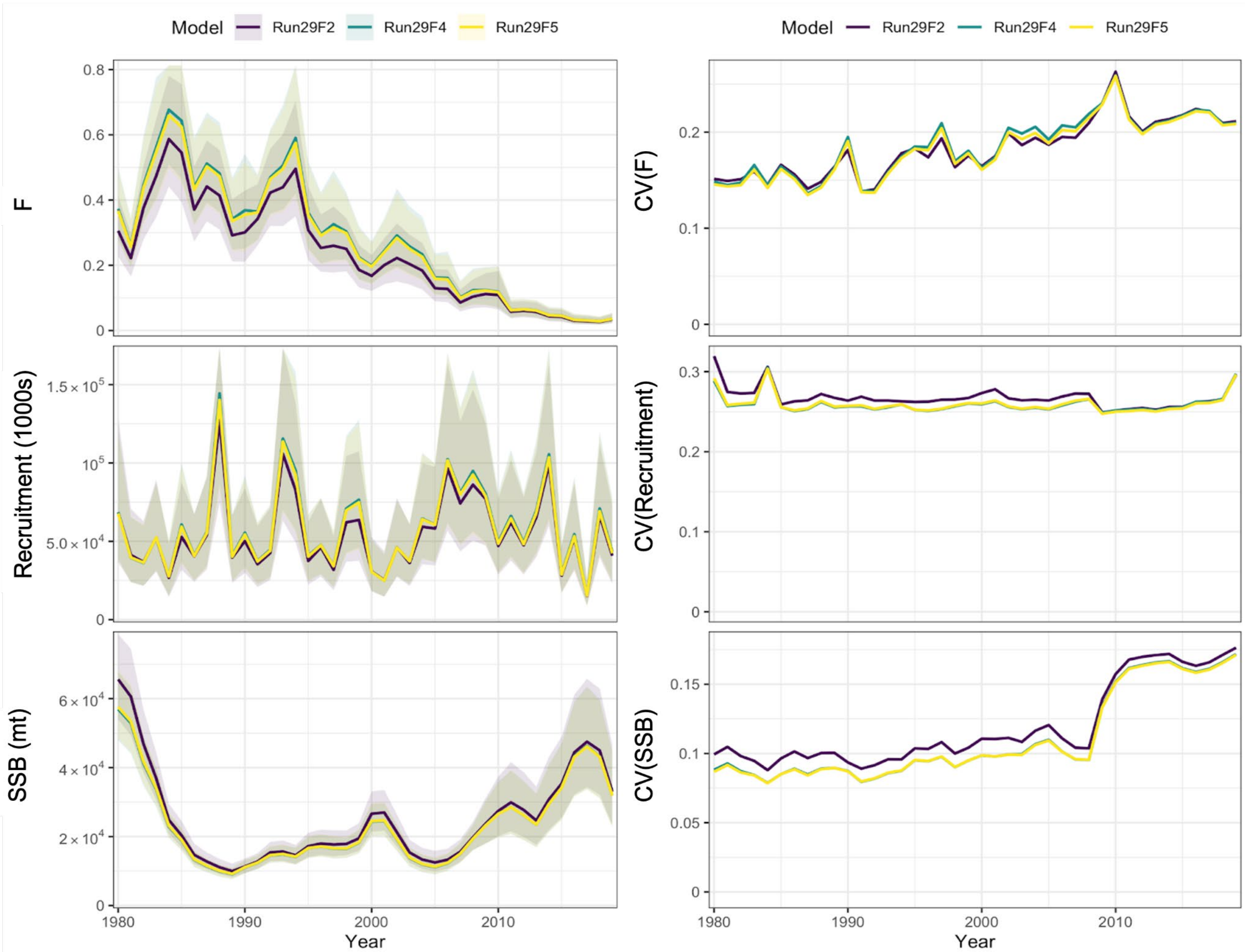


Figure 4.18: Estimates of fishing mortality (F), recruitment, and spawning stock biomass (SSB, left) and CVs around these estimates (right) for candidate runs 29F2, 29F4, and 29F5.

Table 4.3: Description, AIC, and Mohn's rho values for all candidate WHAM runs fo American plaice.

Description	Run	AIC	Rho_R	Rho_SSB	Rho_Fbar
Split Albatross/Bigelow time series with logistic normal age comp, starting values from multinomial age comp, fix Albatross fall age 11+ selectivity at 29B estimate & apply selectivity random effect to Albatross fall index	29F-2	-5480.6	-0.0825	-0.0311	0.0319
Split Albatross/Bigelow time series with logistic normal age comp, Bigelow units & biomass units and fix only 1 age at full selectivity for each index based on initial run that used 0.5 as starting estimate	29F-4	-5482.3	-0.0724	-0.0278	0.0370
Implementation like run 29F-4, but also implement an ar1 random effect on age for the Albatross fall index.	29F-5	-5453.0	-0.1378	-0.0540	0.0607

Table 4.4: American plaice spawning stock biomass (SSB, mt), recruitment (thousands), and fully selected F estimates for WHAM Run 29F4.

Year	SSB	Recruitment	Fully selected F
1980	56942	68642	0.373
1981	52932	39477	0.259
1982	41213	36118	0.446
1983	33744	52479	0.561
1984	22868	27877	0.677
1985	18892	60690	0.643
1986	13271	40971	0.439
1987	11366	55961	0.512
1988	9928	144385	0.482
1989	9104	40636	0.341
1990	11056	55489	0.369
1991	12362	37453	0.365
1992	14546	44661	0.469
1993	14870	115585	0.506
1994	14088	95152	0.590
1995	16633	40840	0.361
1996	17137	47659	0.295
1997	16522	34346	0.326
1998	16446	70895	0.304
1999	18253	76562	0.224
2000	24353	30585	0.199
2001	24563	24907	0.244
2002	19186	46047	0.291
2003	13764	37286	0.258
2004	11937	64530	0.234
2005	11137	60620	0.162
2006	12297	102378	0.161
2007	14969	80895	0.102
2008	19355	94889	0.123
2009	23291	80235	0.124
2010	26649	49331	0.119
2011	28515	66187	0.063
2012	26168	49169	0.066
2013	23433	69620	0.062
2014	29500	105542	0.047
2015	34269	29335	0.046
2016	43291	54561	0.033
2017	46289	15706	0.031
2018	43149	70931	0.028
2019	31858	43350	0.036

Age Structured Assessment Program

The Age Structured Assessment Program (ASAP) was applied to the information available for American plaice in US waters. ASAP is a statistical catch-at-age model that has been widely applied to a range of US and international fisheries (Legault and Restrepo 1998, Dichmont et al. 2016). The ASAP applications to plaice were developed as a potential candidate assessment method, but as the advantages of WHAM emerged (e.g., better fit to fishery age composition with process errors in selectivity), ASAP modeling continued for exploration and supporting information.

The ASAP model building process followed the same general workflow as WHAM. An initial 'bridge' run was configured with data from the 2019 VPA (NEFSC 2022b), natural mortality and maturity assumptions were revised, then alternative models were explored (e.g., alternative stock indices, alternative selectivity models, earlier starting year, likelihood weighting, etc.). Effective sample size for age composition was iteratively derived (Francis 2011), design-based precision was assumed for survey indices (CVs reported in ToR3 section), and precision of fishery catch (CV=0.1) was iterated based on residual variance of an initial model.

Bridge runs were developed to illustrate the effects of transitioning from VPA to ASAP, new data, and new model assumptions (Figure 4.19). They generally fit the data despite no fine tuning. The major source of uncertainty in the 2019 VPA was a retrospective pattern (SSB rho= 0.27, F rho= -0.20), and the ASAP fit to the 2019 data with independent indices of abundance at age (run 1) had nearly the same retrospective pattern (SSB rho= 0.27, F rho= -0.20). ASAP run 2, with revised and updated data, had better retrospective consistency for SSB (rho=0.10) but worse consistency for F (rho= -0.25). ASAP run 3, with revised assumptions of M and maturity, had better retrospective consistency in both (SSB rho= 0.05, F rho= -0.09), and all candidate ASAP runs had minor retrospective patterns (absolute rho <0.1 for SSB and F). Therefore, it appears that the previous assumption of natural mortality contributed to retrospective patterns in previous plaice assessments.

Among the alternative selectivity models and likelihood weightings, results were most sensitive the transition from independent indices of abundance at age to aggregate biomass indices and age composition (Figure 4.20). Age-aggregated indices with age composition are more consistent with survey sampling designs and their precision in an integrated statistical model surveys. Candidate ASAP runs fit the catch, aggregate biomass indices well, but had residual patterns to fishery and survey age composition.

Initial runs that fit to calibrated NEFSC surveys as single spring and fall indices had residual patterns in which survey observations of age-1 composition from the Bigelow survey (calibrated to Albatross units) were all greater than model predictions (Figure 4.21). This residual pattern appears to result from the uncertain estimates of relative efficiency for the Albatross and Bigelow for small plaice (Figure 3.6). Therefore, ASAP runs were revised to fit the NEFSC surveys as separate Albatross and Bigelow indices, as recommended in the previous stock assessment (NEFSC 2022b).

Based on model diagnostics, the best ASAP run fit to 1980-2019 fishery and NEFSC Albatross and Bigelow surveys (run 43; Figures 4.22-4.26). The model fit catch, survey biomass indices and survey age composition well, but patterns in age composition residual persisted despite several explorations of alternative selectivity models, suggesting stochastic variability in annual selectivity (Figure 4.23). Estimates of stock size and fishing mortality were retrospectively consistent (SSB rho= 0.04, F rho= -0.03; Figure 4.26) and well estimated (CV<0.2), except for estimates of recruitment in the most recent five years which were relatively uncertain (CV=0.2 to 0.5). Considering the residual patterns in fishery age composition, estimation of process errors in selectivity by WHAM were preferred by the Working Group.

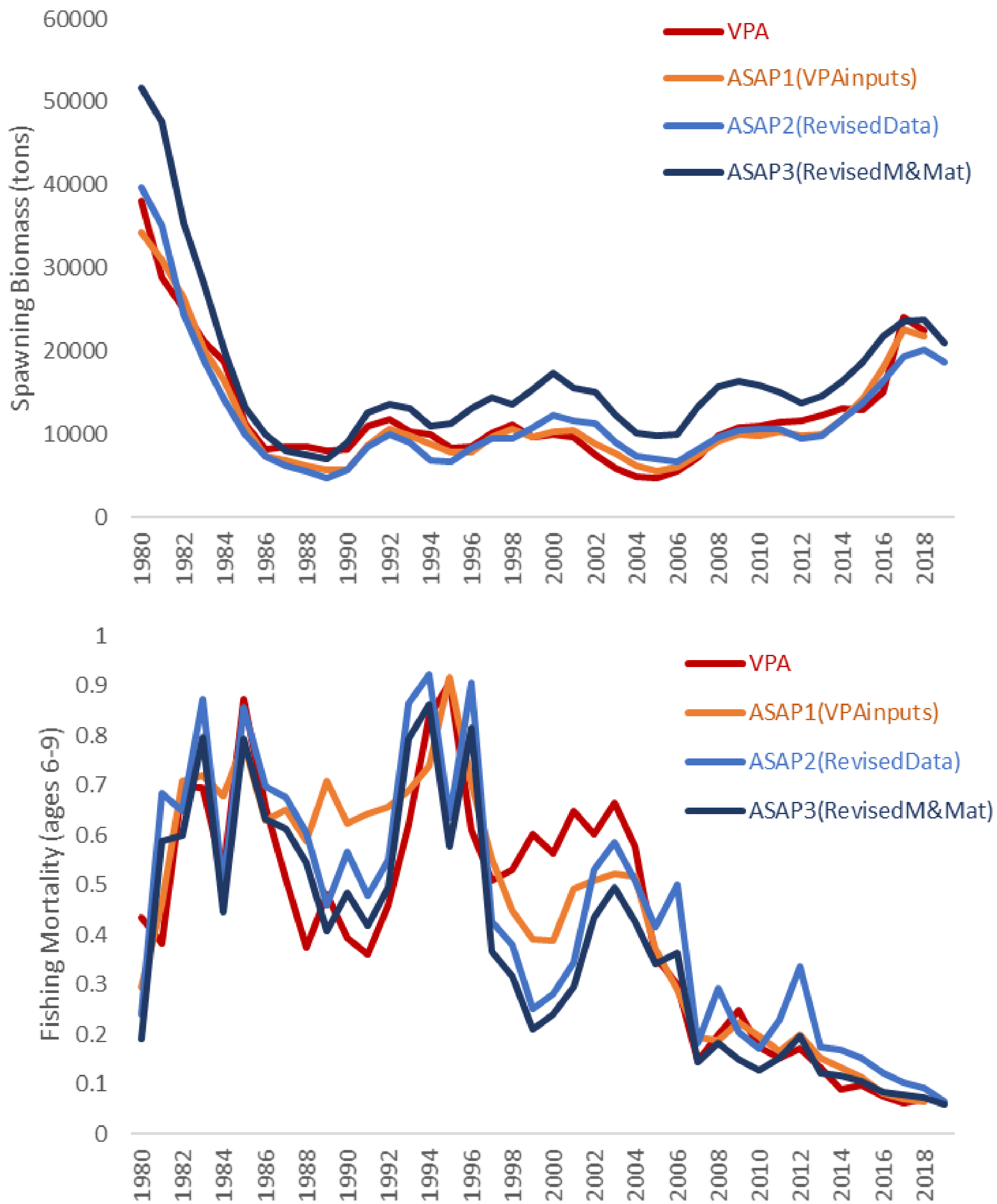


Figure 4.19: Estimates of American plaice spawning stock biomass and fishing mortality from VPA and ASAP bridge runs.

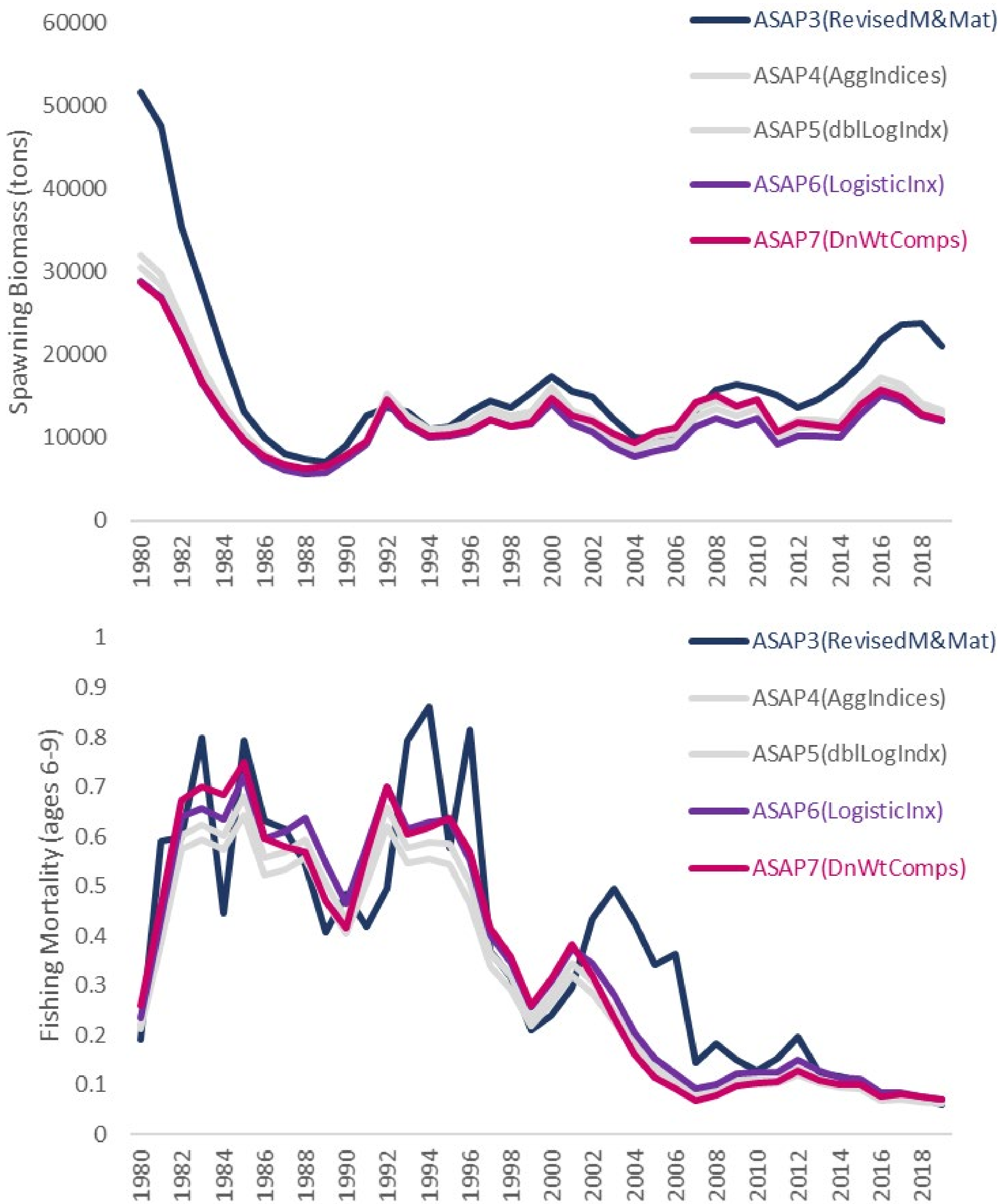
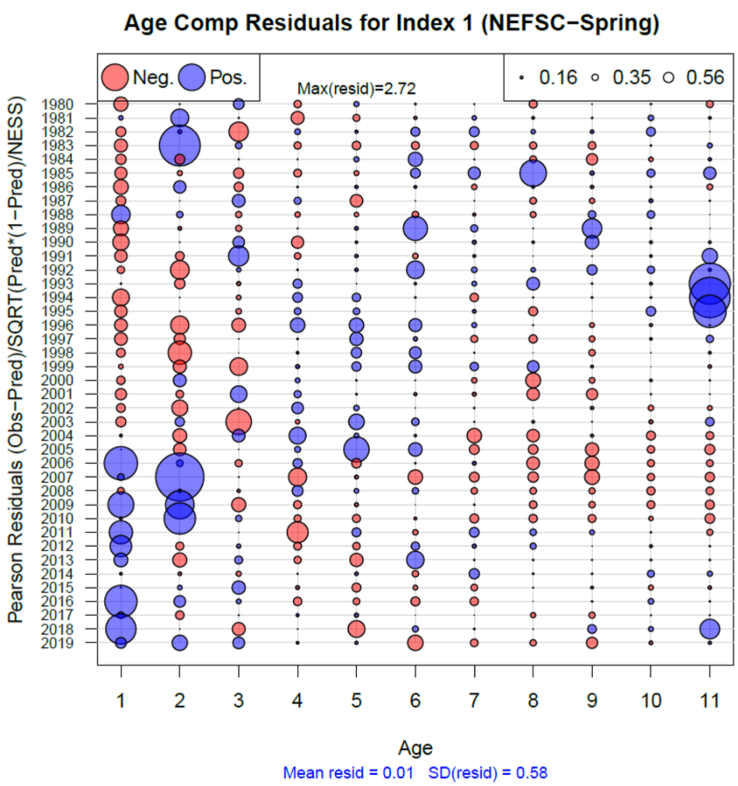


Figure 4.20: Estimates of American plaice spawning stock biomass and fishing mortality from alternative ASAP runs.



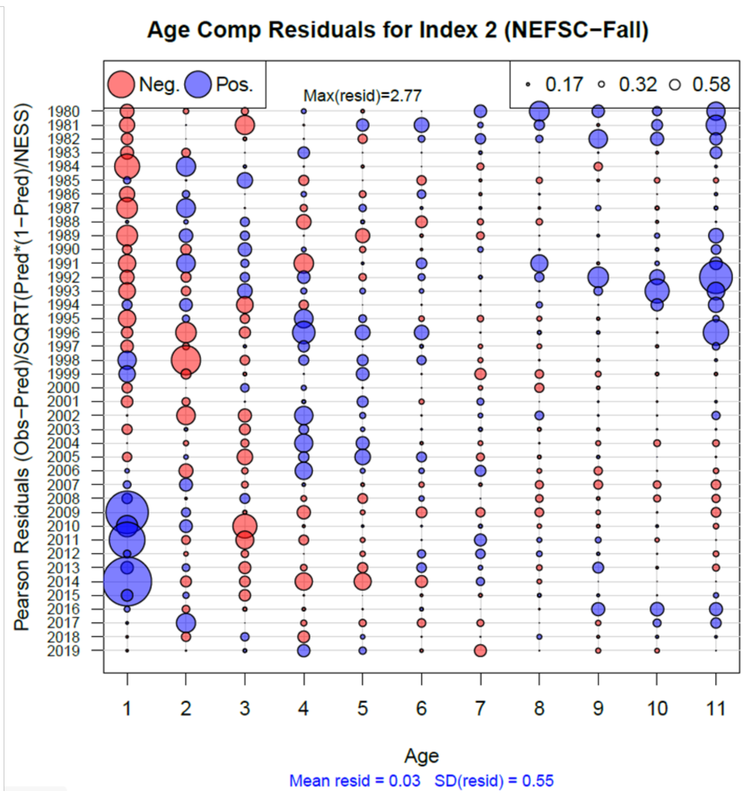


Figure 4.21: Age composition residuals from American plaice ASAP run 27 with calibrated NEFSC surveys fit as single spring and fall indices.

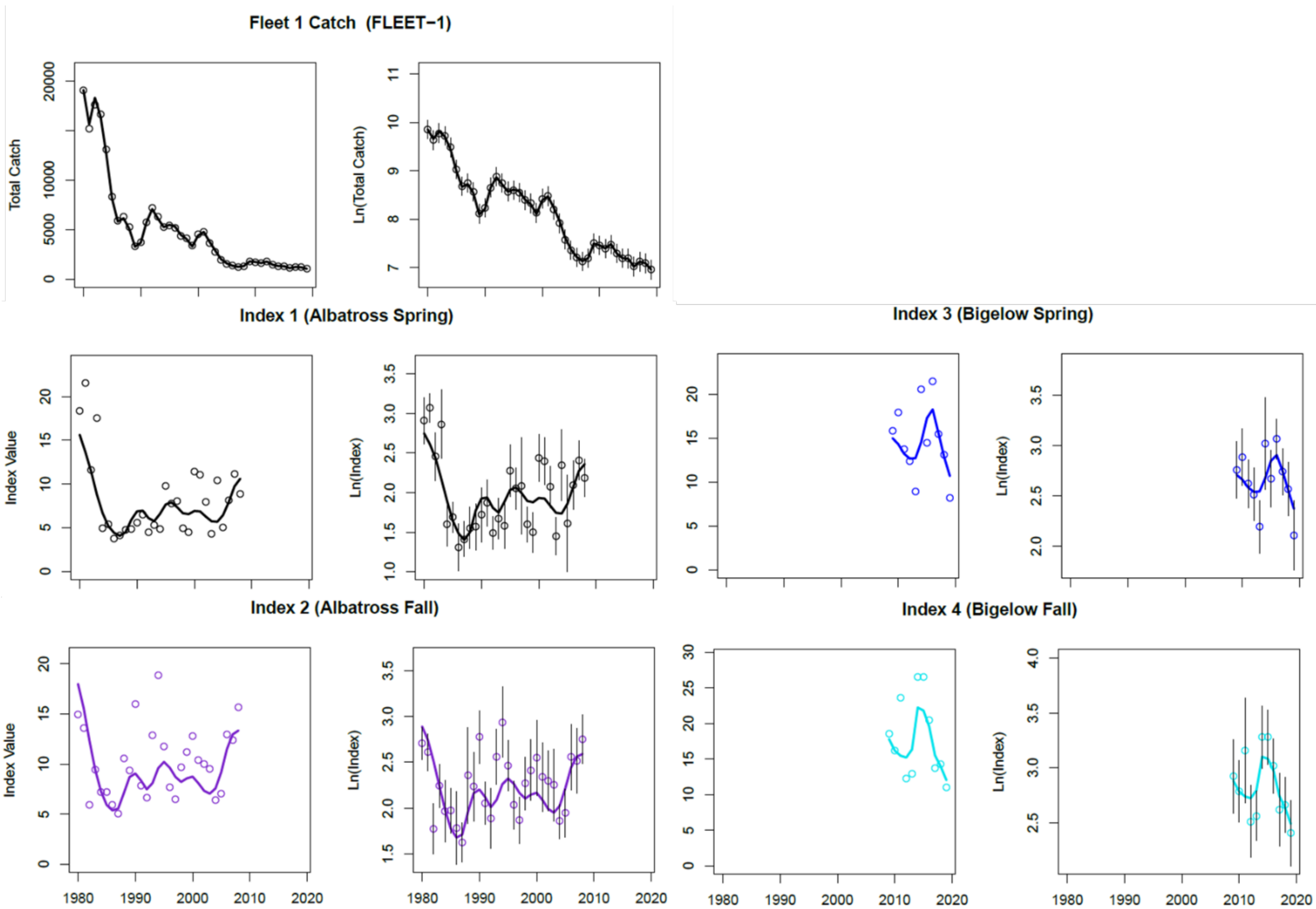


Figure 4.22: American plaice ASAP run 43 fit to aggregate fishery and survey catches.

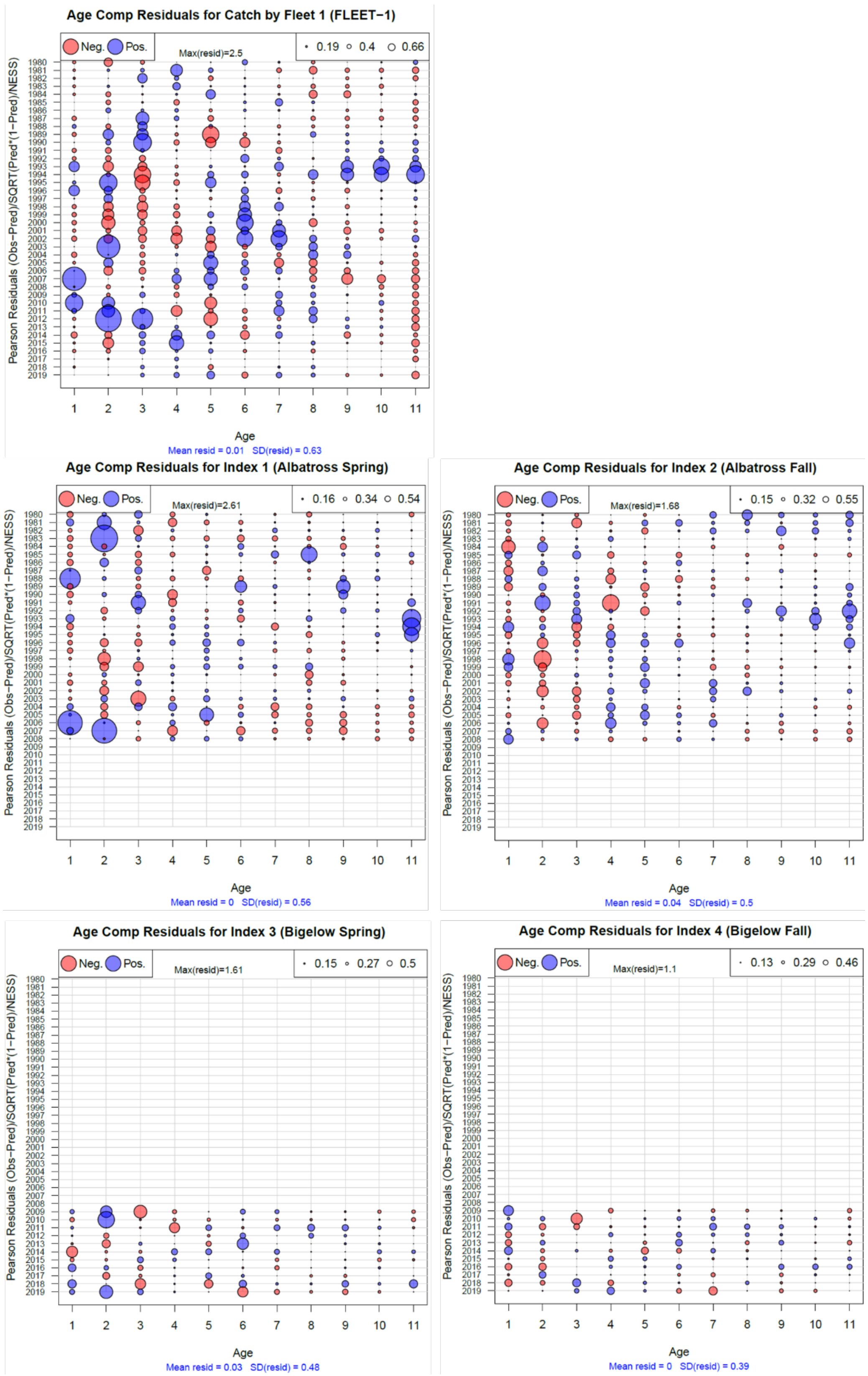


Figure 4.23: American plaice ASAP run 43 fit to fishery and survey age composition.

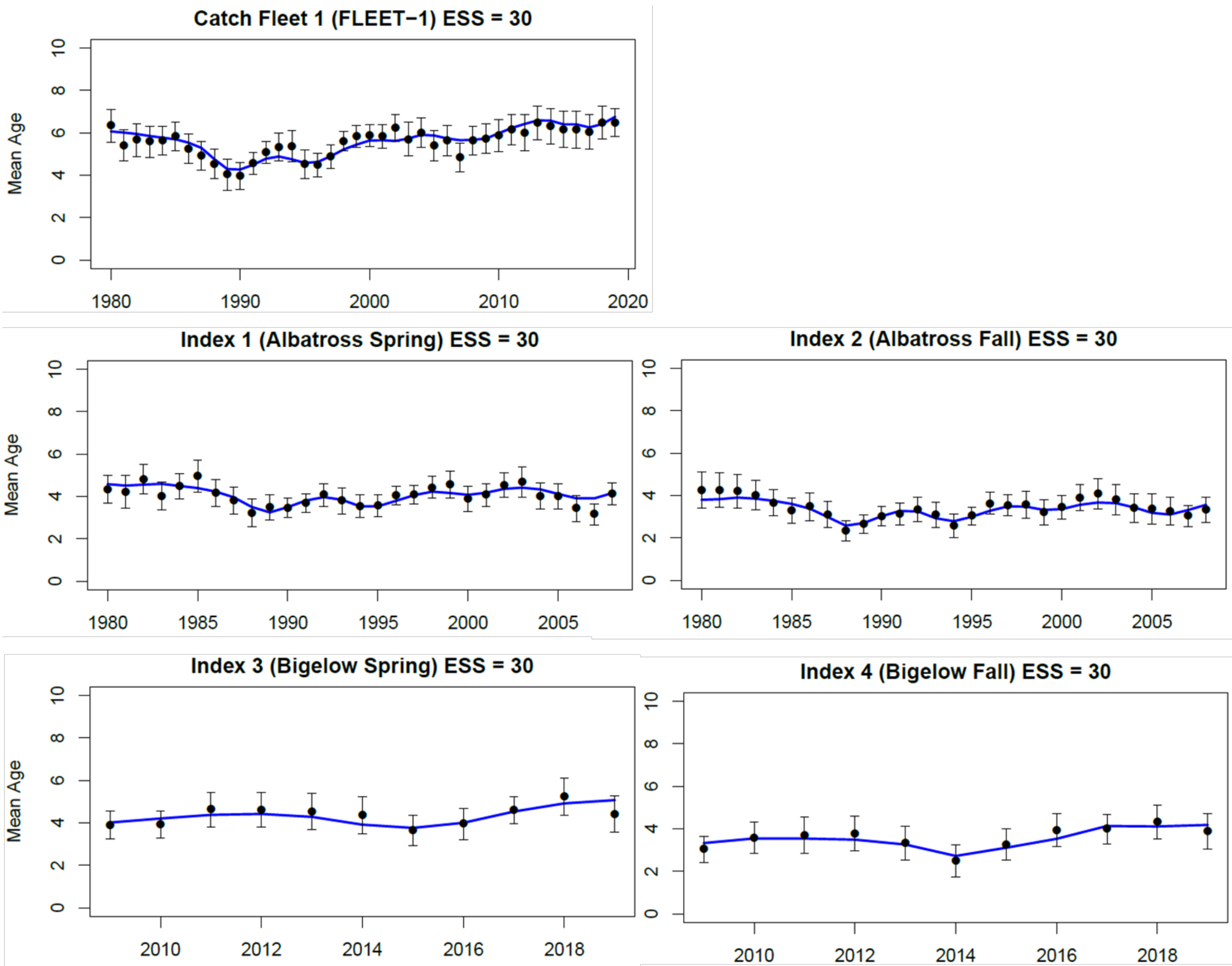


Figure 4.23: Observed mean age (black dot and error bars) of American plaice in the fishery and surveys and predicted values from ASAP run 43 (blue lines) for iterative determination of effective sample size (ESS).

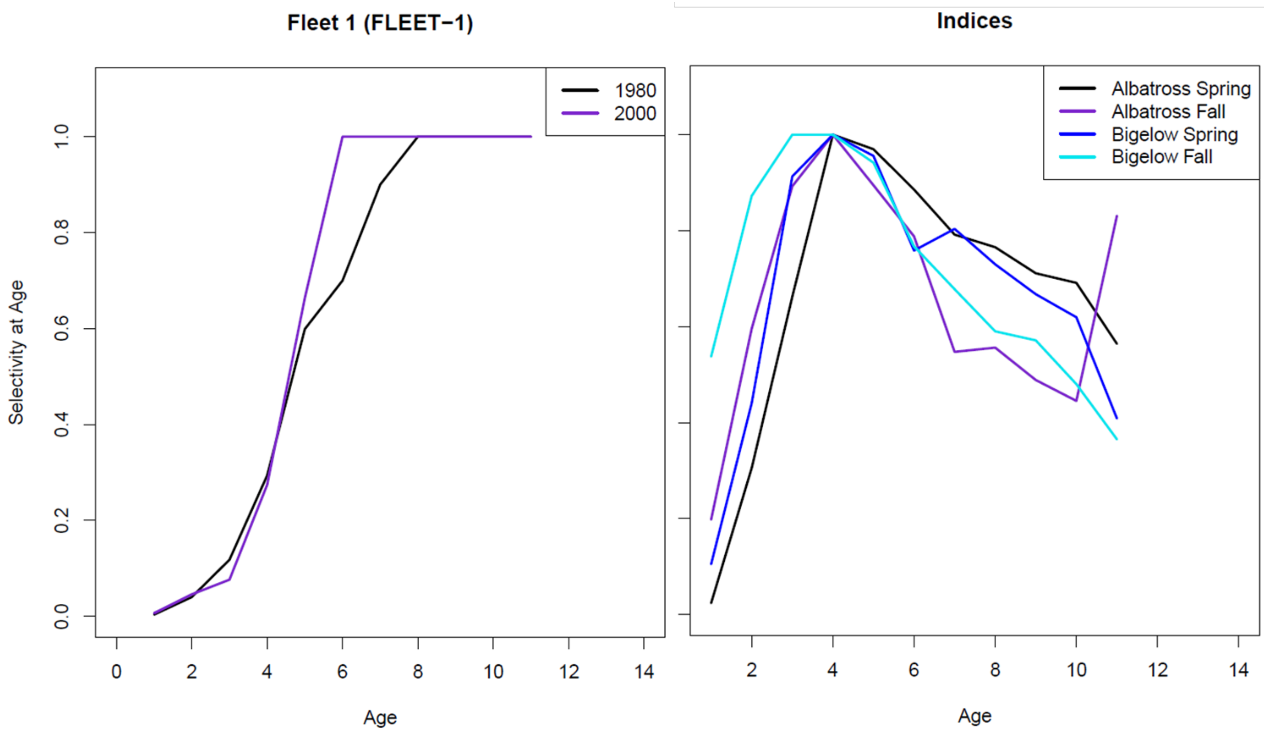


Figure 4.24: Estimates of American plaice selectivity by the fishery (left) and surveys (right) from ASAP run 43.

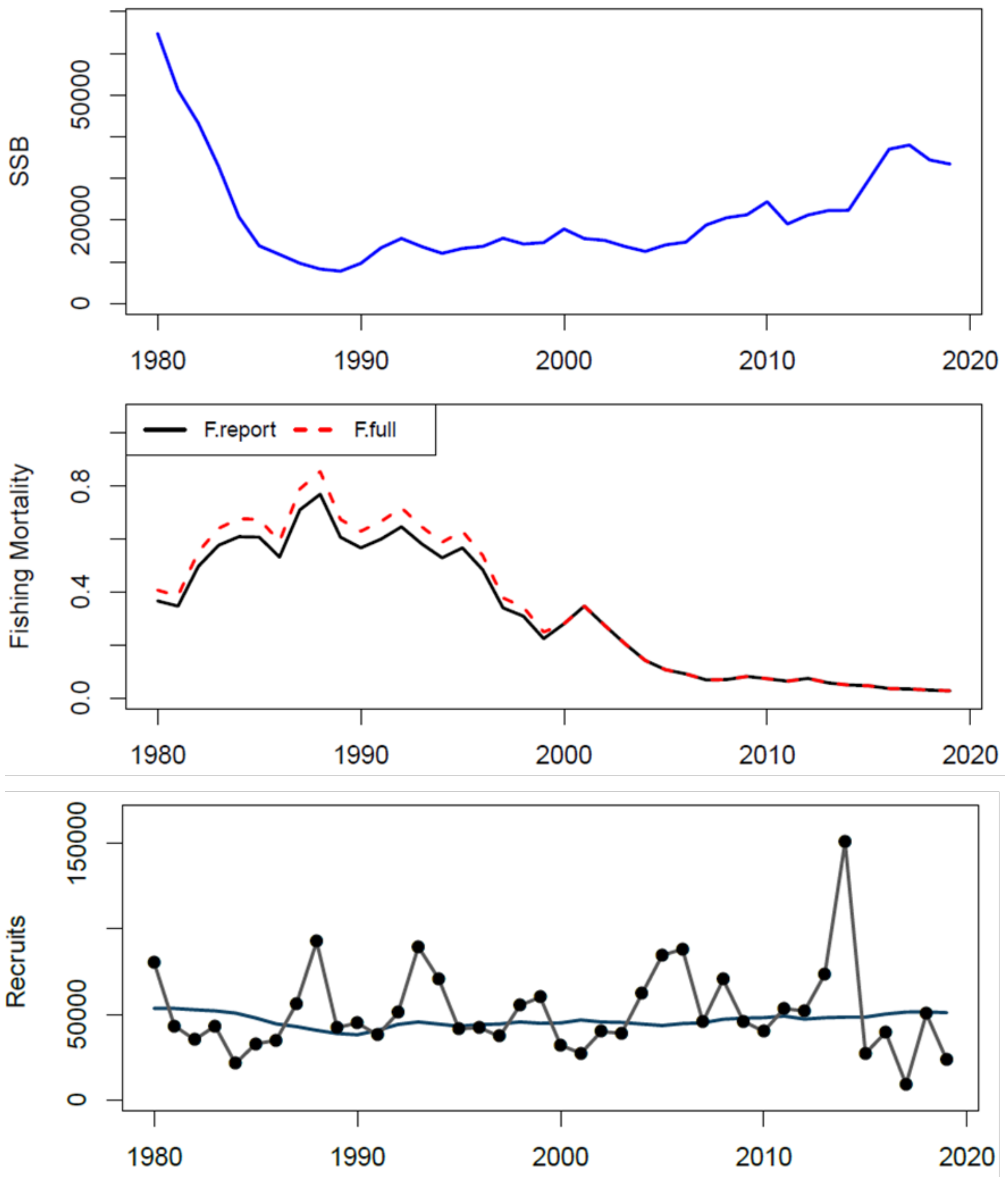


Figure 4.24: Estimates of American plaice spawning stock biomass (SSB, mt) and fishing mortality (Freport: average of ages 6-9, and fully selected F) and age-1 recruitment (thousands) from ASAP run 43.

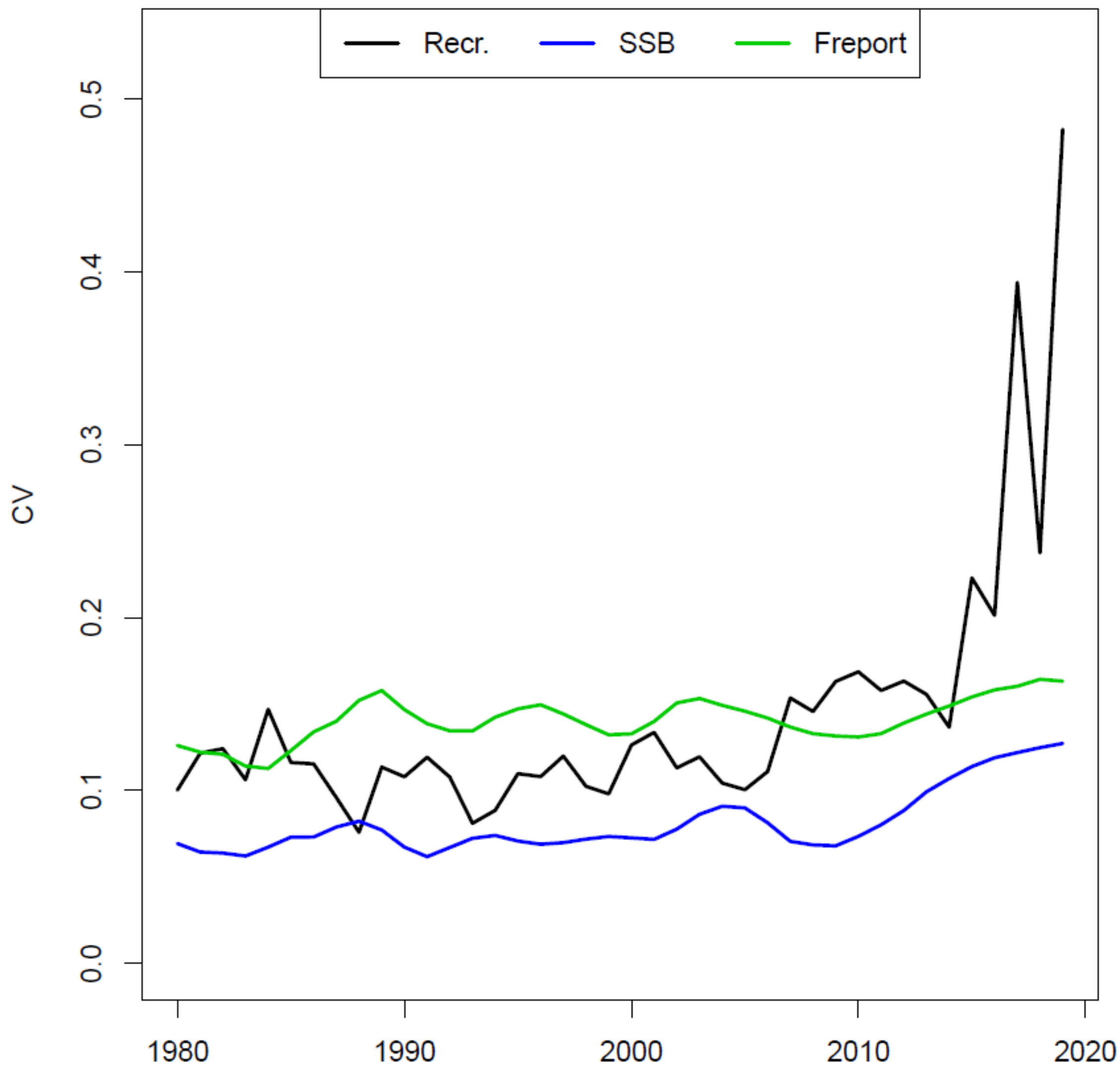


Figure 4.25: Estimated precision (CV: coefficient of variation) for American plaice spawning stock biomass (SSB) and fishing mortality (Freport: average of ages 6-9) and age-1 recruitment from ASAP run 43.

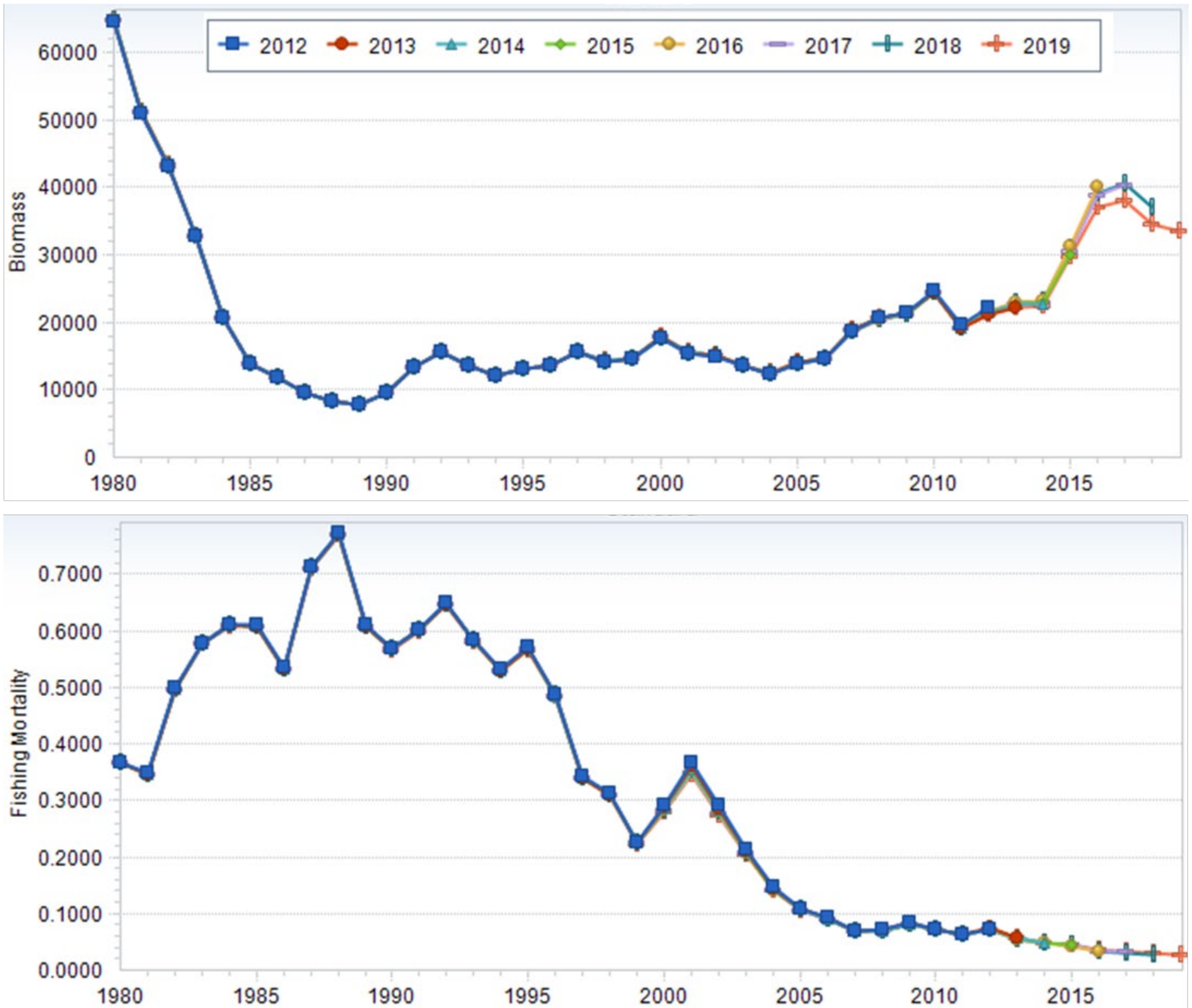


Figure 4.26: Retrospecitve analysis of American plaice spawning stock biomass (SSB) and fishing mortality (Freport: average of ages 6-9) from ASAP run 43.

Stock Synthesis

An American plaice assessment model was developed in Stock Synthesis (SS, Methot and Wetzel 2013), with the objective of providing support for the primary assessment model results (Hennen and Hansell, 2022, Working Paper 17, Appendix A). SS is an integrated statistical catch-at-age model that can fit to unprocessed data and account for important processes (e.g., mortality and growth) that operate in conjunction with catch, size/age and indices of abundance. SS is one of the most commonly used stock assessment packages in the US and globally (Dichmont 2016, 2021) and has many essential features of next generation stock assessment models (Punt et al. 2020). SS provides a few uncommon features that allow inference based on data that are not used in other assessment models. Therefore, if the results from SS using a different structural configuration align with results from the primary assessment model, results can be considered robust to model configuration and platform.

SS was chosen to take advantage of length data, which was not used in other stock assessment platforms. Using conditional age at length data means that there is no need to develop an external age at length key to generate age composition data. Conditional age at length data can be particularly useful in informing estimates of growth. Using length composition data directly allows for estimation of selectivity at length instead of estimating selectivity at age as is done in the other assessment models for American plaice. Selectivity at length has some interpretive advantages because many selective processes act on length rather than age (e.g., mesh size).

The data used in the SS model were generally the same as the primary assessment model, with the addition of length composition data and length conditioned age data (Figure 4.27). Because SS is fit to length composition data, a longer time series of composition data is available for the surveys (e.g., age data is only available since 1980). The SS model starts in 1940 to allow for a burn in period, where the model estimates recruitment.

Some parameters were fixed because they were difficult to estimate. Among these are natural mortality (M=0.3), the Brody growth coefficient (k=0.13) and length at minimum age (an SS specific parameter below which the growth curve becomes linear; fixed at 9.2 cm), maturity

(Table 4.1), weight at length ($W=2.28e\text{-}06L^{3.369}$as well as several others (see Hennen and Hansell, 2022, Working Paper 17, Appendix A for more details). Steepness of the stock-recruit relationship was estimated in some runs and assumed in others based on initial estimates (h=0.61), When possible, each of these were set to match survey-based estimates described in TOR 3 in this report.

Some stock-recruitment parameters were estimated, including the equilibrium recruitment level (R0) and the average recruitment deviation (σR). Survey catchability (Q) was estimated for each survey. Survey size selectivity was estimated with a two-parameter logistic curve. The catch retention curve required three estimated parameters. Initial F was estimated as well, which informs initial depletion of the stock. Fishery selectivity was estimated using three time blocks (1985 - 1988; 1989 - 1997; 1998 - 2019). The size at maximum age (similar to the asymptote in the Von Bertalanfy growth curve) was estimated and allowed to change in 1981.

SS allows for the use of likelihood penalties called “parameter priors” that restrict the estimation of parameters in a variety of potential ways. In this application to plaice, the “symmetric beta” option was chosen as a parameter prior, which introduces a weak penalty and probably did not influence estimation of parameters very much. The input variance for composition data was increased in order to allow the model some flexibility in fitting to the index data (e.g., Francis 2011). Bias adjustments were made to the estimated recruitment deviations according to the methods described in Methot and Taylor (2011).

The model met the criteria for convergence. The maximum gradient observed in the fit of the model was < 0.0002, none of the estimated parameters were strongly correlated or near bounds, and the Hessian matrix was positive definite. Jittering the starting values by 10% resulted in a convergence rate of approximately 60% and produced no solutions with a significantly likelihood less than the likelihood presented here. The low convergence rate indicates problems with model stability, but jitter results support the hypothesis that the solution is likely to be a global rather than local minimum over the likelihood surface. Likelihood profile analysis, in which R0 was fixed at values above and below the converged solution and refit, indicated no important conflicts between the various data. The only apparent conflict was between discards and the rest of the data, and although it may have contributed to model instability, was not generally sufficient to force the model away from the global minimum of the likelihood surface.

In general, the data were fit reasonably well by the model. The fit to the indices of abundance were noisy, but the residuals were randomly distributed (Figure 4.28; see Carvalho et al. 2021). The magnitude of the residuals was comparable to what would be expected given the input variance associated with each survey index. The fit to length composition data was reasonably good. The fit to annual mean length produced residuals with some patterning, but the magnitude of the residuals was close to what would be expected given the input variances associated with the data (Figure 4.29). The fit to the average length composition for all years was tight (Figure 4.30). Fits to individual years were generally good but showed some patterning in residuals (Figure 4.31). The fit to the conditional age at length data was also reasonably good. The residuals from the fit to annual mean age were somewhat patterned but were of a magnitude roughly appropriate to what was expected given the input variances associated with the data (Figure 4.32).

The model estimated a reduction in average maximum length at age (Figure 4.33). Independent analysis of survey data found similar results (Figure 2.17). The model also indicated reduced selectivity at smaller size between the years of 1989 and 1997, while the periods before and after this were broadly similar (Figure 4.34). The northern shrimp fishery discarded many plaice in the 1989-1997 period and the small-mesh in that fishery has different selectivity than the large-mesh trawl fleet that contributed most earlier and later discards. The selectivity of retained catch was estimated and when combined with the fishery selectivity can be used to infer the discard selectivity (selectivity*(retain+(1-retain)); Figure 4.35). Therefore, a constant retention curve combined with time varying fishery selectivity can potentially capture changes in discard selectivity. Time varying selectivity of the discard fleet is likely given the collapse of the northern shrimp fishery. The estimates of survey selectivity were constant through time and similar between the two surveys, and the surveys selected smaller fish than the fishery (Figure 4.36).

The trends and terminal biomass estimates from the SS model are similar to those estimated in other modeling frameworks. Spawning stock biomass appears to be at about one third of the estimated equilibrium level (Figure 4.37). Fishing mortality is currently low but appears to have been high in the past. The estimated recruitment time series is highly variable and uncertain at the beginning and end of the time series when there is little data to inform recruitment estimates. Retrospective analyses do not indicate substantial inconsistency of estimates (Figure 4.38). A revised SS run modeled NEFSC surveys as separate Albatross and Bigelow indices, as recommended in the previous stock assessment (NEFSC 2022b), and provided estimates similar to the candidate WHAM runs (Figure 4.39).

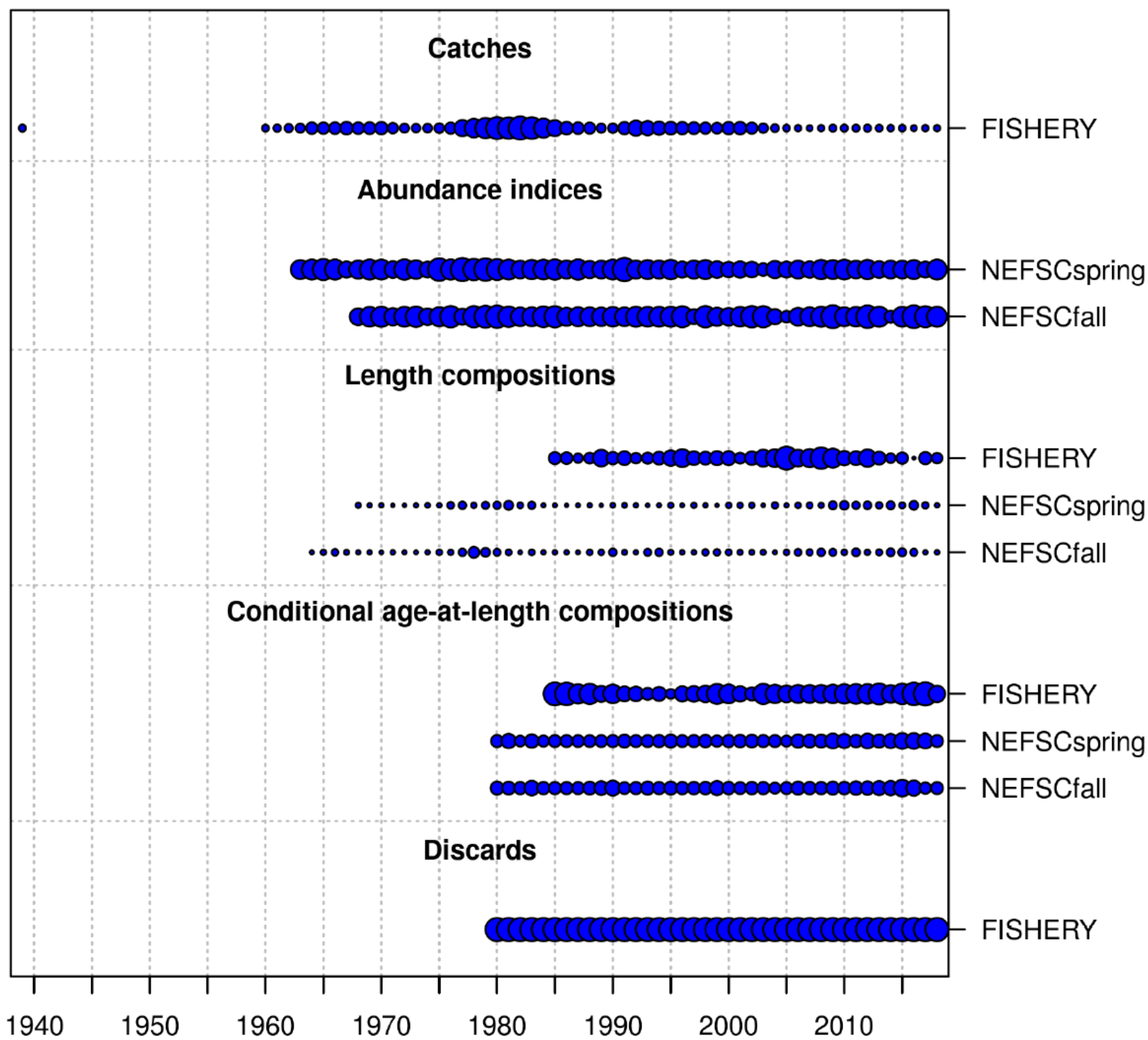


Figure 4.27. Data used for SS application to American plaice by year, where circle area is relative within a data type. Circles are proportional to total catch for catches; to precision for indices, discards, and mean body weight observations; and to total sample size for compositions and mean weight- or length-at-age observations. 'Ghost' observations (not included in the likelihood) have equal size for all years. Note that since the circles are scaled relative to maximum within each type, the scaling within separate plots should not be compared.

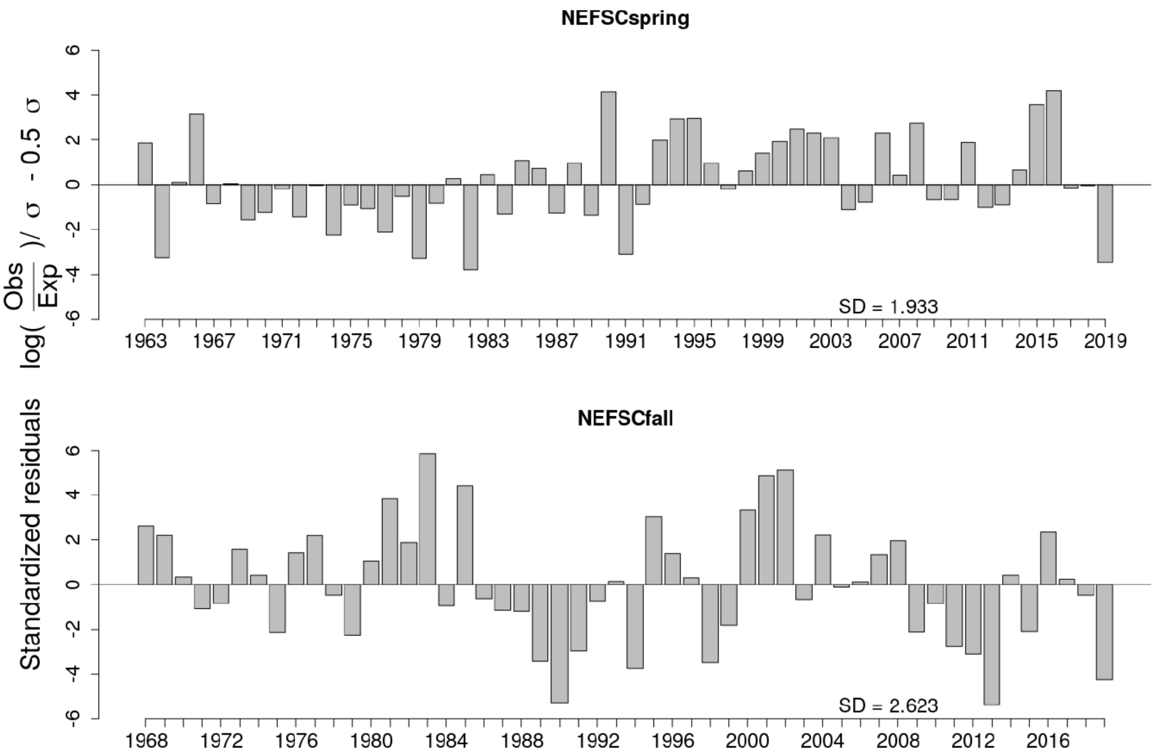


Figure 4.28. SS residuals from the model fits to each survey index used by year. The standard deviation of the residuals over the time series is shown over the horizontal axis.

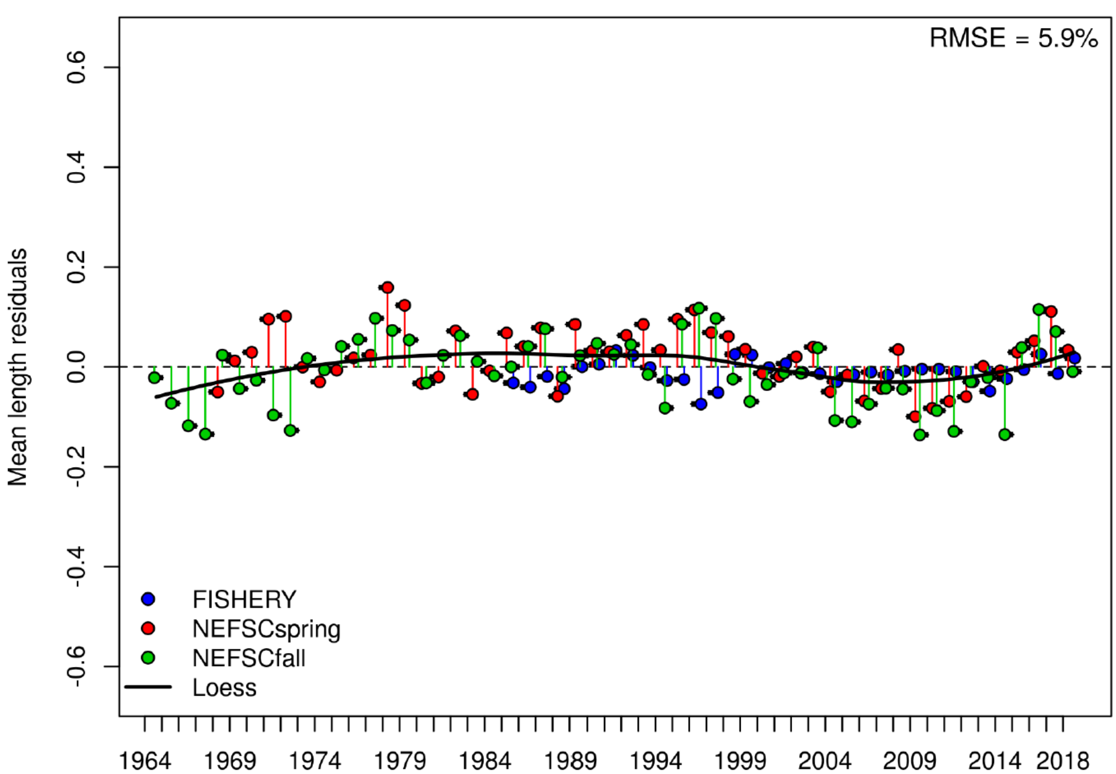


Figure 4.29. SS joint residual plot from fit to annual mean length from length composition data.

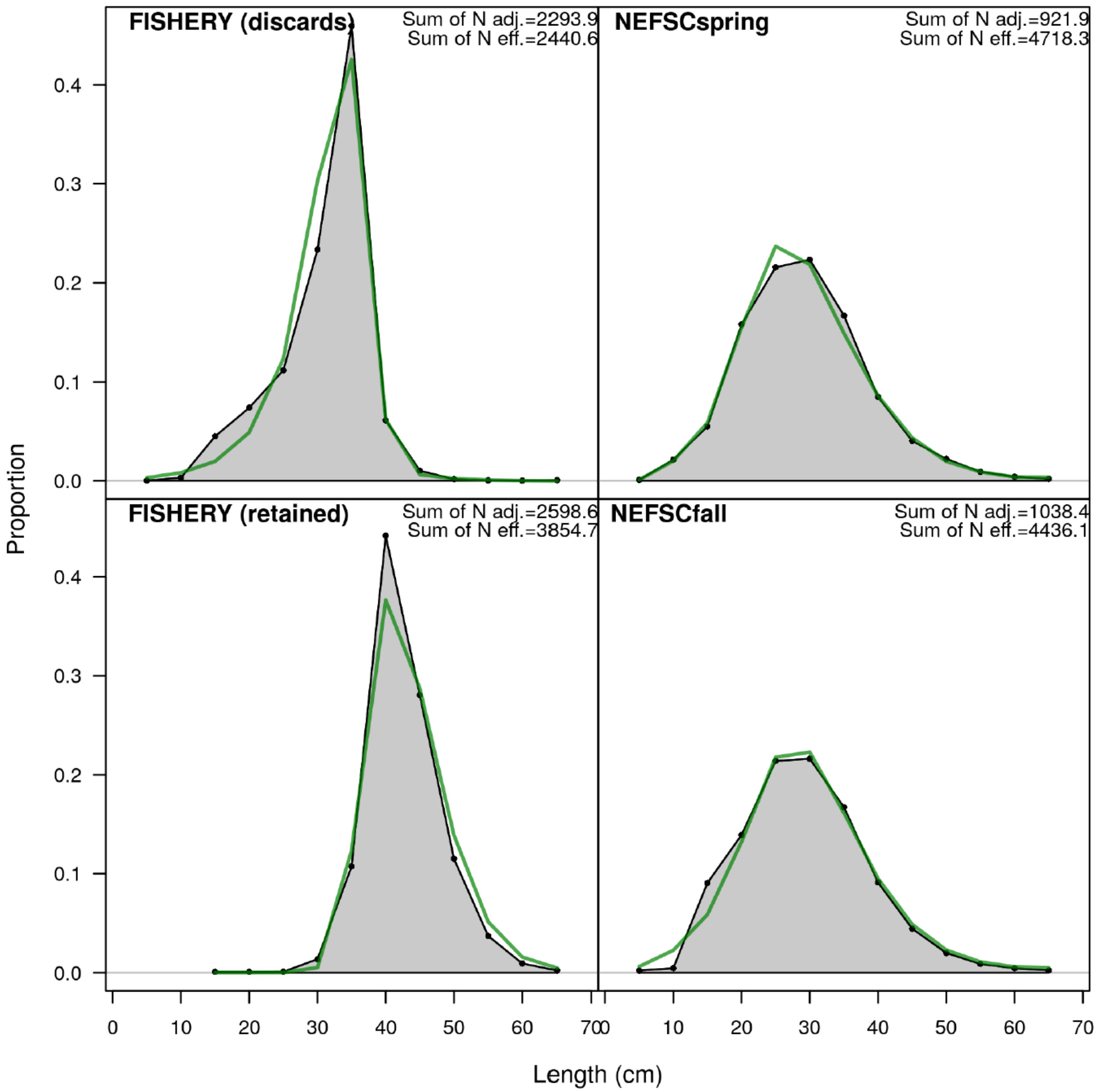


Figure 4.30. SS length comps, aggregated across time by fleet. Labels 'retained' and 'discard' indicate discarded or retained sampled for each fleet.

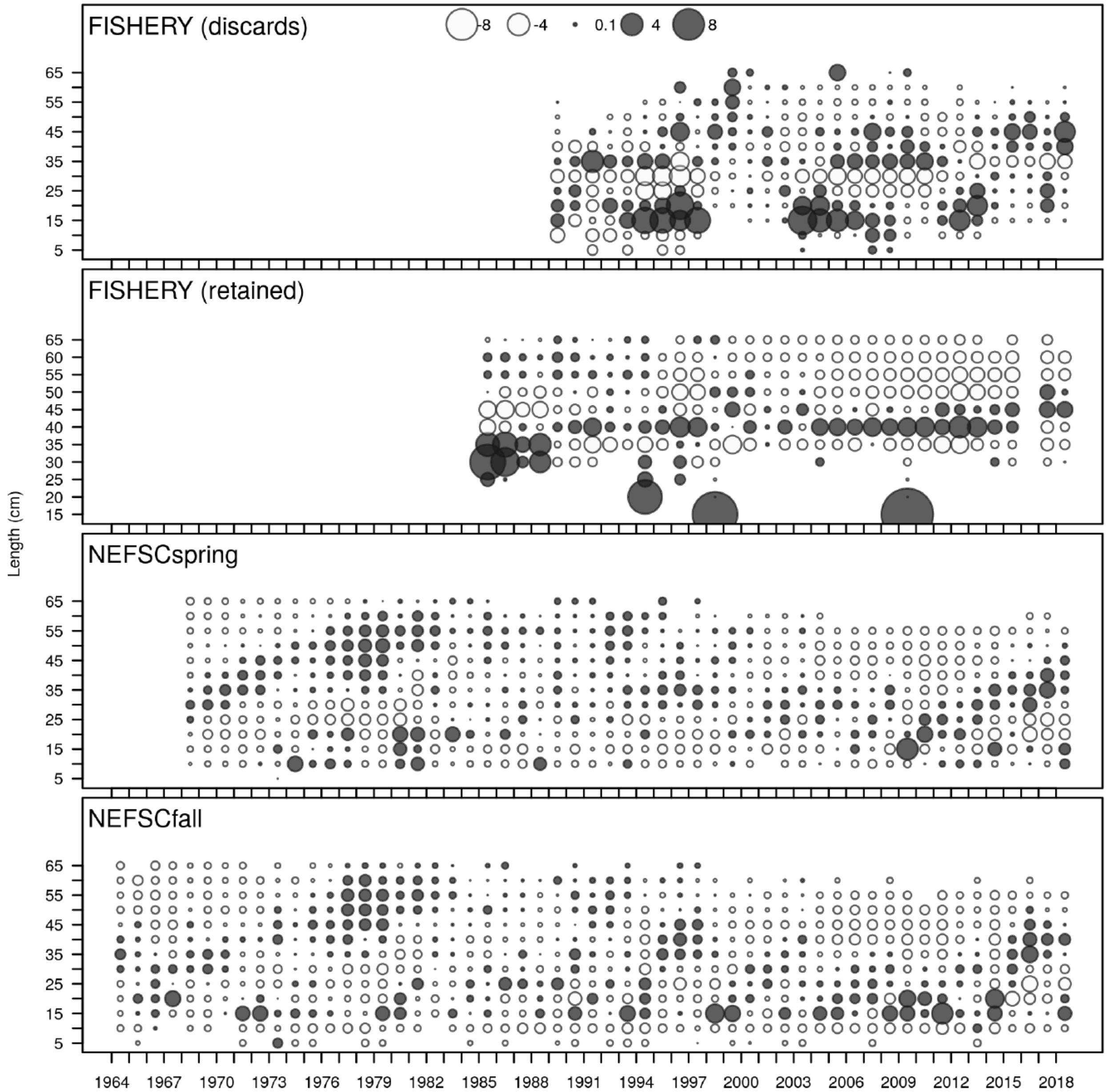


Figure 4.31. SS Pearson residuals, comparing across fleets. Closed bubbles are positive residuals (observed > expected) and open bubbles are negative residuals (observed < expected).

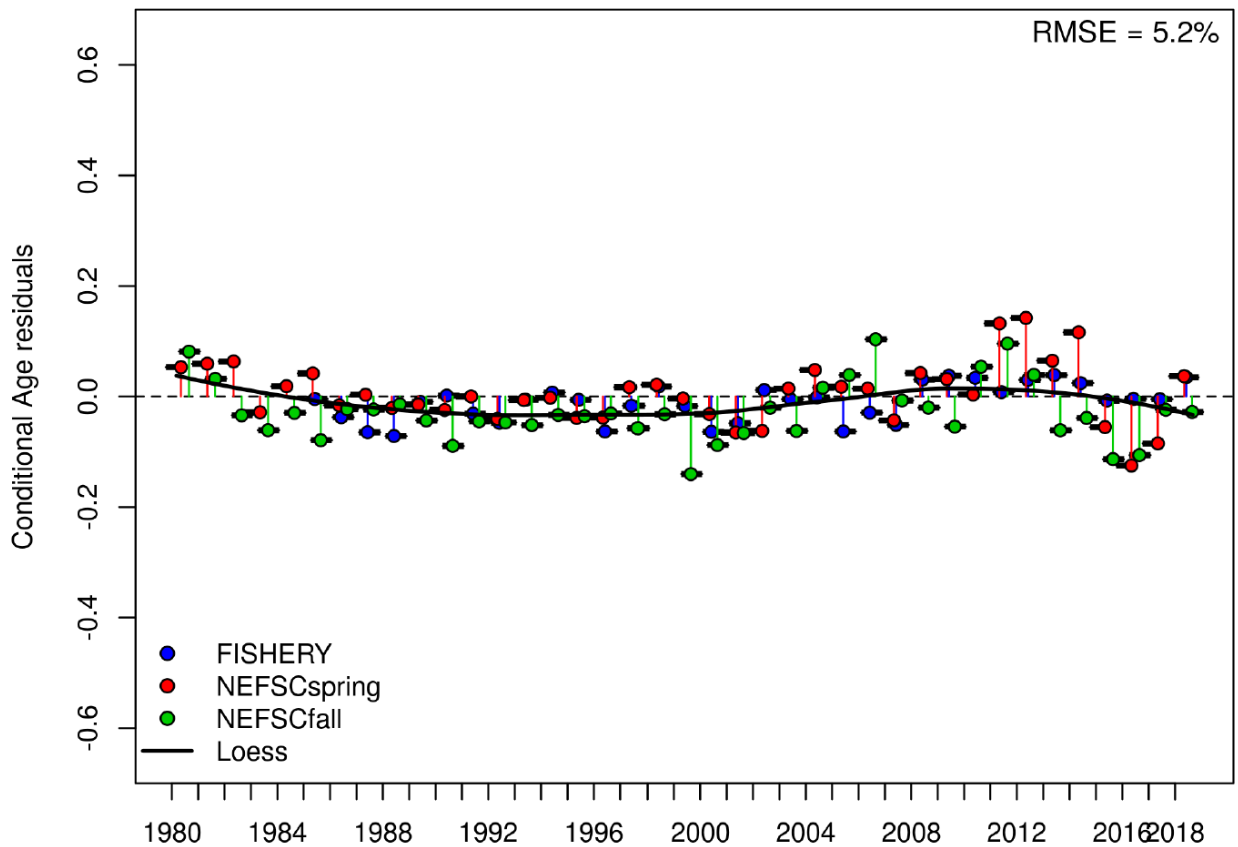


Figure 4.32. SS joint residual plot from fit to annual mean age from conditional age at length composition data.

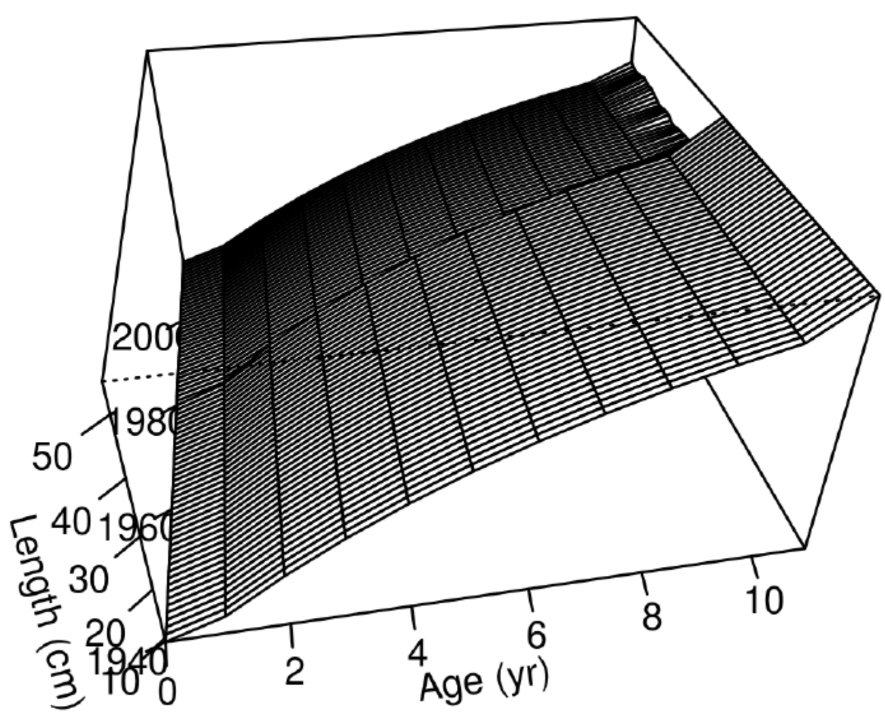


Figure 4.33. SS estimates of time-varying growth.

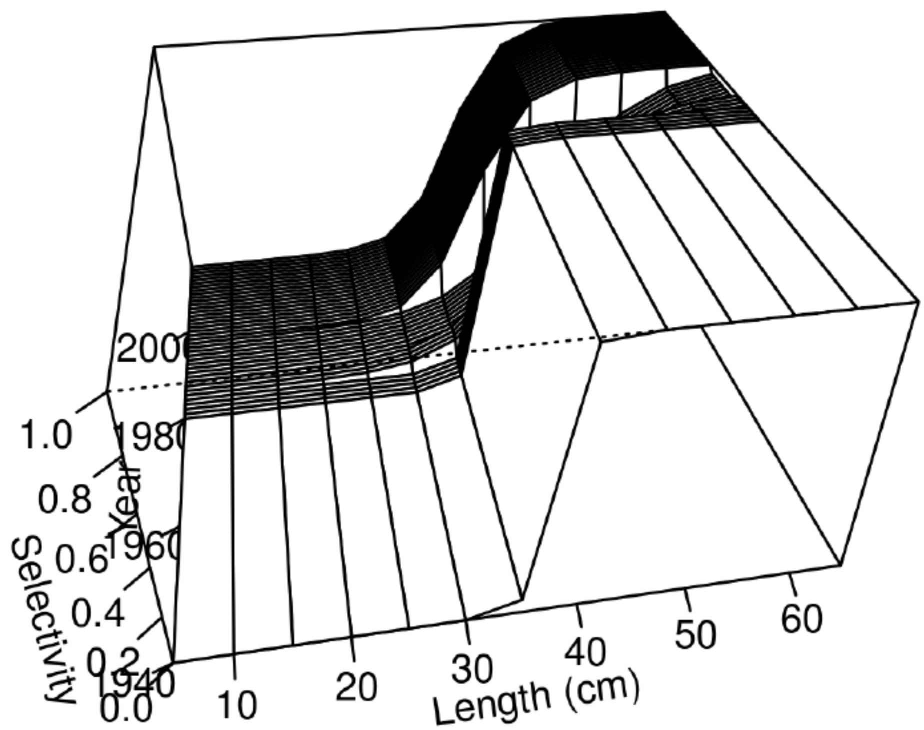


Figure 4.34. SS estimates of time-varying selectivity for the fishery.

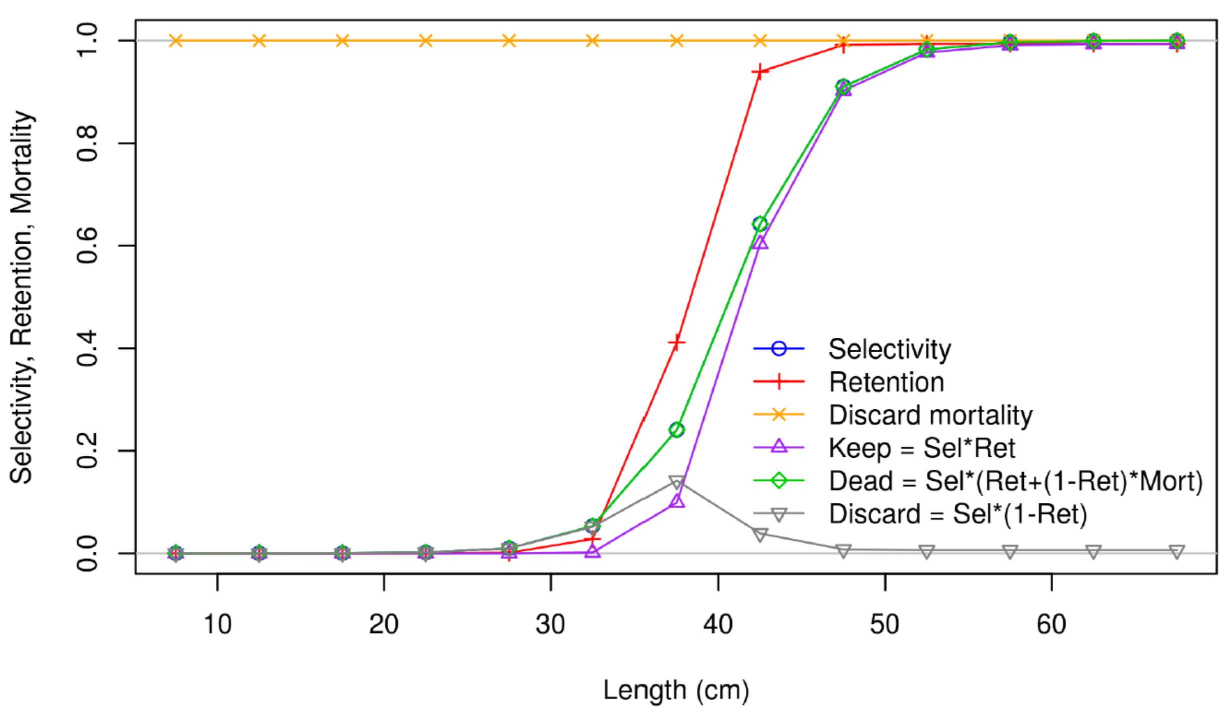


Figure 4.35. SS estimates of selectivity at length in the most recent period for each component of the fishery, including derived discard selectivity.

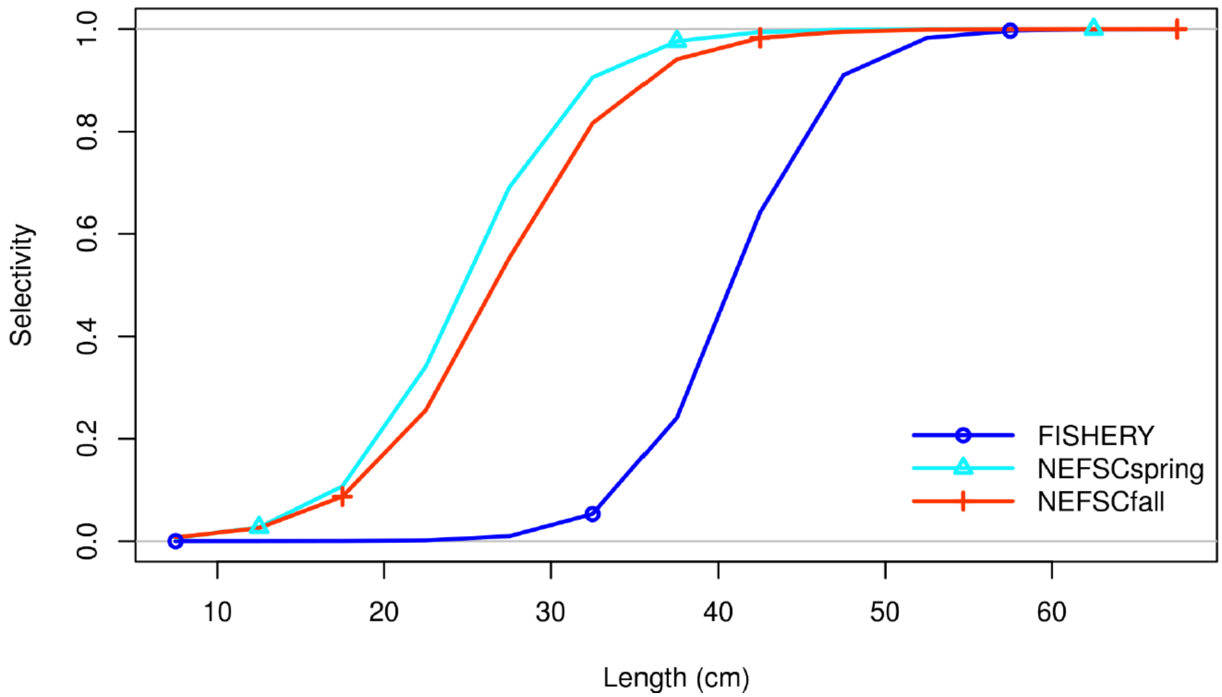


Figure 4.36. SS estimates of selectivity at length for the surveys and fishery in the most recent period.

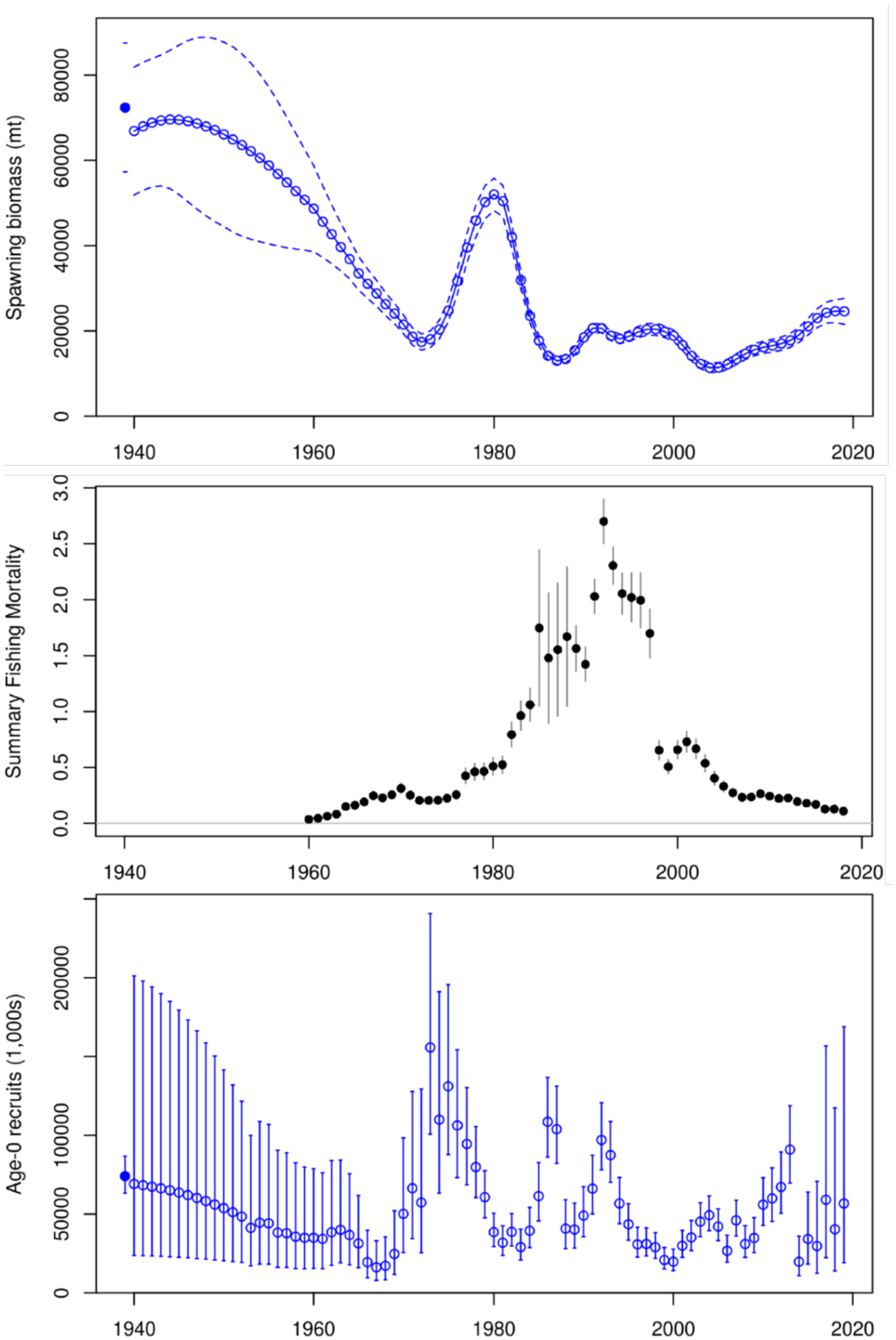


Figure 4.37. SS estimates of spawning biomass (mt), apical fishing mortality, and Age-0 recruits (1,000s) with ~95% asymptotic intervals.

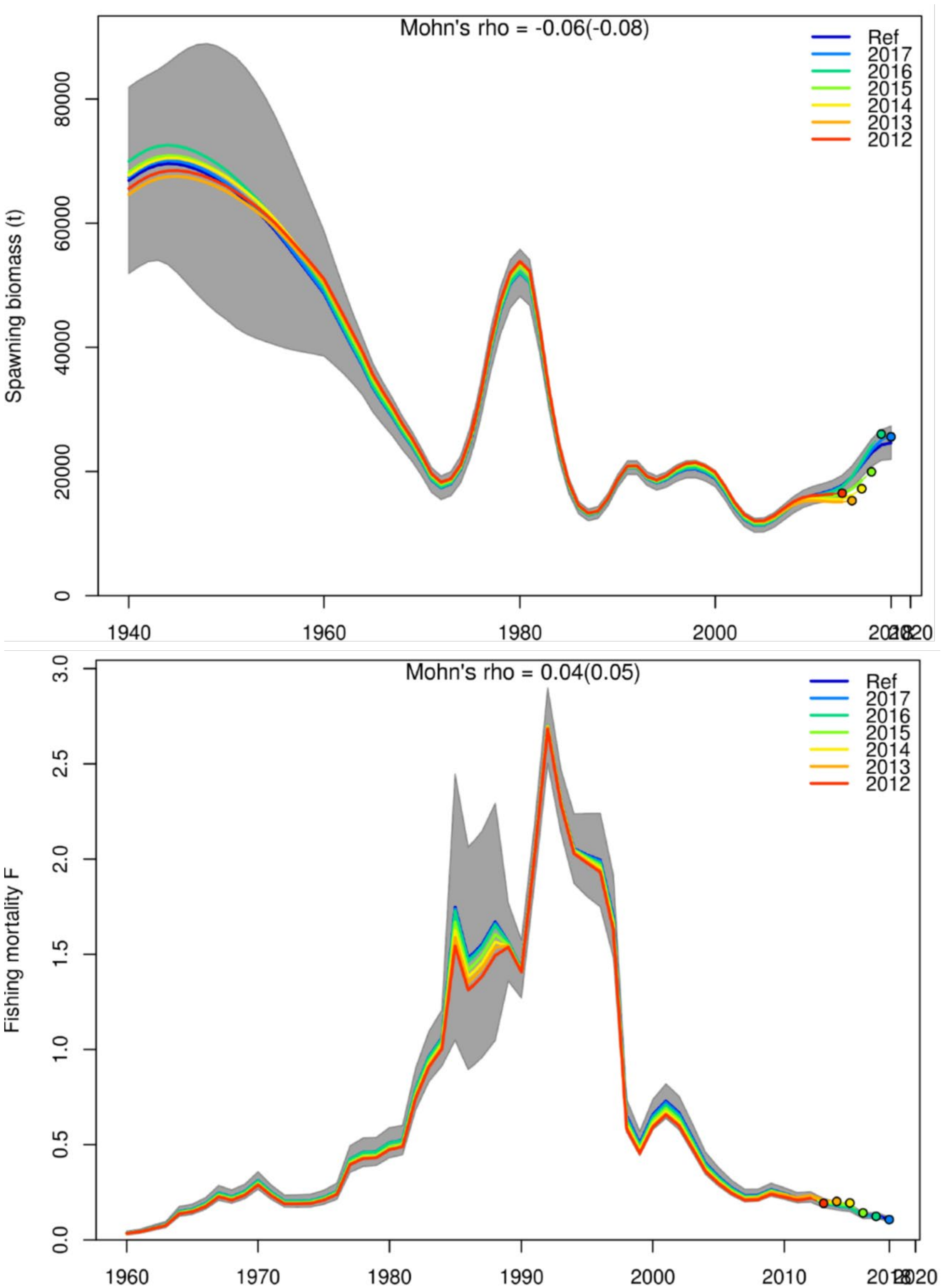


Figure 4.38. Retrospective pattern and Mohn's rho for SSB and F from SS.

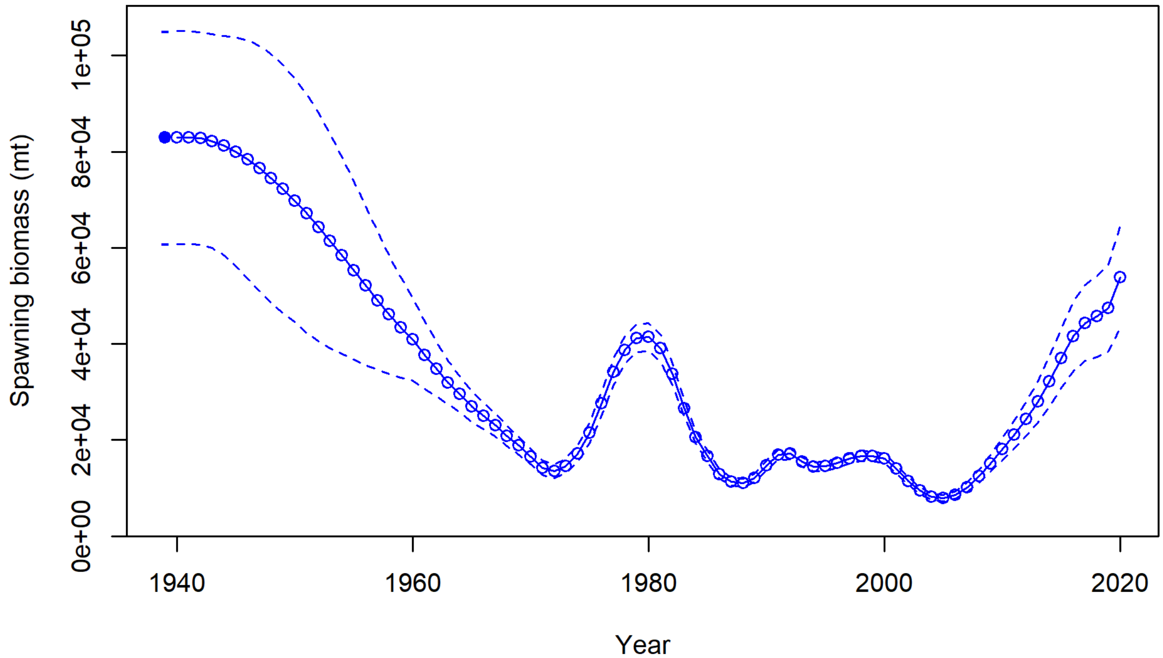


Figure 4.39. SS estimates of spawning biomass (mt) with ~95% asymptotic intervals from a run that modeled NEFSC surveys as separate Albatross and Bigelow indices.

Virtual Population Analysis

Previous stock assessments to American plaice in US waters were based on calibrated VPA (Gavaris 1988). A major source of uncertainty in the most recent assessment was retrospective patterns, so status determinations and short-term projections were based on retrospective-adjusted estimates from the VPA (NEFSC 2022b). Considering these problems with the VPA and research recommendations from several peer reviews (NEFSC 2001, 2015, 2022b), a primary goal of the Working Group was to develop a more advanced stock assessment approach than VPA. However, as the previous benchmark method, the VPA was updated with the revised fishery and survey data presented in ToR2 and ToR3 sections above and revised with new assumptions of natural mortality (M=0.3) and maturity at age to illustrate how those revisions affected VPA results.

The magnitude of retrospective consistency of the 2019 VPA was rho= 0.27 for spawning stock biomass and rho= -0.20 for fishing mortality. The updated VPA had worse retrospective patterns (SSB rho= 0.29, F rho= -0.27; Figure 4.40). The retrospective pattern from the VPA with revised M and maturity was better (SSB rho= 0.09, F rho= -0.11), but substantially worse than those from WHAM, ASAP or SS. Despite the retrospective inconsistency, VPA estimates were similar to those from other models and support the same status determination.

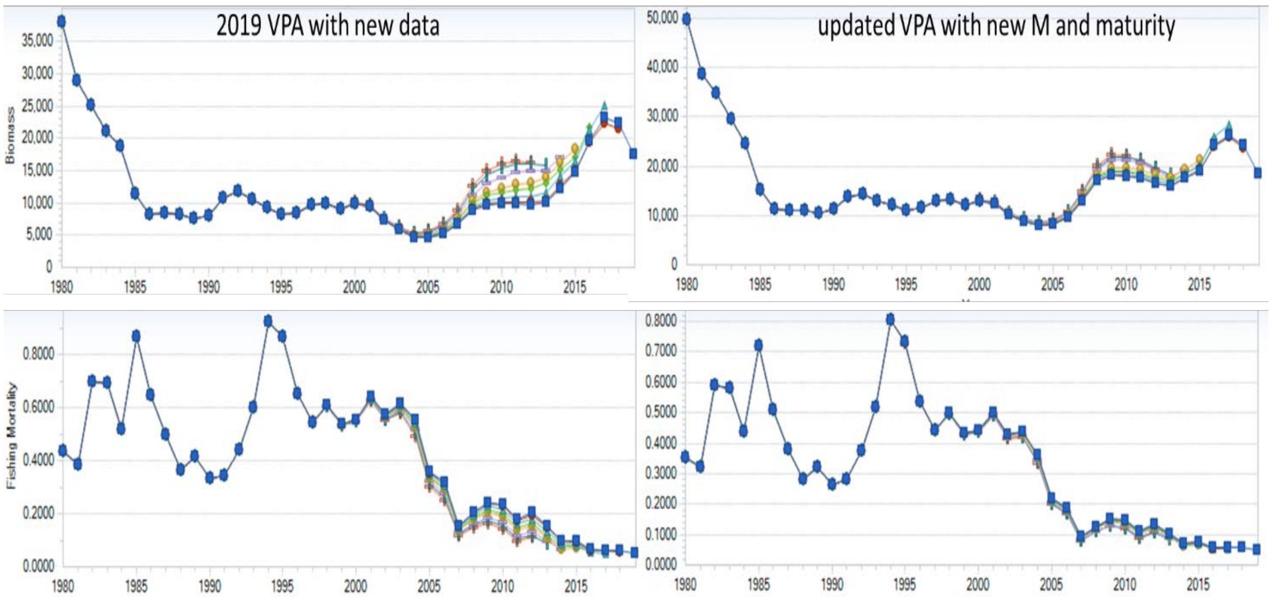


Figure 4.40. Retrospective consistency of American plaice VPAs with revised fishery and survey data (left) and revised with new assumptions of natural mortality (right).

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*."

Contributors:

Steve Cadrin, Amanda Hart and Tim Miller

The objective of the US system for federally managed marine fisheries is to achieve optimum yield, (defined as MSY), while avoiding overfishing and rebuilding overfished stocks (US DOC 2007). MSY reference points or their proxies are required to determine stock status, and MSY should be estimated as "*the largest long-term average catch or yield that can be taken from a stock or stock complex under prevailing ecological, environmental conditions and fishery technological characteristics (e.g., gear selectivity), and the distribution of catch among fleets*" (NOAA 2016). 'Overfishing' status is determined as $F>F_{MSY}$, 'overfished' status is determined as $B<B_{THRESHOLD}$, in which $B_{THRESHOLD}$ is the biomass below which the capacity of the stock to produce MSY on a continuing basis is jeopardized (often $½B_{MSY}$), and B_{MSY} is the rebuilding target.

Throughout its assessment and management history, the fishery for American plaice in US waters has been managed with MSY proxies derived from dynamic pool reference points from yield and spawning biomass per recruit analyses (Thompson and Bell 1934). The first analytical assessment of plaice estimated $F_{0.1}$, F_{max}, and $F_{20\%}$ assuming M=0.2, the most recent average selectivity derived from VPA, and the most recent average weight at age (NEFSC 1992), and the overfishing definition was based on $F_{20\%}$. In 1998, reference points for plaice were re-defined to comply with the current mandate to end overfishing and rebuild stocks, and the overfishing definition for plaice was revised from $F_{20\%}$ to $F_{0.1}$, because the stock-recruit series was too short to estimate F_{MSY}, a SSB_{MSY} proxy was based on average recruitment and spawning potential at $F_{0.1}$, and $SSB_{THRESHOLD}$ was defined as ½SSB_{MSY} (Applegate et al. 1998). In 2002, MSY proxies for American plaice were re-defined based on $F_{40\%}$ (a common proxy for F_{MSY} that is expected to produce at least 75% of MSY for a wide range of life histories and stock-recruit relationships; Clark 1993) and $SSB_{F40\%}$, and the SPR proxies were justified because the 1980-1999 stock recruit relationship was negative (i.e., decreasing recruitment with increasing SSB; NEFSC 2002b). $F_{40\%}$ and $SSB_{F40\%}$ are the proxy reference points used for several New England groundfish stocks (NOAA 2010).

The Working Group re-examined the stock-recruit relationship to confirm the justification for using $F_{40\%}$ and $SSB_{F40\%}$ as proxy reference points. Estimates of recruitment, weight-at-age, and selectivity were examined to determined current prevailing conditions for reference point calculations from WHAM output (https://github.com/ahart1/PlaiceWG2021). Reference points were updated to be consistent with the proposed WHAM run (29F4). Time series of SSB and recruitment estimates do not indicate a strong stock-recruit relationship, because the strongest year-classes (e.g., 1986 and 1991) were produced by relatively low SSB, and year-classes produced by relatively high SSB were average or below average (Figure 5.1). Estimates of SSB and recruitment from longer-term assessments show a similar pattern (Figure 5.2). These results confirm the previous justification for using $F_{40\%}$ and $SSB_{F40\%}$ as proxy reference points.

Environmental factors were considered for determining the appropriate time frame of recruitment to inform biological reference points of plaice. The butterfish research track stock assessment recognized a shift in condition of butterfish in 2011 which was associated with increased bottom temperature, indicating that ocean warming to date does not appear to be negatively impacting butterfish health and may currently be benefiting butterfish (NEFSC 2022a). This finding justified the use of recent average recruitment (2011-2019) to derive reference points for butterfish. The examination of plaice condition (ToR1) identified a similar temporal pattern in which condition increased around 2011. Exploratory analyses suggested a positive relationship between plaice condition and AMO, suggesting that warmer water temperatures negatively impact plaice health. The examination of recruitment rate over time also suggests increased recruitment under warming conditions, but condition decreased in the most recent four years. Therefore, the plaice Working Group decided to retain the entire time series of recruitment in the characterization to inform biological reference points. Recognizing the changes in fish condition (Figure 1.10) and weight-at-age (Figure 2.17), the last five years of weight-at-age represent the relatively stable weights at ages 1-6 for the entire time series and the current period of relatively lighter weight at ages 7+.

WHAM estimates of selectivity-at-age suggest that the last five years represent the current selectivity pattern (Figure 5.3). Therefore, selectivity and weights-at-age for reference points were from the last five years of the stock assessment (2015-2019; Table 5.1), and maturity-at-age from the entire time series of NEFSC spring survey observations, as assumed in stock assessment models (Table 4.1). Considering the lack of a stock-recruit relationship, and the precision of recruitment estimates (CV=0.25 to 0.3; Figure 5.4), the Working Group decided to use the entire time series of recruitment for reference point estimates. These assumptions are consistent with the approach applied to other groundfish stocks in the region (NEFSC 2022b).

The integrated estimate of $F_{40\%}$ is 0.42 from WHAM run 29F4, assuming M=0.3, and the most recent five years of selectivity and weight-at-age (Figure 5.5). The associated estimate of $SSB_{F40\%}$ is 18,000 mt assuming the entire time series of recruitment. Over the time periof of the stock assessment, reference point estimates varied over time because of time-varying weight and selectivity-at-age, (Figure 5.6).

Based on WHAM estimates, historical overfishing (1980-1998) depleted the stock to be overfished in the late 1980s, but fishing mortality has been less than $F_{40\%}$ since 1998, and the stock rebuilt to be significantly greater than $SSB_{F40\%}$ in 2019 (Figure 5.7). According to these analyses, there is high probability that the current stock is not overfished, and overfishing is not occurring (Figure 5.8).

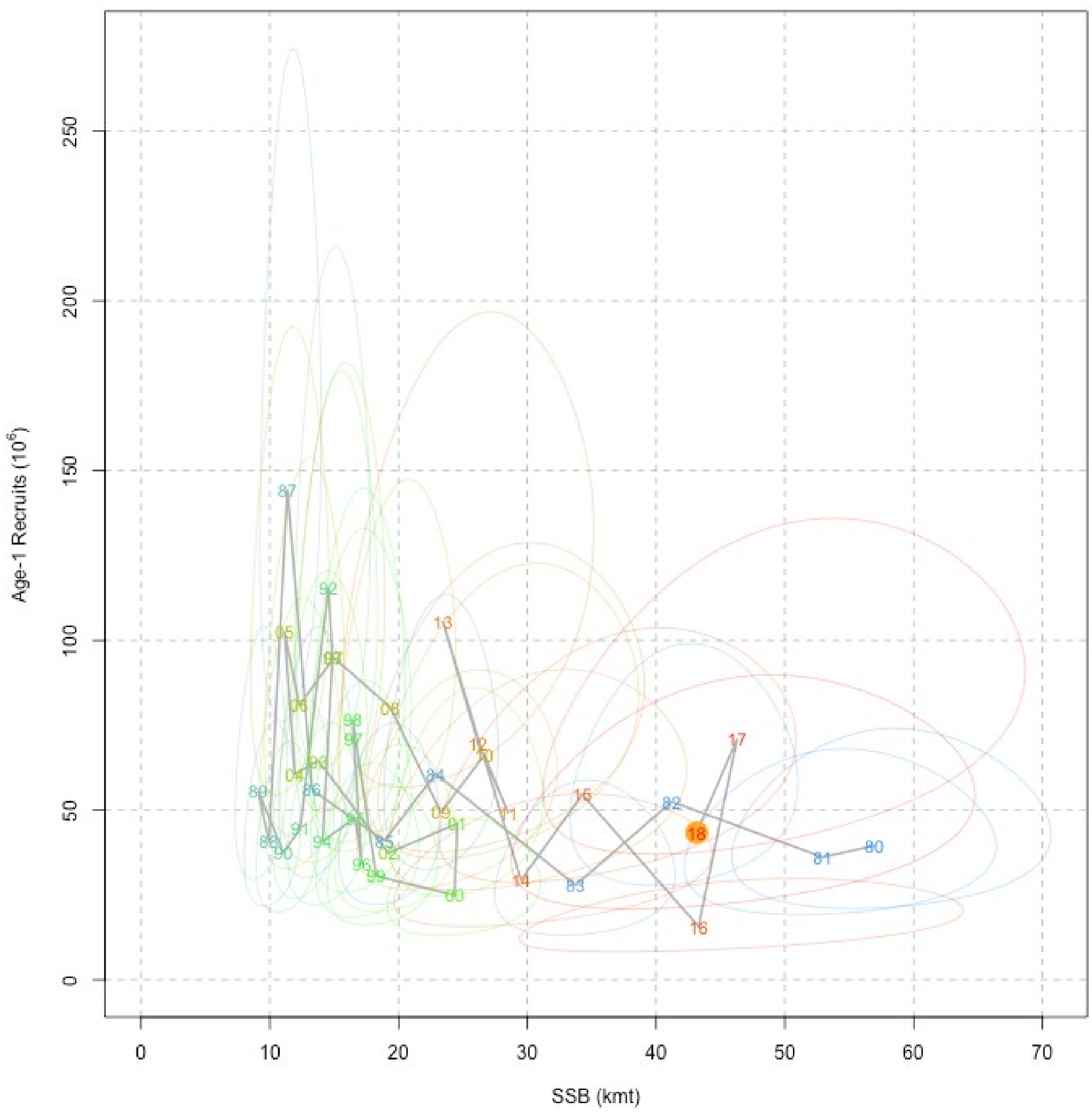


Figure 5.1. WHAM estimates of American plaice SSB and recruitment, labeled by year-class (year at age-1) with joint confidence bounds.

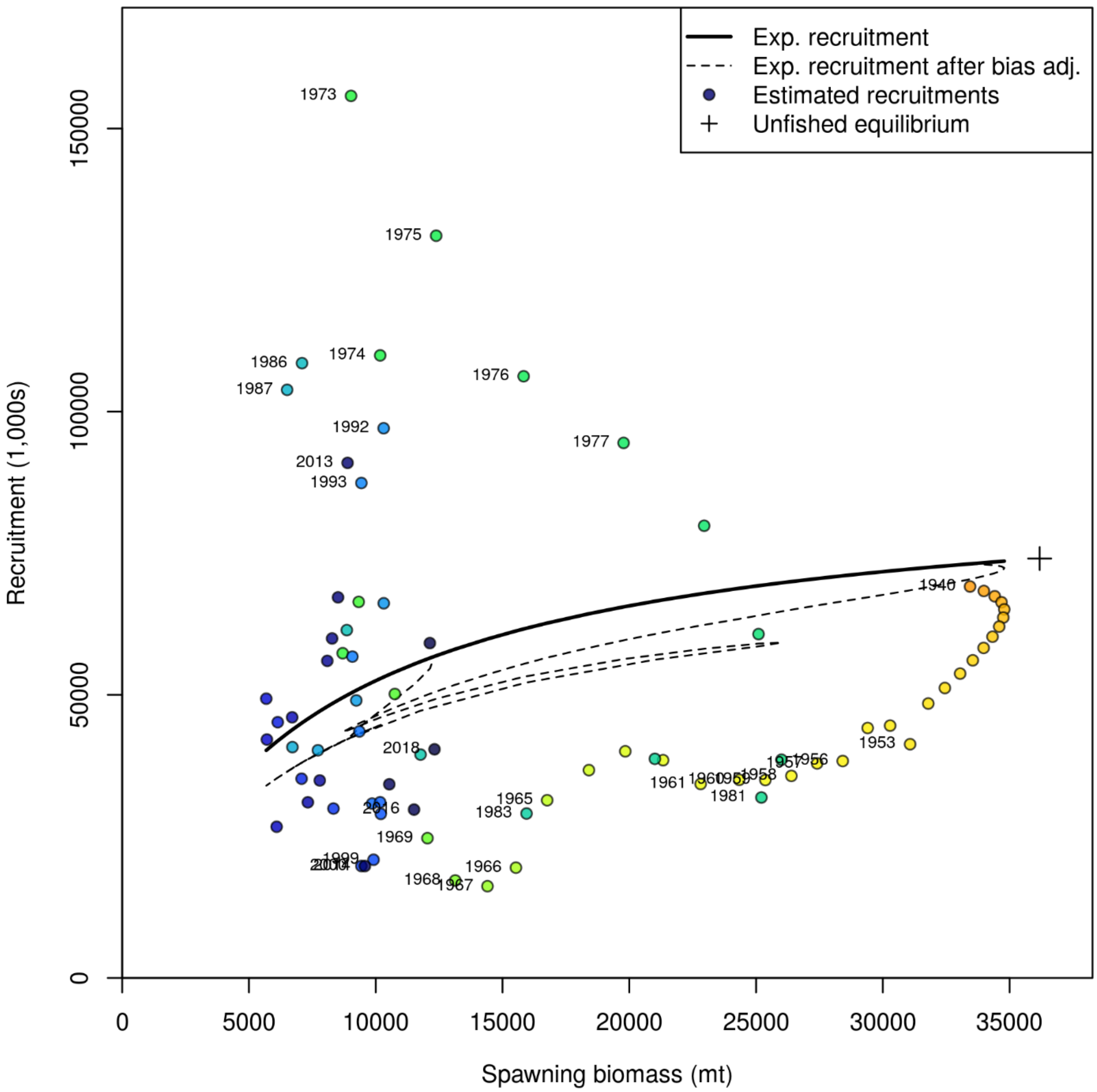


Figure 5.2. SS estimates of American plaice SSB and recruitment labeled 1940-2020, labeled by year-class (year at age-0).

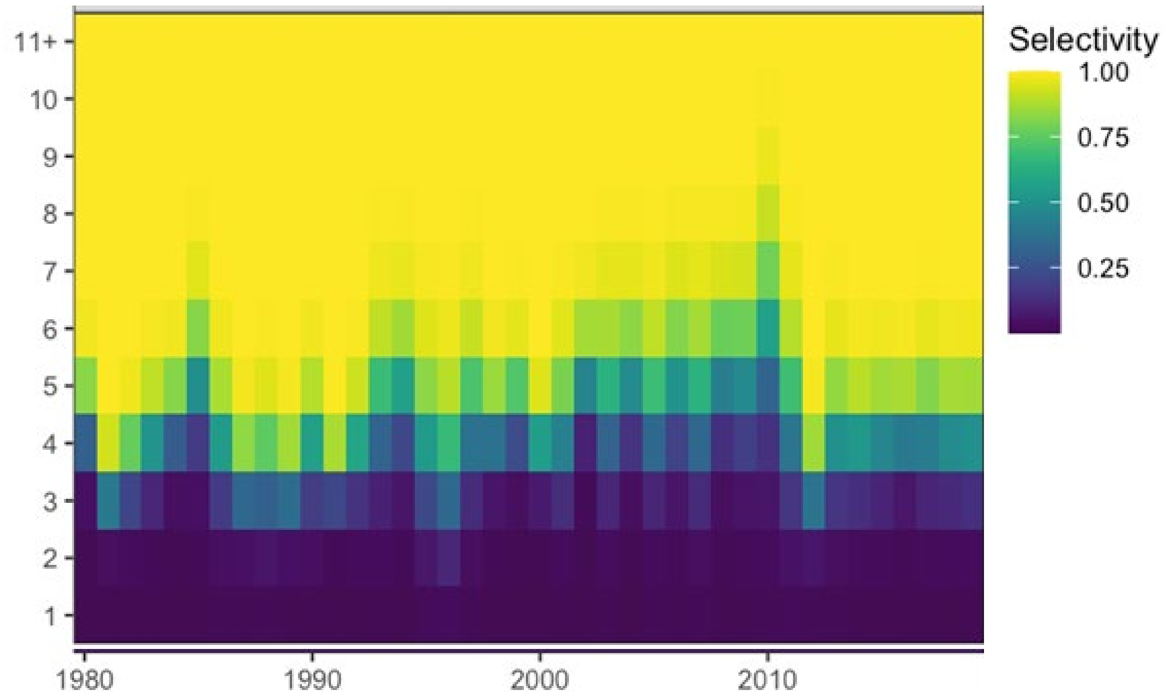


Figure 5.3. Estimates of fishery selectivity at age for American plaice from WHAM run 29F4.

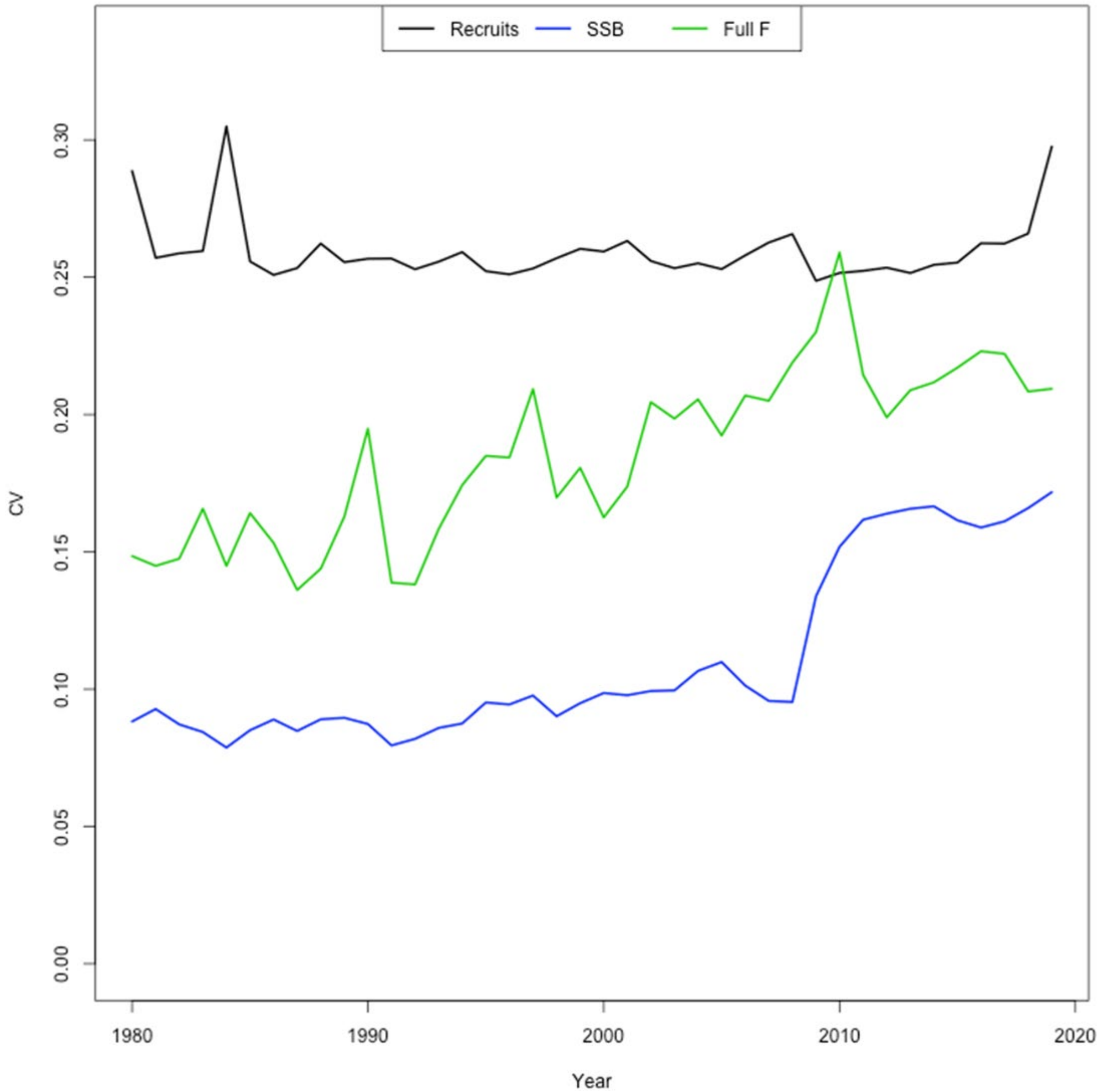


Figure 5.4. Precision of American plaice SSB, fully-selected F and recruitment estimates from WHAM run 29F4.

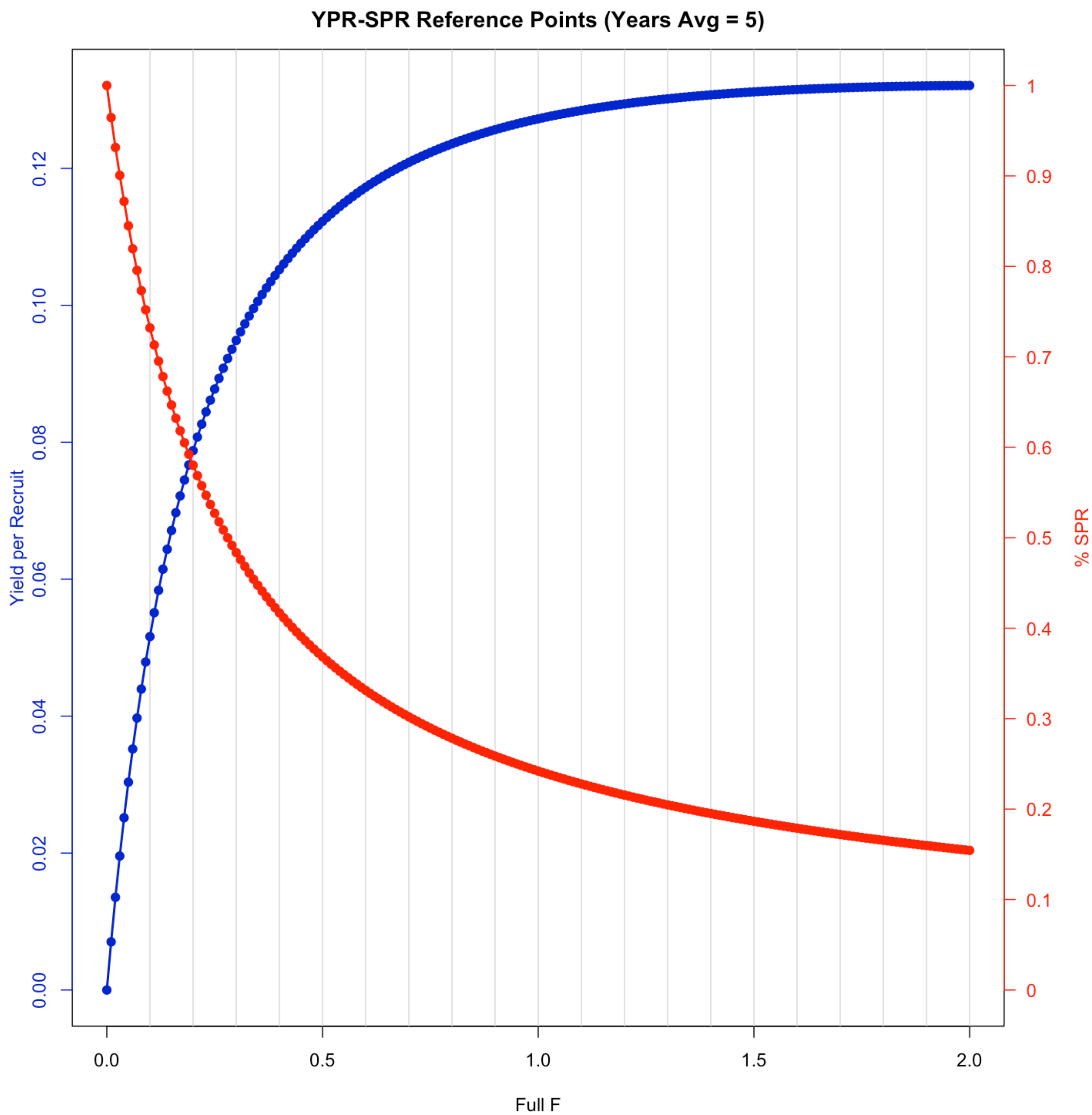


Figure 5.5. Yield (kg) and spawning biomass per recruit (% of unfished) for American plaice based on WHAM run 29F4.

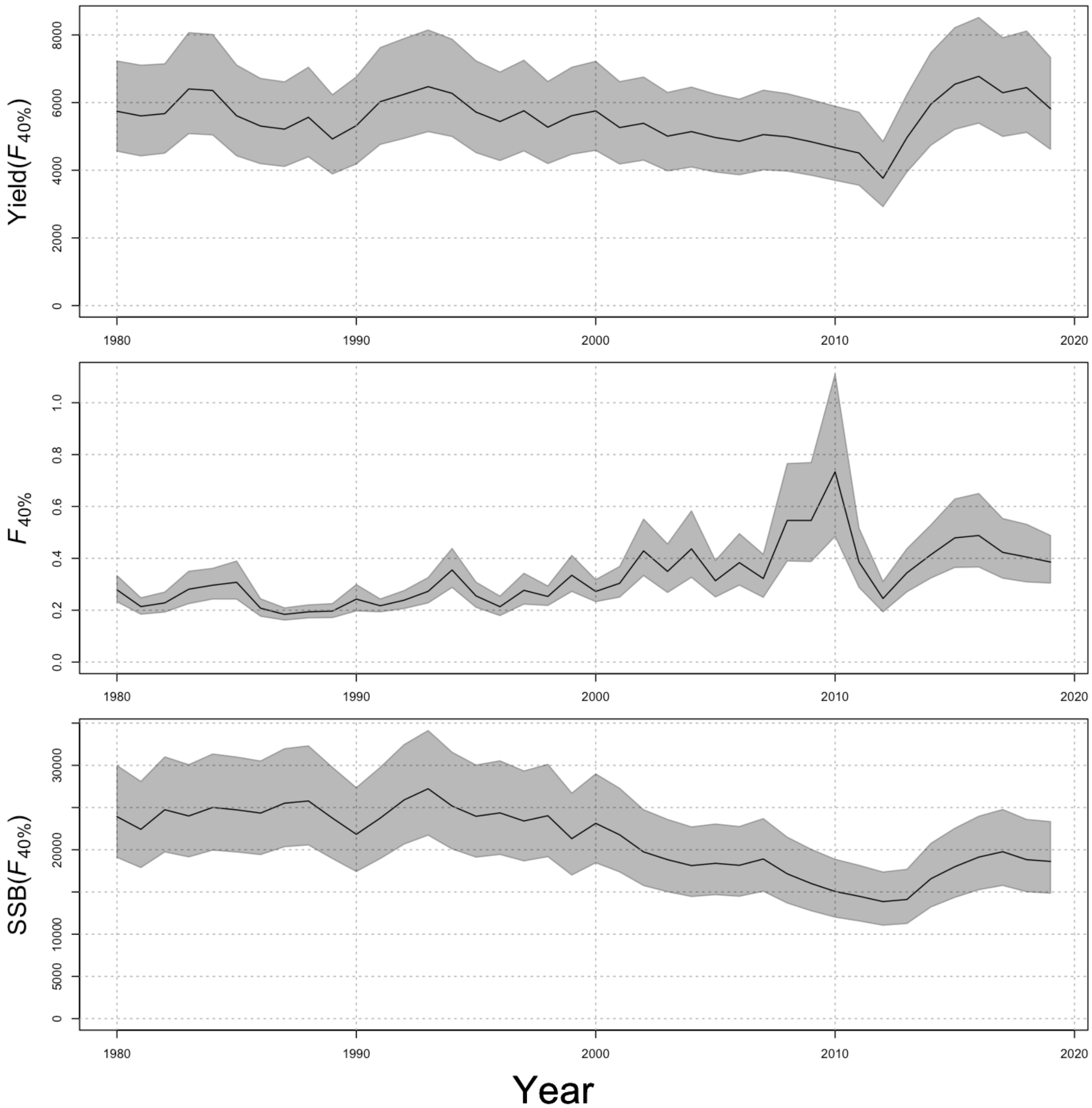


Figure 5.6. Time varying estimates of $F_{40\%}$, $SSB_{F40\%}$ and long-term yield at $F_{40\%}$ for American plaice based on WHAM run 29F4.

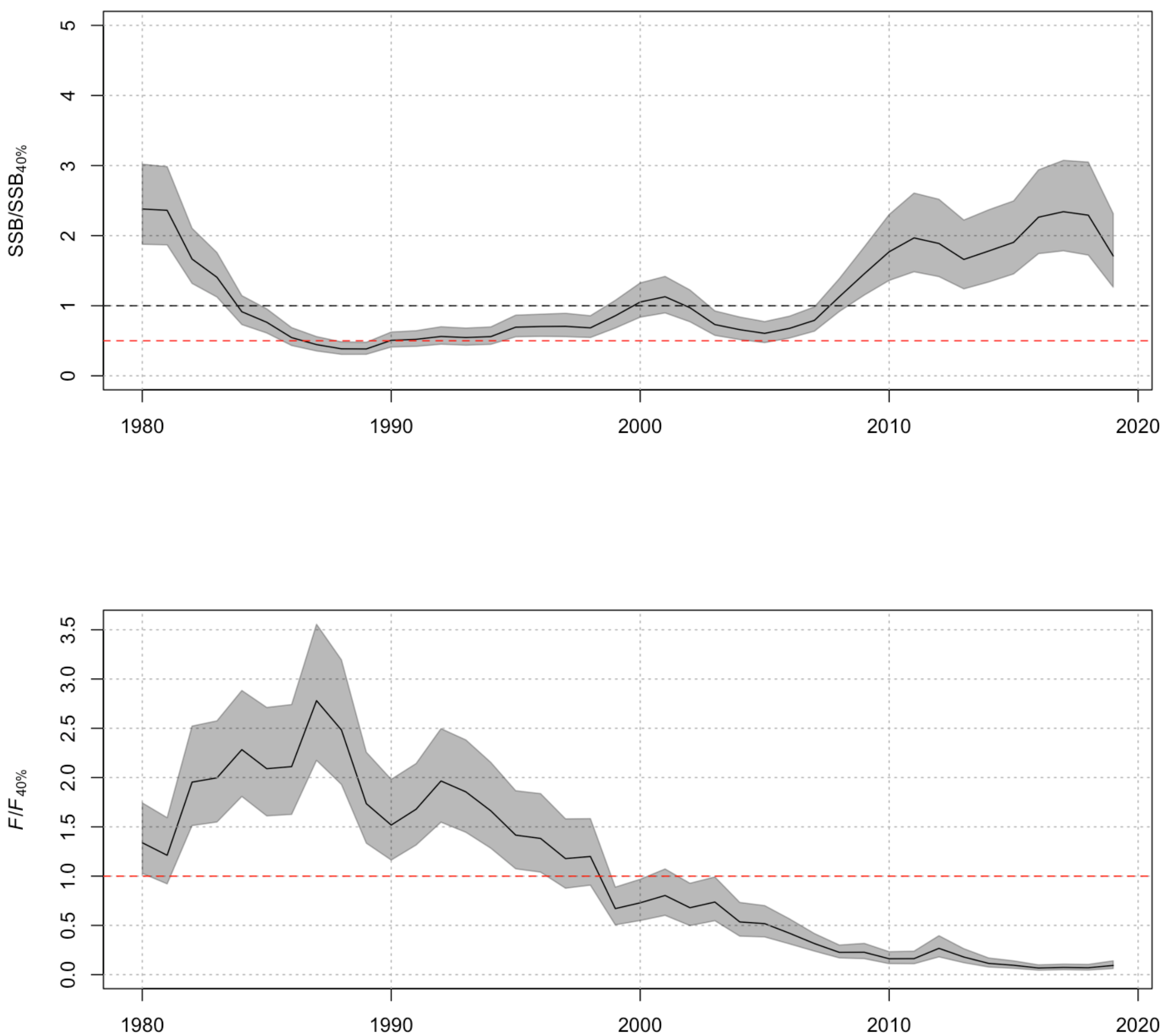


Figure 5.7. Relative stock status (top) and exploitation status (bottom) of American plaice based on WHAM run 29F4. Black dashed lines represent the rebuilding target (top) or the overfishing threshold (bottom). The red line indicates the overfished threshold.

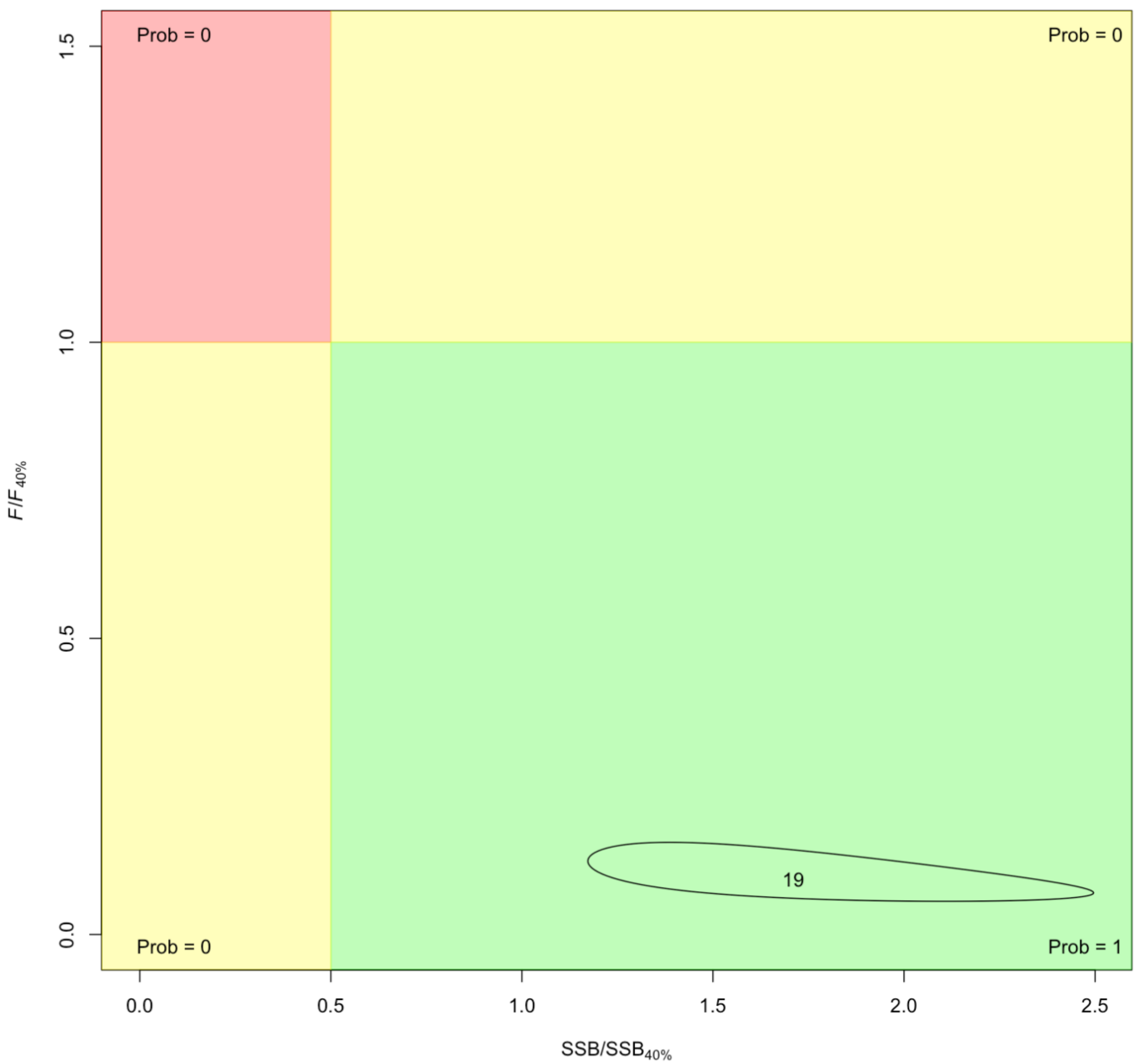


Figure 5.8. Relative stock status (SSSB/$SSB_{F40\%}$) and exploitation status (F/$F_{40\%}$) of American plaice in 2019 based on WHAM run 29F4, with probability (Prob.) for each quadrant.

Table 5.1. American plaice maturity, selectivity and weights at age assumed in reference point estimates.

Age	Maturity	Selectivity	Catch Wts	Jan-1 Wts	SSB Wts
Age-1	0.000	0.002	0.034	0.019	0.025
Age-2	0.140	0.016	0.116	0.061	0.084
Age-3	0.440	0.104	0.237	0.167	0.199
Age-4	0.750	0.451	0.396	0.313	0.352
Age-5	0.890	0.856	0.516	0.457	0.485
Age-6	0.960	0.977	0.602	0.554	0.577
Age-7	0.980	0.997	0.654	0.625	0.639
Age-8	0.990	0.999	0.726	0.680	0.702
Age-9	0.990	1.000	0.773	0.740	0.756
Age-10	1.000	1.000	0.786	0.770	0.778
Age-11+	1.000	1.000	0.916	0.916	0.916

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*."

Contributors:

Larry Alade, Amanda Hart, Tim Miller, and Steve Cadrin

The US system for federally managed marine fisheries requires annual catch limits that are expected to avoid overfishing, and overfishing is defined as $F>F_{MSY}$ (US DOC 2007). This mandate is implemented through short-term projections to determine an overfishing limit and a precautionary 'acceptable biological catch' (NOAA 2016). The expected catch from short-term projections that assume future $F=F_{MSY}$ (or the F_{MSY} proxy) are used to determine the overfishing limit. Regional fishery councils determine an acceptable catch control rule that is expected to avoid overfishing. For American plaice and other New England groundfish, the acceptable catch control rule is based on short-term projections that assume future $F=75\%F_{40\%}$ for stocks that are not overfished (NOAA 2010). The management procedure for New England groundfish requires that projections are based on retrospective-adjusted abundance at age if the stock assessment has 'major' retrospective patterns (i.e., retrospective inconsistency that is greater than estimated confidence limits; NEFSC 2008).

Projection results through 2022 are presented for demonstration, but the WHAM application will be updated with 2020-2021 data in the 2022 management track process, and projections will also be updated. Stochastic projections of stock size and catches for 2020-2023 are consistent with the new biological reference points proposed for the research track assessment (ToR5). WHAM (run 29F4) was used to produce integrated projections that account for uncertainty in all estimated parameters, 2020 abundance at age, and future recruitment, including uncorrelated process variance in survival and recruitment. No retrospective adjustment was applied, because retrospective inconsistency was well within confidence limits. Projections assumed future recruitment based on the entire time series of recruitment and recent 5-year

(2015-2019) estimates of selectivity and observations of weight-at-age, and maturity-at-age from the entire time series to represent current conditions (i.e., the same as assumed for reference point calculations, Table 5.1), which is consistent with the approach applied to other groundfish stocks in the region (NEFSC 2022b). Provisional projections assumed catches of 663 mt in 2020 and 708 mt in 2021 based on estimates used for quota monitoring. Projections scenarios included 1) fishing at the estimate of $F_{40\%}$ in 2022 to demonstrate overfishing limit projections, 2) fishing at 75%$F_{40\%}$ in 2022 to demonstrate acceptable catch projections, 3) fishing at the estimate of 2019 F in 2022 to demonstrate a status quo F scenario, and 4) no fishing, as a basis for comparison to harvest scenarios.

All provisional projections resulted in an increase in catch by 2022 (Tables 6.1 -6.4; Figures 6.1-6.3). Projected spawning stock biomass decreased for all 2022 F>0 scenarios but remained well above the $SSB_{F40\%}$ (Figure 6.2). The projected stock decrease is related to decreased recent recruitment (Figure 4.18).

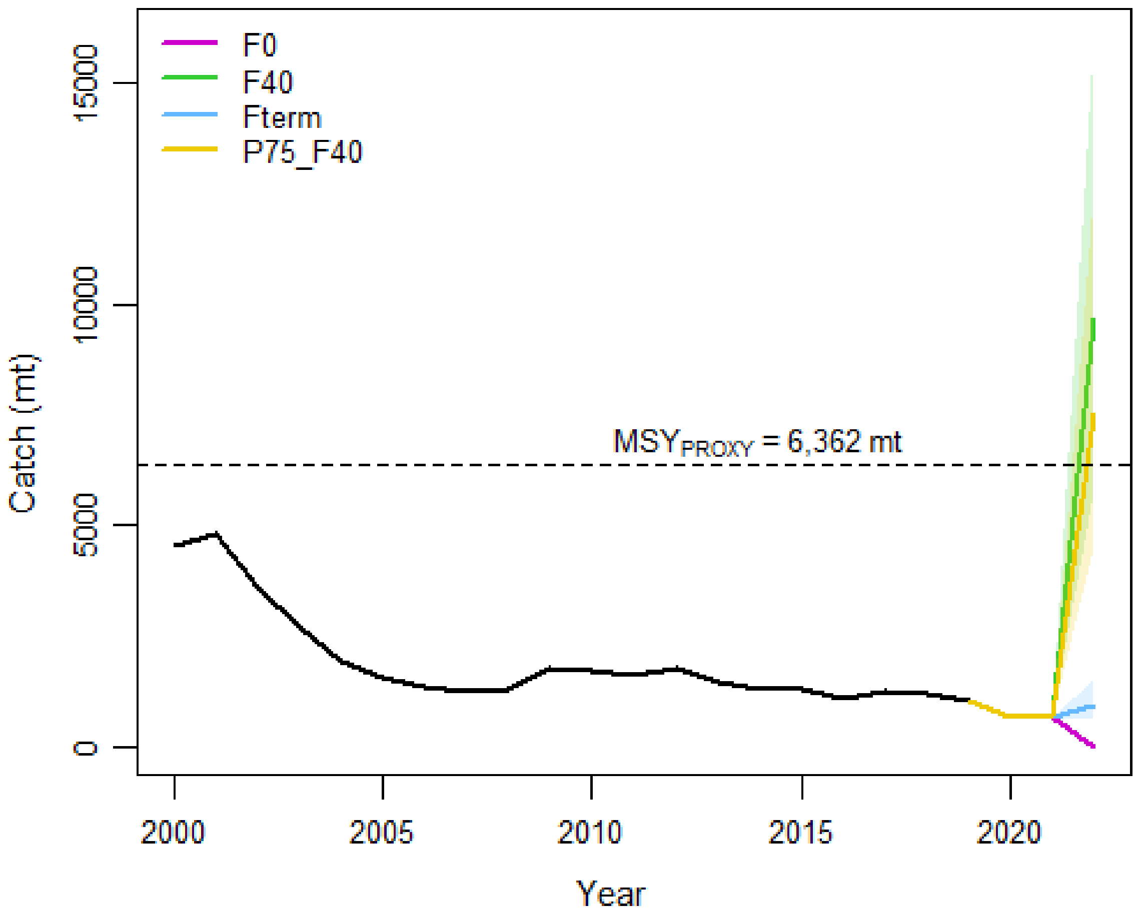


Figure 6.1. Projections of American plaice catch expected from four scenarios of fishing in 2022 (F0: no fishing; F40: fishing at the $F_{40\%}$; Fterm: fishing at 2019 F; and P75_F40: fishing at 75% of $F_{40\%}$).

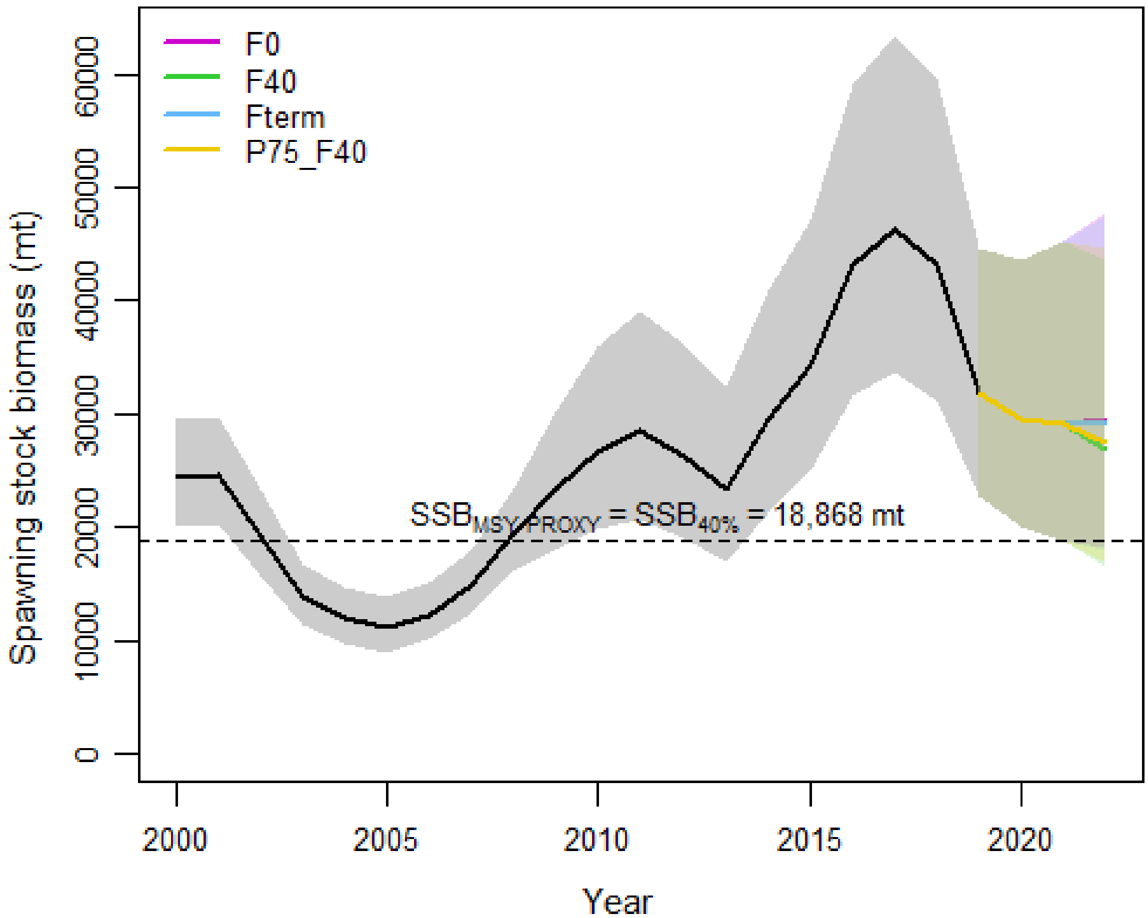


Figure 6.2. Projections of American plaice spawning stock biomass (SSB) expected from four scenarios of fishing in 2022 (F0: no fishing; F40: fishing at the $F_{40\%}$; Fterm: fishing at 2019 F; and P75_F40: fishing at 75% of $F_{40\%}$).

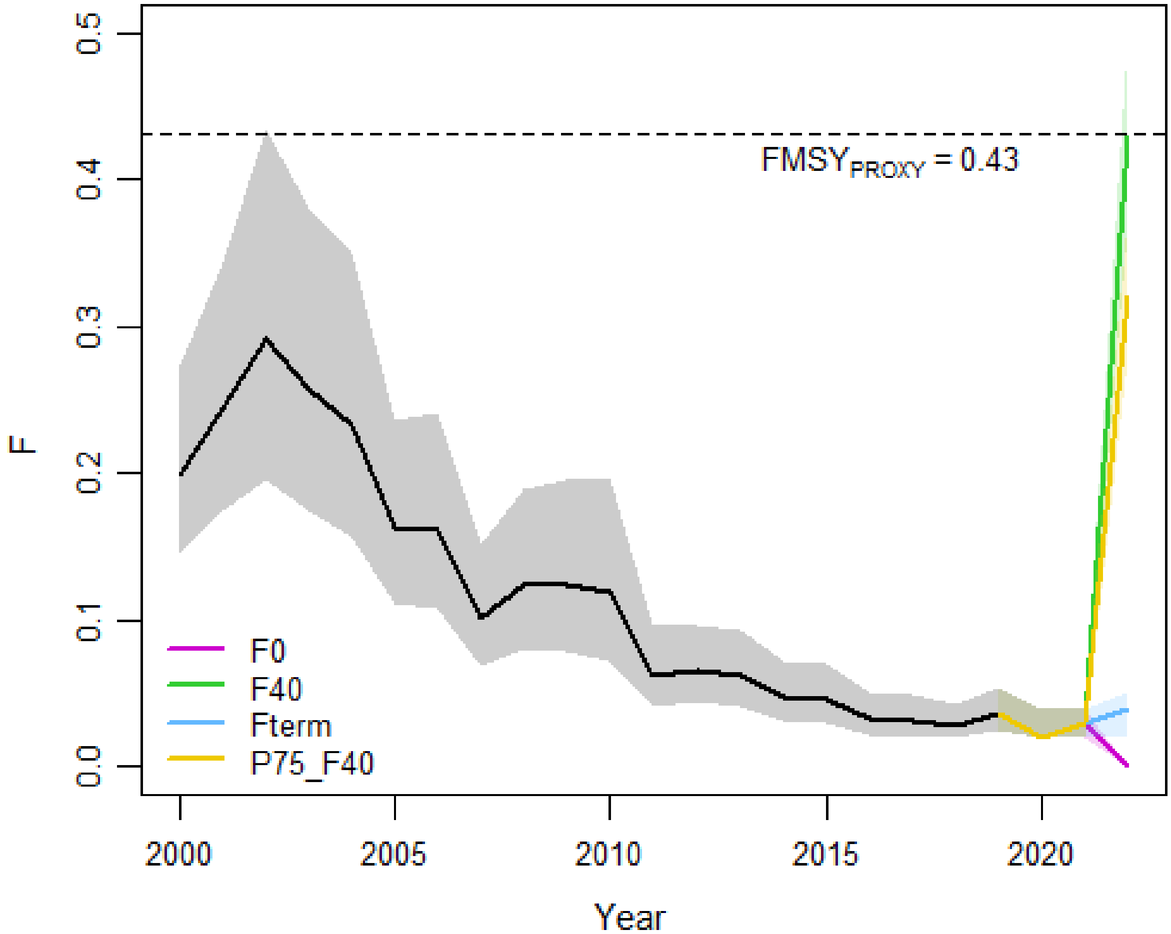


Figure 6.3. Projections of American fishing mortality expected from four scenarios of fishing in 2022 (F0: no fishing; F40: fishing at the $F_{40\%}$; Fterm: fishing at 2019 F; and P75_F40: fishing at 75% of $F_{40\%}$).

Table 6.1. Projections of American spawning stock biomass (SSB), recruitment, catch and fishing mortality expected from fishing at $F_{40\%}$ in 2022.

		2020	2021	2022
SSB(mt)	Estimate	29,553	29,186	26,938
	lower CI	20,001	18,856	16,607
	Upper CI	43,668	45,175	43,695
Age 1 Rec (000's)	Estimate	52,050	52,050	52,050
	lower CI	18,491	18,491	18,491
	Upper CI	146,515	146,515	146,515
Catch (mt)	Estimate	663	708	9,657
	lower CI	663	708	5,807
	Upper CI	663	708	16,058
Fishing Mortality (F)	Estimate	0.020	0.030	0.430
	lower CI	0.020	0.020	0.370
	Upper CI	0.040	0.040	0.500

Table 6.2. Projections of American spawning stock biomass (SSB), recruitment, catch and fishing mortality expected from fishing at 75%$F_{40\%}$ in 2022.

		2020	2021	2022
SSB(mt)	Estimate	29,553	29,186	27,545
	lower CI	20,001	18,856	16,979
	Upper CI	43,668	45,175	44,687
Age 1 Rec (000's)	Estimate	52,050	18,491	146,515
	lower CI	52,050	18,491	146,515
	Upper CI	52,050	18,491	146,515
Catch (mt)	Estimate	663	708	7,556
	lower CI	663	708	4,540
	Upper CI	663	708	12,574
Fishing Mortality (F)	Estimate	0.020	0.030	0.320
	lower CI	0.020	0.020	0.280
	Upper CI	0.040	0.040	0.380

Table 6.3. Projections of American spawning stock biomass (SSB), recruitment, catch and fishing mortality expected from fishing at 2019 F in 2022.

		2020	2021	2022
SSB(mt)	Estimate	29,553	29,186	29,237
	lower CI	20,001	18,856	17,985
	Upper CI	43,668	45,175	47,529
Age 1 Rec (000's)	Estimate	52,050	52,050	52,050
	lower CI	18,491	18,491	18,491
	Upper CI	146,515	146,515	146,515
Catch (mt)	Estimate	663	708	963
	lower CI	663	708	610
	Upper CI	663	708	1,519
Fishing Mortality (F)	Estimate	0.020	0.030	0.040
	lower CI	0.020	0.020	0.020
	Upper CI	0.040	0.040	0.050

Table 6.4. Projections of American spawning stock biomass (SSB), recruitment, catch and fishing mortality expected from fishing at F=0 in 2022.

		2020	2021	2022
SSB(mt)	Estimate	29,553	20,001	43,668
	lower CI	29,186	18,856	45,175
	Upper CI	29,459	18,148	47,821
Age 1 Rec (000's)	Estimate	52,050	52,050	52,050
	lower CI	18,491	18,491	18,491
	Upper CI	146,515	146,515	146,515
Catch (mt)	Estimate	663	708	-
	lower CI	663	708	-
	Upper CI	663	708	-
Fishing Mortality (F)	Estimate	0.020	0.030	-
	lower CI	0.020	0.020	-
	Upper CI	0.040	0.040	-

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 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*."

Status of Previous Research Recommendations

Research recommendations from previous stock assessments of American plaice were compiled, and the status of each recommendation was evaluated by the Working group.

1st Northeast Stock Assessment Workshop (NEFC 1985)

- 'Stock structure is currently unresolved' – NEFSC (1986) responded that 'although definitive studies of stock structure (e.g., tagging programs, genetic studies) have not been undertaken, trends in abundance from the Gulf- of Maine and Georges Bank appear identical over time.' This research track assessment reviewed the limited information available, which suggests separate phenotypic stock of American plaice in US waters that grow and mature more rapidly and have independent recruitment than those in Canadian waters. Although there are growth differences between the Gulf of Maine and Georges Bank, trends in abundance are similar.
- 'Reliable recruitment indices need to be developed' – Recruitment indices were included in the first analytical assessment (NEFC 1992) and subsequent assessments.
- 'Ageing data need to be integrated for correlating observed trends in research vessel survey indices' – Age data were included in the first analytical assessment (NEFC 1992) and subsequent assessments.
- 'Quantitative estimates of Z are generally lacking' – Mortality was estimated by the first analytical assessment (NEFC 1992) and subsequent assessments.

- 'Further analysis of catch by market category will be undertaken' – Catch by market category was included in the first analytical assessment (NEFC 1992) and subsequent assessments.

14th Northeast Stock Assessment Workshop (NEFSC 1992)

- 'Incorporate large mesh sea sampling discard data into future discard estimates' – Observer data was used to estimate large mesh trawl discards in the 1999 stock assessment (NEFSC 1999b) and subsequent assessments.
- 'Examine all sources of sea sampling data prior to 1989' – All sources of observer data were included in the 2012 stock assessment (NEFSC 2012) and subsequent assessments.
- 'Apportion catch by the proportion landed and proportion discarded in catch projections; simulate effect of removal of shrimp fishery on stock status and biological reference points' – Catch projections by fleet were provided by the 1999 stock assessment (NEFSC 1999b)
- 'Include Massachusetts state survey as abundance indices in the VPA' – The Massachusetts survey was included in the 1999 stock assessment (NEFSC 1999b) and subsequent assessments until 2019, when it was removed from the VPA (NEFSC 2019)
- 'Continue aging of commercial samples to improve representativeness of age-length keys' – Age length keys were updated by the 1999 stock assessment (NEFSC 1999b) and subsequent assessments.

28th Northeast Stock Assessment Workshop (NEFSC 1999b)

- 'The sea sample data used to estimate discards in the shrimp fishery could be further stratified to take account of variations in discard rates by depth' – Discard estimates were re-stratified by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).
- 'Use of another effort measure of effort such as days fished should be evaluated as an effort multiplied in the survey-based method for calculating discards in the shrimp fishery' – The discard rate and expansion to total discards were revised by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).
- 'Examine the feasibility of including Massachusetts sea sampling data and VTR data in the calculation of discards' – The discard estimation method was replaced by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).

- 'Examine the USSR data to determine if catches of American plaice may have been underestimated during the late 1960s.' – USSR catch data was included in extended series WHAM and SS runs, but the quality of the data was not verified.
- 'Examine the available data to characterize the seasonality and spatial variability of spawning in the Gulf of Maine.' – No new information was available to characterize spawning. Spring surveys by NEFSC and states have limited temporal coverage to evaluate seasonality.
- 'Derive estimates of discards for the small-mesh otter trawl component, particularly for the years 1980,1981 and 1983.' – Discards from the small mesh trawl fleet were estimated by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005), the 2015 assessment update that included small-mesh otter trawl discards, and subsequent assessments.

32nd Northeast Stock Assessment Workshop (NEFSC 2001)

- 'Statistically test maturity ogives for differences before pooling or separating ogives' – Statistical tests were used to derive time-varying maturity by the 2012 stock assessment (NEFSC 2012), and those analyses were updated in this research track assessment. However, results supported a single maturity ogive.
- 'Investigate the most appropriate choice of maximum age in the VPA and method for estimating F on the oldest age.' – Alternatives for maximum age were evaluated in the 2008 stock assessment and was revised from age-9+ to age-11+ to reduce the retrospective pattern (NEFSC 2008).
- 'Given the importance of discards in the stock, an appropriate at sea monitoring program needs to be developed and maintained' – The sampling design of the observer program was revised by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).
- 'Investigate using the shrimp and Massachusetts inshore survey in the indirect method for estimating discards' – Mayo and Terceiro (2005) concluded that survey-based methods of estimating discards (i.e., the 'indirect method') were more uncertain than those derived from observers, and the indirect method was replaced by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).
- 'Re-examine the indirect methods and other methods for estimating discards' – Mayo and Terceiro (2005) concluded that survey-based methods of estimating discards were more

uncertain than those derived from observers, and the indirect method was replaced by the Northeast Standardized Bycatch Reporting Method (Rago et al. 2005).

- ‘Investigate using statistical catch at age models to account for ageing errors in the catch at age. This recommendation applies to all the analytical assessments reviewed by the SARC and should be taken as a general recommendation’ – Many northeast US stock assessments transitioned from VPA to the ASAP statistical catch at age, and this research track assessment proposes replacing the VPA with an integrated assessment method’
- ‘Age archived samples from the Massachusetts inshore survey’ – Archived samples from the MADMF survey have not been processed.
- ‘Examine trends of survey indices by geographic area in order to evaluate the appropriateness of pooling biological parameters by area’ – This research track assessment investigated regional variation in growth and maturity.

Groundfish Assessment Review Meeting (GARM) (NEFSC 2002a)

- ‘The GARM panel recommended that sensitivity analyses be conducted to evaluate effects of uncertainty in discard estimates on assessment results’ – Although historical discards were uncertain, estimates from recent observer estimates with higher coverage have been relatively precise (e.g., CV<=0.15 for large-mesh discards, the primary source of recent discards).
- ‘The survey time series could be split into two tuning indices based on time periods corresponding to changes in methods for estimating discards’ – The proposed assessment models split survey indices in 2008-2009 to account for the change in NEFSC survey systems, but that may also address this recommendation.

Re-Evaluation of Biological Reference Points for New England Groundfish (NEFSC 2002b)

- ‘The current VPA time series of stock recruit data was considered insufficient to apply to any parametric stock-recruit model.’ – This research track assessment explored assessment models with extended time series (1960-2019) but the stock-recruit relationship was not well defined, and results suggest that the early 1980s represent the historical period of productivity.

2015 Update Assessment (NEFSC 2015)

- 'For the next benchmark assessment, the Panel recommended that a statistical catch-at-age model, which can potentially handle the observed conflict between offshore and inshore surveys, should be explored.' – Statistical catch at age models were developed for this research track assessment.
- 'In addition, the assessment team should consider the inclusion of the Maine-New Hampshire survey as another calibration index.' – The Maine-New Hampshire survey was considered in this research track assessment.

2017 Update Assessment (NEFSC 2017)

- 'The Gulf of Maine-Georges Bank American plaice assessment could be improved with updated studies on growth of Georges Bank and Gulf of Maine fish. A difference in growth rates between Gulf of Maine and Georges Bank fish has been documented; however, historical catch data for Georges Bank may not be sufficient to conduct a separate assessment. The panel recommends continuation of research on growth rates and implications for stock structure. The growth rate difference actually may not persist in the most recent years so this could all be explored further in a benchmark review.' – Regional growth differences were investigated in this research track assessment.
- 'Finally, the panel recommends further research and consideration of survey catchability estimates.' – Estimates of survey efficiency and catchability were considered in the 2017 and 2019 update assessments (NEFSC 2019) as well as this research track assessment.

Scientific & Statistical Committee Discussion from 2017 Update Assessment

- 'There was a discussion on improving the analysis for plaice in the future given new age information on the stock.' – All age information available from NEFSC surveys were included in this research track, but age data from state surveys was not processed.

- 'The SSC noted some data conflicts in the information namely that the trends were different inshore and offshore, however, given that stock status appeared to be good, the SSC was comfortable with the 75%FMSY projection recommendation.' – The conflict between NEFSC and state surveys was resolved in the 2019 assessment update by excludinig state surveys, justified by the shift in plaice distribution to deeper habitats, and after explorations of assessment models including state surveys, that decision was maintained in this research track assessment.

2019 Update Assessment (NEFSC 2019)

- 'The Panel recommend the development of a statistical catch at age or state-space model for this stock in the 2021 research track assessment, which would make it easier to split the Bigelow and Albatross time series into two separate indices. Perhaps it would be useful in a research track to examine how the information on the younger fish appearing in the MADMF survey data might be used given the concern with movement of the stock offshore.' – This research track assessment developed statistical catch at age assessments, state-space models, decided to split the NEFSC Albatross and Bigelow time series, and explored the effect of depth on surveys.
- 'Consideration of regionally-stratified catch at age estimation for Gulf of Maine and Georges Bank could be considered in the next assessment to account for potential growth differences.' – This research track assessment considered regional variation in growth.

Scientific & Statistical Committee Discussion from 2019 Update Assessment

- 'There was considerable discussion about whether to use the projected (and declining) 75%Fmsy values for the ABCs. This is what the SSC elected to do in 2017 for this stock. The SSC noted that using the projected values assumes a constant level of uncertainty and thus, a high degree of confidence in the projections. Furthermore, to use the projections in this way, the SSC suggested that the rho adjustment should be applied in each projected year, rather than just the terminal year. This would lead to more aggressive declines than the projections alone indicate. The SSC decided to use the minimum (year 3, 2022) 75% Fmsy value and hold that value constant for three years. This provides an additional uncertainty buffer to increase the chances that overfishing does not occur on this stock.' – The retrospective pattern in previous VPA assessments does not apply to the WHAM assessment method proposed by the research track Working Group, but 75%$F_{40\%}$ scenarios were included in projections to support the current approach to determining ABC.

- 'As noted, there is a strong retrospective bias in the assessment for this stock. This was an area of focus in the previous advice of the SSC and continues to be an uncertainty that concerns the SSC. Given the improved performance of the projections, the SSC felt that the retrospective adjustment made prior to implementing the projection methods adequately mitigates this uncertainty; therefore, additional buffering is not needed'– The retrospective pattern in previous VPA assessments does not apply to the WHAM assessment method proposed by the research track Working Group, but 75%$F_{40\%}$ scenarios were included in projections to support the current approach to determining ABC as well as alternative approaches to ABC based on uncertainty in projected catch at $F_{40\%}$.

New Research Recommendations

1. Investments are needed to streamline the estimation of commercial catch and promote reproducibility of estimates.
2. Future assessments should consider deriving discards from electronic monitoring for vessels in those programs when an integrated catch monitoring system is developed, and information on length distribution from electronic monitoring should be considered for estimating composition of discards.
3. As the Gulf of Maine scallop fishery expands, it should be included in discard estimation.
4. As recommended by NEFSC (2011), age samples from state surveys could help to use those surveys in the assessment. Although the exclusion of state surveys was primarily based on the shift in plaice distribution to deeper habitats, using offshore age-length keys may have contributed to the poor model fit to state surveys, and age composition from state surveys could improve integrated analyses of inshore and offshore surveys.
5. Exploration of spatiotemporal integration of federal and state surveys should continue.
6. Continue to monitor shifts in distributions of plaice, particularly the noted increase depth associated with increasing temperature, to evaluate whether there are improvements to model performance by including an environmental covariate on catchability.
7. The relationship between recruitment and ocean temperature should continue to be monitored in future analyses and considered in the context of model fitting (i.e., including environmental covariates) and recruitment assumptions for reference points and projections.

8. To ensure that future model explorations with environmental covariates are comparable via AIC, problems with calculation and comparability of AIC with and without environmental covariate linkage should be resolved.

9. If the proposed assessment approach does not meet the standards of peer review or is rejected in a future management track assessment, an alternative model be developed to integrate information from catch, age composition and indices.

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*."

The Northeast US stock assessment process requires an accepted stock assessment method to provide best scientific information available for fishery management, including a contingency plan if the proposed assessment method fails peer review in the research track process or subsequently fails peer review in the routine management track process. Many northeast US assessments specify an empirical backup approach based on survey data, either swept-area estimates of stock biomass and a target exploitation rate or survey biomass trends and recent catch.

The Index-Based Research Track Working Group simulation-tested the performance of several empirical assessment methods (NEFSC 2020a). They found that empirical approaches had similar performance within two groups. One group of empirical approaches tended to produce low long- term F/F_{MSY} and high SSB/SSB_{MSY}, but low short-term catch/MSY. The other group of approaches had a more linear tradeoff between long-term SSB/SSB_{MSY} and catch/MSY. Based on their performance results. They recommended that the first group of approaches (catch curve and $F_{40\%}$, catch curve and F=M, dynamic linear model, survey trends 'PlanB Smooth', expanded survey biomass and recent F, common trends) would be most appropriate for stocks that are at or above B_{MSY}, but the catch performance was relatively poor. they concluded that "*For stocks that have had an age-based assessment rejected due to a strong retrospective pattern, there is no expectation that an index-based assessment will perform better than a rho-adjusted statistical catch at age analysis*" (NEFSC 2020a).

Efficiency-adjusted swept-area biomass estimates are available for plaice in US waters (Figure 3.5) and could be used as a contingency plan for monitoring the stock. However, age composition of the fishery (Figure 2.12) and NEFSC surveys (Figure 3.8) suggest that older plaice are not fully selected by the surveys. Model estimates also indicate domed selectivity of NEFSC surveys (e.g., Figures 4.8 and 4.24). Therefore, catch biomass from the fishery is not directly comparable to survey estimates of biomass using a simple catch/biomass exploitation ratio. Biomass reference points derived from area-swept approaches and dynamic pool models require selectivity assumptions, and comparisons of area-swept biomass to the biomass reference point implicitly assume equal survey catchability of recruits and spawners. Unfortunately, the estimate of survey efficiency is uncertain for small plaice (Figure 3.4). This problem precluded the method being used for red hake in a recent research track peer review (NEFSC 2020b). Another challenge with area-swept approaches is that the survey area is not entirely consistent with the stock area (e.g., surveys include Canadian waters of eastern Georges Bank but exclude some shallow habitats).

Catch curves were also considered as a contingency approach to stock assessment. The assumption of constant mortality is consistent with the perception of flat-topped fishery selectivity, and fishery catch-at-age was reasonably log-linear for ages older than the age at peak catch (Figure 8.1). However, estimates of total mortality rate (Z) were less than the assumed M for recent year-classes, producing negative estimates of fishing mortality. Survey catch curves would violate the assumption of constant mortality, because older plaice do not appear to be fully selected by the surveys.

Data-limited approaches based on surveys or catch curves would not include all the available information, including decades of information available for fishery landings and discards, multiple survey indices, and age composition. Procedurally, once a model-based assessment is replaced by an empirical approach, the model-based assessment cannot be re-considered in the management track process and requires another research track assessment. The most recent benchmark assessment for plaice in US waters was in 2008 (NEFSC 2008), and another research track assessment is not expected in the next decade. Therefore, based on results of simulation testing and recommendations by the Index-Based Methods Working Group (NEFSC 2020a), the apparent domed selectivity of surveys for plaice (Figures 2.12, 3.8, and 4.8), and challenges deriving reference points for survey biomass (NEFSC 2020b), the Working Group recommends that if the proposed assessment approach (WHAM run 29F4) does not meet the standards of peer review or is rejected in a future management track assessment, an alternative model be developed to integrate information from catch, age composition and indices (e.g., alternative WHAM configurations or ASAP run 43).

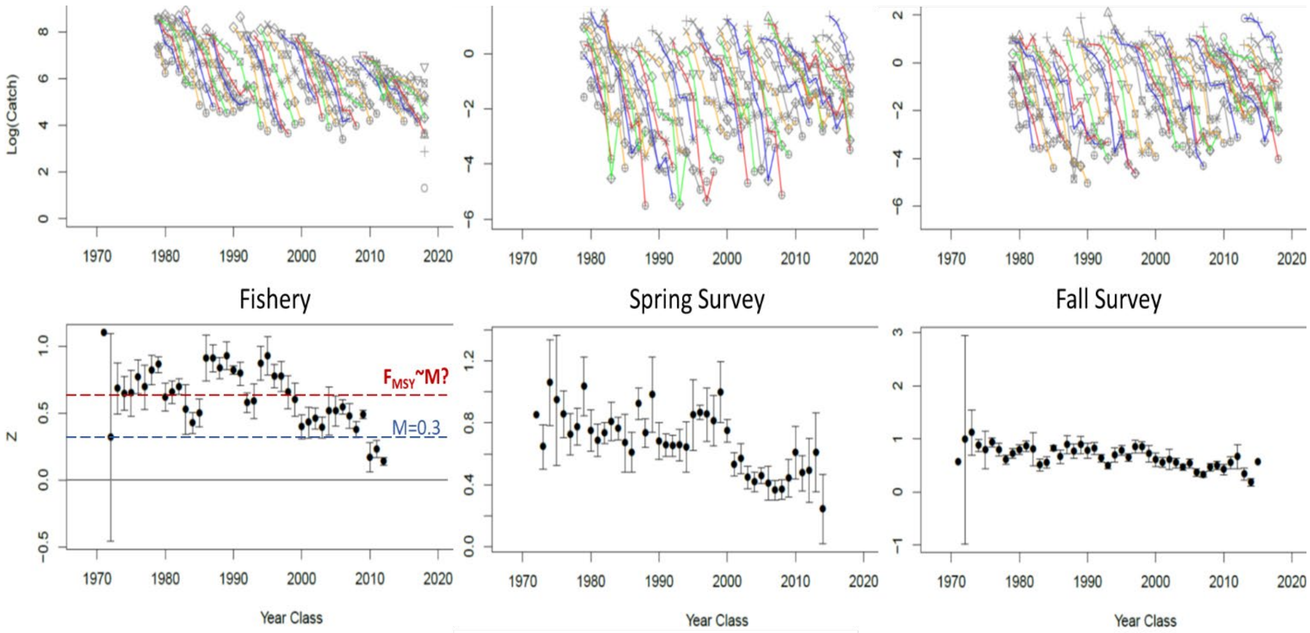


Figure 8.1 Catch Curves for American plaice from fishery and survey catch at age, with references for the assumed rate of natural mortality and the proxy F_{MSY}=M.

ACKNOWLEDGMENTS

This report reflects approximately a year of work from a large group of people, including Working Group members and participants. The Working Group specifically thanks Amanda Hart for her contributions to developing WHAM applications for plaice. We also thank previous assessment leads (Loretta O'Brien and Mark Terceiro), previous age readers (Louise Dery, Christine Esteves), survey participants, observers, port samplers and fishermen who contributed to the data used in this assessment. Cate O'Keefe, Jackie O'Del, Vito Giacalone, and Maggie Raymond helped to organize meetings with fishermen in Gloucester and Portland. Fishermen participated in workshops and individual meetings. Greg DeCelles and John Manderson contributed to the Fishermen's Ecological Knowledge study. Keith Hankowsky, Max Grezlik, Cole Carrano, Lucy McGinnis, and Jamie Behan served as meeting rapporteurs. Michele Traver and Alexander Dunn provided administrative support. UMES students were supported by the NOAA Office of Education, Educational Partnership Program with Minority Serving Institutions, Experiential Research and Training Opportunity. Research was partially funded by the NOAA Cooperative Institute for the North Atlantic Region.

REFERENCES

Adams CF, Alade LA, Legault CM, O'Brien L, Palmer MC, Sosebee KA, Traver ML. 2018. Relative importance of population size, fishing pressure and temperature on the spatial distribution of nine Northwest Atlantic groundfish stocks. PLoS ONE 13(4): e0196583.

Alexander AB, Moore HF, Kendall WC. 1915. Otter-trawl fishery. US Bureau of Fisheries Document No. 816.

Alexander M, Shin S, Scott J, Curchitser E, Stock C. 2019. The response of the Northwest Atlantic Ocean to climate change. Journal of Climate 33: 405-428.

Anderson SC, Keppel EA, Edwards AM. 2019. A reproducible data synopsis for over 100 species of British Columbia groundfish. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/041.

Applegate A, Cadrin S, Hoenig J, Moore C, Murawski S, Pikitch E. 1998. Evaluation of existing overfishing definitions and recommendations for new overfishing definitions to comply with the Sustainable Fisheries Act. New England Fishery Management Council Report (s3.amazonaws.com/nefmc.org/Evaluation-of-Existing-Overfishing-Definitions.pdf).

Benoît HP, Hurlbut T, Chassé J, Jonsen ID. 2012. Estimating fishery-scale rates of discard mortality using conditional reasoning. Fisheries Research 125–126: 318– 330.

Bigelow HB, Schroeder WC. 1953. Fishes of the Gulf of Maine. Fisheries Bulletin 54: 1-577.

Bowen WD, editor. 1987. A review of stock structure in the Gulf of Maine area: a workshop report. CAFSAC Res. Doc. 87/21.

Brickman D, Alexander MA, Pershing A, Scott JD, Wang Z. 2021. Projections of physical conditions in the Gulf of Maine in 2050. Elementa: Science of the Anthropocene 9.1 DOI: https://doi.org/10.1525/elementa.2020.20.00055

Brodziak J, O'Brien L. 2005. Do environmental factors affect recruits per spawner anomalies of New England groundfish? ICES Journal of Marine Science 62: 1394-1407.

Brooks DA. 1985. Vernal circulation in the Gulf of Maine. Journal of Geophysical Research: Oceans 90(C3): 4687–4706.

Brown SK, Mahon R, O'Boyle R, Zwanenburg K, Buja K, Claflin L, Atkinson B, Howell G, Sinclair M, Monaco M. 1996. East coast of North America groundfish: initial explorations of biogeography and species assemblages. Silver Spring, MD: Strategic Environmental Assessments Division, National Oceanic and Atmospheric Administration, and Dartmouth, NS: Marine Fish Division, Department of Fisheries and Oceans. 102 pp.

Burnett J, O'Brien L, Mayo RK, Darde J, Bohan M. 1989. Finfish maturity sampling and classification schemes used during Northeast Fisheries Center bottom trawl surveys, 1963-1989. NOAA Tech. Mem. NMFS-F/NEC-76.

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 (https://s3.amazonaws.com/nefmc.org/200220_Fishery-Data-Working-Group-report_Final_with-appendices.pdf).

Cadrin SX., Hansell AC, Wright BL. 2020. Cape Cod Commercial Fishermen’s Alliance Groundfish Electronic Monitoring Data Analysis. Final Report. 23 p.

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.

Carvalho F, Winker H, Courtney D, Kapur M, Kell L, Cardinale M, Schirripa M, Kitakado T, Yemane D, Piner KR, Maunder MN, Taylor I, Wetzel CR, Doering K, Johnson KF, Methot RD. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research 240: 105959.

Chang J-H, Chen Y, Halteman W, Wilson C. 2016. Roles of spatial scale in quantifying stock–recruitment relationships for American lobsters in the inshore Gulf of Maine. Canadian Journal of Fisheries and Aquatic Sciences 73: 885–909.

Charnov EL, Gislason H, and Pope JG., 2013. Evolutionary assembly rules for fish life histories. Fish and Fisheries 14: 213-224.

Chen S, and Watanabe S. 1989. Age dependence of natural mortality coefficient in fish population dynamics. Nippon Suisan Gakkaishi 55: 205-208.

Chouinard P-M, Dutil J-D, Proulx S, Borcard D, Crocker. 2014. A spatial approach to matching marine fish diversity and abundance with habitat features. Hydrobiologia 734: 39–55.

Clark W. 1993. The effect of recruitment variability on the choice of a target level of spawning biomass per recruit. University of Alaska Sea Grant College Program, Report Number 93-02:233-246.

Colton JB Jr., Temple RF. 1961. The enigma of Georges Bank spawning. Limnology and Oceanography 6: 280-291.

Cope JM. 2021. Upgrading from M 0.2: an application-based method for accessible estimation, evaluation and uncertainty characterization of natural mortality. Fisheries Research (in press).

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 2009. COSEWIC assessment and status report on the American plaice, *Hippoglossoides platessoides*: Maritime population, Newfoundland and Labrador population, Arctic population, in Canada www.sararegistry.gc.ca/status/status_e.cfm.

Demarest C. 2019. Evaluating the observer effect for the Northeast US groundfish fishery. In Amendment 23 to the Northeast Multispecies Fishery Management Plan Including a Draft Environmental Impact Statement (DEIS) Volume II - Appendices (https://s3.amazonaws.com/nefmc.org/200221-Draft-Amendment-23-Volume-II-Appendices.pdf).

Dery LM. 1988. American plaice *Hippoglossoides platessoides*. pp. 111-118 In Age Determination Methods for Northwest Atlantic Species. NOAA Technical Report NMFS 72.

Dichmont CM, Deng R, Punt AE, Brodziak J, Chang Y.-J, Cope JM, Ianelli JN, Legault CM, Methot RD, Porch CE, Prager MH, Shertzer K. 2016. A review of stock assessment packages in the United States. Fisheries Research 183: 477–460.

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

Esteves C, Burnett J. 1993. A comparison of growth rates for American plaice, *Hippoglossoides platessoides*, in the Gulf of Maine-Georges Bank region derived from two different data sources. NEFSC Ref. Doc. 93-09.

Fisch N, Camp E, Shertzer K, Ahrens R. 2021. Assessing likelihoods for fitting composition data within stock assessments, with emphasis on different degrees of process and observation error. Fish. Res. 243:106069.

Fitz-Gerald C, Rossi N, Cronin H, Sanderson M. 2019. Electronic Monitoring in the Groundfish Fishery: A summary of current programs for the Groundfish Advisory Panel and Groundfish Committee. NOAA and New England Fishery Management Council. https://s3.amazonaws.com/nefmc.org/3c_GF-EM-in-the-northeast-for-AP-and-Cmte-8_6.pdf.

Francis, RICC. 2011. Data Weighting in Statistical Fisheries Stock Assessment Models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124–38.

Fredston-Hermann A, Selden R, Pinsky M, Gaines SD, Halpern BS. 2020. Cold range edges of marine fishes track climate change better than warm edges. Global Change Biology 26: 2908–2922.

Fuiman LA, Batty RS. 1994. Susceptibility of Atlantic herring and plaice larvae to predation by juvenile cod and herring at two constant temperatures. Journal of Fish Biology 44: 23–34.

Gavaris S. 1988. An adaptive framework for the estimation of population size. CAFSAC Res. Doc 88/29: 12 p.

Gislason H, Daan N, Rice JC, and Pope JG. 2010. Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries 11: 149-158.

Greene, CH, Pershing, AJ. 2003. The flip-side of the North Atlantic Oscillation and modal shifts in slope-water circulation patterns. Limnology and Oceanography 48: 319–322.

Haltuch MA, Punt AE, Dorn MW. 2008. Evaluating alternative estimators of fishery management reference points. Fisheries Research 94: 290–303.

Hamel OS., 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science 72: 62-69.

Han D, Tian S, Zhang Y, Chen Y. 2021. An evaluation of temporal changes in the trophic structure of the Gulf of Maine ecosystem. Regional Studies in Marine Science 42: 101635.

Hare JA, Morrison WE, Nelson MW, Stachura MM, Teeters EJ, Griffis RB, Alexander MA, Scott JD, Alade L, Bell RJ, Chute AS. 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.

Harris LG, Tyrrell MC. 2000. Changing community states in the Gulf of Maine: Synergism between invaders, overfishing and climate change. Biological Invasions 3: 9–21.

Henry A, Demarest C, Errend M. 2019. Modelling discard incentives for northeast multispecies (groundfish) stocks. In Amendment 23 to the Northeast Multispecies Fishery Management Plan Including a Draft Environmental Impact Statement (DEIS) Volume II - Appendices (https://s3.amazonaws.com/nefmc.org/200221-Draft-Amendment-23-Volume-II-Appendices.pdf).

Howell WH, Caldwell MA. 1984. Influence of temperature on energy utilization and growth of embryonic and prolarval American plaice, *Hippoglossoides platessoides* (Fabricius). J. Exp. Mar. Biol. Ecol. 19: 173-189.

Huntsman AG. 1918. Histories of new food fishes. I. The Canadian plaice. Bull. Biol. Board Can. 1. 32p.

ICCAT (International Commission for the Conservation of Atlantic Tunas). 2021. Report of the 2021 Western Bluefin Stock Assessment Meeting (online, 30 August-1 September 2021). https://www.iccat.int/Documents/Meetings/Docs/2021/REPORTS/2021_WBFT_SA_ENG.pdf

Iles TC. 1994. A review of stock-recruitment relationships with reference to flatfish populations. Netherlands Journal of Sea Research 32: 399–420.

Jensen AL. 1997. Origin of the relation between K and L_{inf} and synthesis of relations among life history parameters. Can. J. Fish. Aquat. Sci. 54: 987-989.

Jensen AC. 1967. A brief history of the New England offshore fisheries. US Bureau of Commercial Fisheries, Fishery Leaflet 594.

Johnson DL, Berrien PL, Morse WW, Vitaliano JJ. 1999. American Plaice, *Hippoglossoides platessoides*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-123. 31 pp.

Johnson DL. 2004. Essential fish habitat source document. American plaice, *Hippoglossoides platessoides*, life history and habitat characteristics, Second Edition. NOAA Technical Memorandum NMFS-NE-187.

Jørgensen C, Holt RE. 2013. Natural mortality: Its ecology, how it shapes fish life histories, and why it may be increased by fishing. Journal of Sea Research 75: 8–18.

Kell LT, Sharma R, Kitakado T, Winker H, Mosqueira I, Cardinale M, Fu D. 2021. Validation of stock assessment methods: is it me or my model talking?. ICES J. Mar. Sci. 78(6):2244-2255.

Kenchington TJ. 2013. Natural mortality estimators for information-limited fisheries. Fish and Fisheries 15: 533-562.

King JR, Camisa MJ, Manfredi VM. 2010. Massachusetts Division of Marine Fisheries Trawl Survey Effort, Lists of Species Recorded, and Bottom Temperature Trends, 1978- 2007. Massachusetts Division of Marine Fisheries Technical Report TR-38. 157 p.

Kleisner KM, Fogarty MJ, McGee S, Hare JA, Moret S, Perretti CT, Saba VS. 2017. Marine species distribution shifts on the US Northeast Continental Shelf under continued ocean warming. Progress in Oceanography 153: 24-36.

Kulka DW. 2012. History and Description of the International Commission for the Northwest Atlantic Fisheries. (https://www.nafo.int/Portals/0/PDFs/icnaf/ICNAF_history-kulka.pdf)

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

Lange AMT, Lux FE. 1978. Review of the other flounder stocks (winter flounder, American plaice, witch flounder, and windowpane flounder) off the northeast United States, August 1978. Northeast Fisheries Center, Woods Hole Lab. Ref. Doc. 78-44.

Lesser, M. P. 2016. Climate change stressors cause metabolic depression in the blue mussel, *Mytilus edulis*, from the Gulf of Maine. Limnology and Oceanography 61: 1705–1717.

Levangie PEL, Blanchfield PJ, Hutchings JA. 2021. The influence of ocean warming on the natural mortality of marine fishes. Environmental Biology of Fishes https://doi.org/10.1007/s10641-021-01161-0.

Lux FE. 1969. Length-weight relationships of six New England flatfishes. Transactions of the American Fisheries Society 4: 617-621.

Lux FE. 1970. Note on growth of American plaice, *Hippoglossoides platessoides* (Fabr.), in ICNAF Subarea 5. ICNAF' Research Bulletin 7: 5-7.

Lynch PD, Methot RD, Link JS (editors). 2018. Implementing a Next Generation Stock Assessment Enterprise. An Update to the NOAA Fisheries Stock Assessment Improvement Plan. NOAA Tech. Memo. NMFS-F/SPO-183.

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

Mayo RK, Terceiro M., editors. 2005. Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, 15-19 August 2005. NEFSC Ref. Doc. 05-13.

McBride, RS, Alade LA, Ricard D, Morgan J. 2018. Sexual Size Dimorphism in United States and Canadian Populations of American plaice *Hippoglossoides platessoides*. pp. 31 in Mercaldo-Allen R, Chambers C, Davis D, Dixon M, Dwyer S, Fairchild E, Howell P, Munroe T, Powell C, Sutherland S. 2018. 16th Flatfish Biology Conference 2018 Program and Abstracts. NEFSC Ref. Doc. 18-14.

McElderry H. 2006. At-Sea Observing Using Video-Based Electronic Monitoring. ICES CM 2006/N:14.

Melnychuk, MC, Kurota H, Mace PM, Pons M, Minto C, Osio GC, Jensen OP, de Moor CL, Parma AM, Little LR, Hively D, Ashbrook CE, Baker N, Amoroso RO, Branch TA, Anderson CM, Szuwalski CS, Baum JK, McClanahan TR, Ye Y, Ligas A, Bensbai J, Thompson GG, DeVore J, Magnusson A, Bogstad B, Wort E, Rice J, Hilborn R. 2021. Identifying management actions that promote sustainable fisheries. Nature Sustainability 4: 440–449.

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: 1744–60.

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

Methratta ET, Link JS. 2007. Ontogenetic variation in habitat associations for four flatfish species in the Gulf of Maine-Georges Bank region. Journal of Fish Biology 70: 1669–1688.

Miller TJ. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for US Northwest Atlantic fish stocks. Can. J. Fish. Aquat. Sci. 70: 1306–1316.

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 Fish Sci Cent Ref Doc. 10-05.

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56: 473-488.

Murphy T, Ardini G, Vasta M, Kitts A, Demarest C, Walden J, Caless D. 2018. 2015 Final Report on the Performance of the Northeast Multispecies (Groundfish) Fishery (May 2007 – April 2016). NEFSC Ref. Doc. 18-13.

NEFC (Northeast Fisheries Center). 1985. Rapporteur's Summary, First Northeast Fisheries Center Stock Assessment Workshop 8-12 July, 1985, Falmouth, Massachusetts (https://apps-nefsc.fisheries.noaa.gov/rcb/publications/series/whlrd/1saw1985.pdf)

NEFC (Northeast Fisheries Center). 1986. Report of the Second NEFC Stock Assessment Workshop (Second SAW). Northeast Fisheries Center, Woods Hole Lab. Ref. Doc. 86-09.

NEFSC (Northeast Fisheries Science Center). 1992. Report of the Fourteenth Northeast Regional Stock Assessment Workshop (14th SAW). NEFSC Ref. Doc. 92-07.

NEFSC (Northeast Fisheries Science Center). 1996. Report of the 22nd Northeast Regional Stock Assessment Workshop (22nd SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref. Doc. 96-13.

NEFSC (Northeast Fisheries Science Center). 1999a. Report of the 28th Northeast Regional Stock Assessment Workshop (28th SAW): Public Review Workshop. NEFSC Ref. Doc. 99-07.

NEFSC (Northeast Fisheries Science Center). 1999b. Report of the 28th Northeast Regional Stock Assessment Workshop (28th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref. Doc. 99-08.

NEFSC (Northeast Fisheries Science Center). 2001a. Report of the 32nd Northeast Regional Stock Assessment Workshop (32nd SAW): Public Review Workshop. NEFSC Ref. Doc. 01-04.

NEFSC (Northeast Fisheries Science Center). 2001b. Report of the 32nd Northeast Regional Stock Assessment Workshop (32nd SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref. Doc. 01-05

NEFSC (Northeast Fisheries Science Center). 2002a. Assessment of 20 Northeast groundfish stocks through 2001: a report of the Groundfish Assessment Review Meeting (GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, October 8-11, 2002. NEFSC Ref. Doc. 02-16.

NEFSC (Northeast Fisheries Science Center). 2002b. Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. NEFSC Ref. Doc. 02-04.

NEFSC (Northeast Fisheries Science Center). 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. NEFSC Ref. Doc. 08-15.

NEFSC (Northeast Fisheries Science Center). 2012a. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. NEFSC Ref. Doc. 12-18.

NEFSC (Northeast Fisheries Science Center). 2012b. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. NEFSC Ref. Doc.12-06.

NEFSC (Northeast Fisheries Science Center). 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. NEFSC Ref. Doc. 15-24.

NEFSC (Northeast Fisheries Science Center). 2017a. 62nd Northeast Regional Stock Assessment Workshop (62nd SAW) assessment summary report. NEFSC Ref. Doc. 17-01.

NEFSC (Northeast Fisheries Science Center). 2017b. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. NEFSC Ref. Doc. 17-17.

NEFSC (Northeast Fisheries Science Center). 2021. State of the Ecosystem 2021: New England Revised. https://repository.library.noaa.gov/view/noaa/29524.

NEFSC (Northeast Fisheries Science Center). 2020a. Draft Report of the Index Based Methods Working Group. NEFSC Ref. Doc. (in prep). https://static1.squarespace.com/static/511cdc7fe4b00307a2628ac6/t/60391d7c0a8d8429d50f72b2/1614355837553/Draft+Index+Based+Methods+Peer+Review+Report.pdf.

NEFSC (Northeast Fisheries Science Center). 2020b. Final report of red hake stock structure working group. NEFSC Ref. Doc. 20-07; 185 p.

NEFSC (Northeast Fisheries Science Center). 2022a. Butterfish (*Peprilus triacanthus*) Research Track Assessment Report. NEFSC Ref. Doc. 22-xx (https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php)

NEFSC (Northeast Fisheries Science Center). 2022b. Stock Assessment Update of 14 Northeast Groundfish Stocks Through 2018. NEFSC Ref. Doc. 22-06.

NEFMC (New England Fishery Management Council). 1985. Fishery Management Plan, Environmental Impact Statements, Regulatory Impact Review, and Initial Regulatory Flexibility Analysis for the Northeast Multi-Species Fishery. (https://s3.amazonaws.com/nefmc.org/MultiSpecies-FMP.pdf)

NEFMC (New England Fishery Management Council).2009. Amendment 16 to the Northeast Multispecies Fishery Management Plan. (https://s3.amazonaws.com/nefmc.org/091016_Final_Amendment_16.pdf)

NEFMC (New England Fishery Management Council). 2016. American plaice. Pp 86-90 In. Omnibus Essential Fish Habitat Amendment 2, Volume 2. (https://s3.amazonaws.com/nefmc.org/OA2-FEIS_Vol_2_FINAL_171025.pdf)

NEFMC (New England Fishery Management Council). 2021. Northeast Multispecies Fishery Management Plan Amendment 23. (https://s3.amazonaws.com/nefmc.org/210430_Groundfish_A23_FEIS_preliminary_submission.pdf)

Neilson JD, Hurley P. 1986. Stock structure of American plaice, witch and winter flounder in the Gulf of Maine area: implications for management. CAFSAC Res. Doc. 86/63.

Nielsen A, Berg CW. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158, 96–101.

NMFS (National Marine Fisheries Service). 2021. Fisheries of the United States 2019. Current Fishery Statistics No. 2019 (https://media.fisheries.noaa.gov/2021-05/FUS2019-FINAL-webready-2.3.pdf?null=)

NOAA (National Oceanic and Atmospheric Administration). 2010. Magnuson-Stevens Fishery Conservation and Management Act Provisions; Fisheries of the Northeastern United States; Northeast (NE) Multispecies Fishery; Amendment 16. 15 CFR Part 902 50 CFR Part 648 [Docket No. 0808071078–0019–02] RIN 0648–AW72.

NOAA (National Oceanic and Atmospheric Administration). 2013. Magnuson-Stevens Act Provisions; National Standard 2—Scientific Information. 50 CFR Part 600 [Docket No. 0808041047–3587–03] RIN 0648–AW62

NOAA (National Oceanic and Atmospheric Administration). 2016. Magnuson-Stevens Act Provisions; National Standard Guidelines. 50 CFR Part 600 [Docket No. 120416013-6270-03] RIN 0648-BB92.

NOAA (National Oceanic and Atmospheric Administration). 2020. The National Electronic Monitoring Workshop Report 2019/2020. Office of Science and Technology. https://s3.amazonaws.com/media.fisheries.noaa.gov/2020-09/2020-EM-National-Workshop-Report-FINAL-4-webready.pdf?ci7Mq1XPdpkHw2yzVtxGTtWXXObKWlPrNRC (National Research Council) 1998. Review of Northeast Fishery Stock Assessments. National Academies Press. https://doi.org/10.17226/6067.

Nye J. 2010. Climate Change and Its Effects on Ecosystems, Habitats and Biota: State of the Gulf of Maine report. Gulf of Maine Council on the Marine Environment.

http://www.gulfofmaine.org/state-of-the-gulf/docs/climate-change-and-its-effects-on-ecosystems-habitats-and-biota.pdf.

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'Boyle RH, Sinclair M, Conover RJ, Mann KH, Kohler AC. 1982. Temporal and spatial distribution of ichthyoplankton communities of the Scotian Shelf in relation to biological, hydrological and physiological features. Rapp. P. V. Reun. Cons. int. Explor. Mer 183:27-40.

O'Brien L, Burnett J, Mayo RK. 1993. Maturation of nineteen species of finfish off the northeast coast of the United States, 1985-1990. NOAA Tech. Rep. NMFS 113. 66 p.

O'Brien L, Esteves C. 2001. Update assessment of American plaice in the Gulf of Maine - Georges Bank region for2000. NEFSC Ref. Doc. 01-02.

O'Brien L, Mayo RK, Esteves C. 1999. Assessment of American Plaice in the Gulf of Maine-Van Helmond, A.T.M, Mortenson, L.O., Plet-Hansen, K.S., Ulrich, C., Needle, C.L., Oesterwind, D., Kindt-Larsen, L., Catchpole, T., Mangi, S., Zimmerman, C., Olesen, H.J., Bailey, N., Bergsson, H., Dalskov, J., Elson, J., Hosken, M., Peterson, L., McElderry, H., Ruiz, J., Pierre, J.P., Dykstra C., and Poos, J.J. 2020. Electronic monitoring in fisheries: Lessons from global experiences and future opportunities. Fish and Fisheries 21: 162-189. Georges Bank Region for 1998. NEFSC Ref. Doc. 99-09.

Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421(6918): 37–42.

Pauly D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. J. Cons. Int. Explor. Mer: 175-192.

Pershing AJ, Alexander MA, Hernandez CM, Kerr LA, Le Bris A, Mills KE, Nye JA, Record NR, Scannell HA, Scott JD, Sherwood GD, Thomas AC. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science 350(6262): 809–812.

Pettigrew NR, Churchill JH, Janzen CD, Mangum LJ, Signell RP, Thomas AC, Townsend DW, Wallinga,JP, Xue H. 2005. The kinematic and hydrographic structure of the Gulf of Maine Coastal Current. Deep Sea Research Part II: Topical Studies in Oceanography, 52(19), 2369–2391.

Pitt TK. 1967. Age and growth of American plaice, (*HippoglossQides platessoides*), in the Newfoundland area of the Northwest Atlantic. J. Fish. Res. Board Can. 24:1077-1099.

Pitt TK. 1972. Estimates of natural mortality coefficients of American plaice *Hippoglossoides platessoides* in the Magdalen shallows. J. Fish. Res. Board Can. 22: 565-598.

Politis PJ, Galbraith JK, Kostovivk P, Brown RW. 2014. Northeast Fisheries Science Center bottom trawl survey protocols for NOAA Ship Henry B. Bigelow. NEFSC Ref. Doc. 14-06.

Powles PM. 1965. The life history and ecology of the American plaice, (*Hippoglossoides. platessoides*), in the Magdalen shallows. J. Fish. Res. Board Can. 22: 565-598.

Powles PM. 1966. Validity of ageing young American plaice from otoliths. Int. Comm. Northwest All. Fish. Res. Bull. 3:103-105.

Powles PM. 1969. Size changes, mortality, and equilibrium yields in an exploited stock of American plaice (*Hippoglossoides platessoides*). J. Fish. Res. Bd. Canada 26: 1205-1235.

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, Wigley SE, Fogarty MJ. 2005. NEFSC bycatch estimation methodology: allocation, precision and accuracy. NEFSC Ref. Doc. 05-09.

Rahmstorf S, Cazenave A, Church JA, Hansen JE, Keeling RF, Parker DE, Somerville RCJ. 2007. Recent Climate Observations Compared to Projections. Science 316(5825): 709–709.

Rikhter VA, and Efanov VN. 1976. On one of the approaches to estimation of natural mortality of fish populations. ICNAF Res. Doc. 79/VI/8, 12.

Robertson MD, Gao, J., Regular, P. M., Morgan, M. J., & Zhang, F. 2021. Lagged recovery of fish spatial distributions following a cold-water perturbation. Scientific Reports, 11(1), 9513.

Roff D A. 1984. The evolution of life history parameters in teleosts. Can. J. Fish. Aquat. Sci. 41: 989-1000.

Ross MR, Hokenson SR. 1997. Short-term mortality of discarded finfish bycatch in the Gulf of Maine Fishery for Northern Shrimp *Pandalus borealis*. North American Journal of Fisheries Management 17: 902-909.

Rounsefell GA. 1948. Development of Fishery Statistics in the North Atlantic. US Fish & Wildlife Service Special Scientific Report No. 47.

Rounsefell GA. 1957. A method of estimating abundance of groundfish on Georges Bank. Fish. Bull., US 113:264-278.

Runge JA, Kovach AI., Churchill, JH, Kerr LA, Morrison JR, Beardsley RC, Berlinsky DL, Chen C, Cadrin SX, Davis CS, Ford KH, Grabowski JH, Howell WH, Ji R, Jones RJ, Pershing AJ, Record NR, Thomas AC, Sherwood GD, Tallack S, Townsend DW. 2010. Understanding climate impacts on recruitment and spatial dynamics of Atlantic cod in the Gulf of Maine: Integration of observations and modeling. Progress in Oceanography, 87(1), 251–263.

Saba VS, Griffies SM, Anderson WG, Winton M, Alexander MA, Delworth TL, Zhang R. 2016. Enhanced warming of the Northwest Atlantic Ocean under climate change. 121(1), 118–132.

Salter J. 2018. Northeast Multispecies Fishery Management Plan Resource: American plaice (*Hippoglossoides platessoides*). NCRL subject guide 2018-08. National Oceanic and Atmospheric Administration Central Library. (https://repository.library.noaa.gov/view/noaa/20207)

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: 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.

Smith PC, Houghton RW, Fairbanks RG, Mountain DG. 2001. Interannual variability of boundary fluxes and water mass properties in the Gulf of Maine and on Georges Bank: 1993 19971, 2. Deep Sea Research Part II: Topical Studies in Oceanography 48: 37–70.

Smith WG. 1985. Temporal and spatial spawning patterns of the principal species of fish and invertebrates in the Georges Bank region. National Marine Fisheries Service, Sandy Hook Lab. Rep. 85-04

Smith WG, Sibunka JD, Wells A. 1975. Seasonal distributions of larval flatfishes (Pleuronectiformes) on the continental shelf between Cape Cod, Massachusetts, and Cape Lookout, North Carolina, 1965-66. NOAA Tech. Rep. NMFS SSRF-691. 66 p.

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.

Swain D, Chouinard G, Morin R, drinkwater K. 1998. Seasonal variation in the habitat associations of Atlantic cod (*Gadus morhua*) and American plaice (*Hippoglossoides platessoides*) from the southern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences 55: 2548–2561.

Then AY, Honeig JM, Hall NG, and Hewitt DA. 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: 82-92.

Thomas A, Townsend D, Weatherbee R. 2003. Satellite-measured phytoplankton variability in the Gulf of Maine. Continental Shelf Research 23: 971–989.

Thompson WF, Bell FH. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. Rep. Int. Fish. (Pacific Halibut) Comm. 8: 49p.

Thorson JT, Shelton AO, Ward EJ, Skaug,H J. 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.

Thygesen UH, Albertsen CM, Berg CW, Kristensen K, Nielsen A. 2017. Validation of ecological state space models using the Laplace approximation. Environ. Ecol. Stat. 24(2):317-339.

Townsend DW, Pettigrew NR, Thomas MA, Neary MG, McGillicuddy DJ, O'Donnell J. 2015. Water masses and nutrient sources to the Gulf of Maine. Journal of Marine Research 73(3–4): 93–122.

Townsend DW, Rebuck ND, Thomas MA, Karp-Boss L, Gettings R.M. 2010. A changing nutrient regime in the Gulf of Maine. Continental Shelf Research 30(7): 820–832.

Townsend DW, Thomas AC, Mayer L, Thomas M, Quinlan J. 2006. Oceanography of the northwest Atlantic continental shelf (1, W). Chapter 5 in: Robinson AR, Brink KH (eds). The Sea: The Global Coastal Ocean: Interdisciplinary Regional Studies and Syntheses. Harvard University Press.

USCG (United States Coast Guard). 2019. Summary of stock area analysis and investigation of misreporting in the northeast multispecies fishery. Report to New England Fishery Management Council. https://s3.amazonaws.com/nefmc.org/USCG-Groundfish-Misreporting-Investigation-and-Analysis.pdf

USDOC (United States Department of Commerce). 2007. Magnuson-Stevens Act. 2006. Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006. Pub. L. No. 109-479, 120 Stat. 3575 (2007).

Van der Veer HW, Berghahn R, Miller JM, Rijnsdorp AD., 2000. Recruitment in flatfish with special emphasis on North Atlantic species: Progress made by the Flatfish Symposia. ICES J. Mar. Sci. 57: 202–215.

Van Helmond ATM, Mortenson LO, Plet-Hansen KS, Ulrich C, Needle CL, Oesterwind D, Kindt-Larsen L, Catchpole T, Mangi S, Zimmerman C, Olesen HJ, Bailey N, Bergsson H, Dalskov J, Elson J, Hosken M, Peterson L, McElderry H, Ruiz J, Pierre JP, Dykstra C, Poos JJ. 2020. Electronic monitoring in fisheries: Lessons from global experiences and future opportunities. Fish and Fisheries 21: 162-189.

Wake C, Burakowski L, Lines G, McKenzie K. 2006. Cross Border Indicators of Climate Change over the Past Century: Northeastern United States and Canadian Maritime Region. https://www.gulfofmaine.org/council/publications/cross-border-indicators-of-climate-change.pdf.

Walsh SJ. 1994a. Life history traits and spawning characteristics in populations of long rough dab (American plaice) *Hippoglossoides platessoides* (Fabricus) in the North Atlantic. Netherlands Journal of Sea Research 32: 241-254.

Walsh SJ. 1994b. Recruitment variability in populations of long rough dab (American Plaice) *Hippoglossoides platessoides* (Fabricius) in the North Atlantic. Netherlands Journal of Sea Research 32: 421-431.

Walsh SJ, Simpson M, Morgan MJ. 2004. Continental shelf nurseries and recruitment variability in American plaice and yellowtail flounder on the Grand Bank: Insights into stock resiliency. Journal of Sea Research 51: 271–286.

Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O, Bairlein F. 2002. Ecological responses to recent climate change. Nature, 416(6879): 389–395.

Wang SDH, Rosenberg AA. 1997. US New England Groundfish Management Under the Magnuson-Stevens Fishery Conservation and Management Act. Marine Resource Economics 12: 361-366.

Wigley SE, Hersey P, Palmer JE. 2008. A description of the allocation procedure applied to the 1994 to 2007 commercial landings data. NEFSC Ref Doc. 08-18.

Wigley SE, Rago PJ, Sosebee KA, Palka DL. 2006. The Analytic Component to the Standardized Bycatch Reporting Methodology Omnibus Amendment: Sampling Design, and Estimation of Precision and Accuracy. NEFSC. Ref. Doc. 06-22.

Winger PD, He P., Walsh S J. 1999. Swimming endurance of American plaice (*Hippoglossoides platessoides*) and its role in fish capture. ICES Journal of Marine Science, 56: 252-265.

Zhang C-I, and Megrey BA. 2006. A revised Alverson and Carney model for estimating the instantaneous rate of natural mortality. Transactions of the American Fisheries Society 135: 620-633.

Zhang F, Regular PM, Wheeland L, Rideout RM, Morgan MJ. 2020. Accounting for non-stationary stock–recruitment relationships in the development of MSY-based reference points. ICES Journal of Marine Science 78: 2233–2243.

Zhang F, Reid KB, Nudds TD. 2015. Counterintuitive patterns of spawning stock age structure and recruitment in Lake Erie yellow perch (*Perca flavescens*). Fisk A, editor. Can J Fish Aquat Sci. 72:1494–1502.

Zheng N, Robertson M, Cadigan N, Zhang F, Morgan J, Wheel L. 2020. Spatiotemporal variation in maturation: A case study with American plaice (*Hippoglossoides platessoides*) on the Grand Bank off Newfoundland. Canadian Journal of Fisheries and Aquatic Sciences 77: 1688–1699.

APPENDIX A. WORKING PAPERS

Available on shared folder (https://drive.google.com/drive/folders/17e_34Am2w0zG4V9Mwpiv-_VfMEcL-3Yj) and upon request to working paper authors.

Ecological Influences (ToR1)

14. Ecosystem and Climate Influences, by Jamie Behan, Lisa Kerr, Amanda Hart, Alex Hansell, Tyler Paklovitch and Steve Cadrin (November 16, 2021)
16. Plaice Ecosystem Drivers by Jamie Behan and Lisa Kerr (June 21, 2022)

Fishery Data (ToR2)

5. Fishing Industry Knowledge of American plaice, by Tyler Pavlowich, David Richardson, John Manderson and Greg DeCelles (November 9, 2021)
6. Exploration of Fishery Data to Evaluate Catch Rates of American Plaice, by Max Grezlik, Lucy McGinnis, Keith Hankowsky, Gavin Fay, Steve Cadrin and Alex Hansell (November 10, 2021)
7. Catch Rates of American Plaice Trawl Fishery, by Keith Hankowsky, Max Grezlik, Lucy McGinnis, Gavin Fay, Steve Cadrin and Alex Hansell (November 12, 2021)
8. American plaice catch rate analysis using a spatial model, by Andy Jones, Tyler Pavlowich, David Richardson and Anna Mercer (November 13, 2021)
9. Fishery Dependent Data Indices of Abundance (LPUE or CPUE) for American Plaice, by Mark Terceiro (November 16 2021)
10. Electronic Monitoring Data: American Plaice, by Cate O'Keefe, Mel Sanderson and Liz Moore (December 4 2021)

Survey Data (ToR3)

11. Seasonal Variation in Size-at-Age of American Plaice from Survey Data, by Steve Cadrin (November 22 2021)
12. Spatio-temporal dynamics of American plaice (*Hippoglossoides platessoides*) in US waters of the northwest Atlantic, by Alexander Hansell, Larry Alade, Andrew Allyn, Lauran Brewster, Steve Cadrin and Lisa Kerr (December 1 2021)

13. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and biomass estimates for American plaice, by Timothy J. Miller, David E. Richardson, Andrew Jones and Phil Politis (December 9 2021)

Biology (ToR4)

1. Size distribution analysis of American plaice, by Tyler Pavlowich (August 2021)
2. Overview of American Plaice ageing in the Northwest Atlantic, by Josh Dayton and Eric Robillard (September 10 2021)
3. Updating Parameters for Length and Weight Relationships and Length at Age of American Plaice, by Ashley Silver, Tyler Pavlowich and Larry Alade (September 10, 2021)
4. Maturity Analyses of American Plaice in the Georges Bank and Gulf of Maine region, by Shakira Goffe, Daniel Hennen and Larry Alade (September 10, 2021)

15. Approximation of Natural Mortality Rate for American Plaice in US Waters Based on Life History Traits, by Steve Cadrin (January 6, 2022)

Assessment Models (ToR4)

17. American Plaice Assessment Model Developed in Stock Synthesis, by Dan Hennen and Alex Hansell (April 25 2022)
18. A state-space assessment of American plaice using the Woods Hole Assessment Model (WHAM), by Amanda Hart, Lisa Kerr and Tim Miller (June 27 2022)

APPENDIX B. SUMMARY OF REGULATIONS

Summary of changes in management regulations on the US commercial fisheries for American plaice, with minimum mesh size regulations in bold (adapted and updated from NEFSC 2017a).

1953-1977 ICNAF era

- **1953 minimum mesh in body and codend 4 ½ inches.**
- 1970 haddock spawning closures, March-April.
- 1972-1974 closures extended to March-May.
- 1975 closures extended to February-May.

1977 - Present Extended Jurisdiction and National Management

- 1977 USA Fishery Conservation and Management Act of 1976 (FCMA) in effect.
- 1977-1982 Fishery Management Plan (FMP) for Atlantic groundfish: Seasonal spawning closures for haddock (Georges Bank Areas 1 and 2), quotas for haddock, etc
- **1982 Mesh size 5 1/8 in. (130 mm).**
- 1982-1985 The 'Interim Plan' for Atlantic groundfish: Eliminated all catch controls, retained closed area and mesh size regulations, implemented minimum landing sizes.
- **1983 Mesh size increased to 5 ½ inches (140 mm).**
- 1984 October Hague Line separating USA and Canadian fishing zones in the Gulf of Maine and Georges Bank region.
- 1985 Fishery Management Plan for the Multispecies Fishery.
- 1987 American plaice minimum size increased to 12 in. (30 cm).
- 1991 Amendment 4 established overfishing definitions for American plaice as $F_{20\%MSP}$.
- 1992 April Nordmore grate required in Gulf of Maine shrimp fishery; no bycatch of groundfish allowed in shrimp fishery
- 1993 Georges Bank Area 2 closure extended from January 1 - June 30.
- **1994** January, Amendment 5 implemented: expanded Area 2, Area 1 closure not in effect.
    - **May, 6 inch (152 mm) mesh restriction implemented, square or diamond mesh allowed.**
    - December, Georges Bank Area 1, Area 2 and Nantucket Lightship Area closed year-around.

- 1996 July, Amendment 7 implemented: Days-at-sea restrictions. Haddock trip limits Raised to 1000 pounds
- 1997 May, Additional scheduled days at sea restrictions from Amendment 7 accelerated.
- 1998 May, Western Gulf of Maine Closure Area adopted: Jeffery's Ledge area closed to all groundfishing.
    - Rolling closures in the western Gulf of Maine
    - October, Amendment 9: revised overfishing definitions as required by Sustainable Fisheries Act.
- **1999 May, Codend mesh regulations changed to 6-inch diamond mesh, 6 1/2-inch square mesh.**
    - Additional rolling closures adopted in the western Gulf of Maine
    - Cashes Ledge seasonal closure adopted
    - Roller gear limited to a maximum of 12 inches in an area of the western Gulf of Maine.
    - Gulf of Maine cod trip limit ranged from 30 to 400 lbs. in this fishing year.
- 2000 May, May closure implementation on northern Georges Bank.
    - Changes to large mesh permit category, granting additional days at sea to vessels using larger than 6-inch diamond / 6-inch square mesh.
- 2002 June, Additional restrictions adopted during this fishing year (result of lawsuit over FW33):
    - Vessels limited to 25% of allocated days at sea May to July;
    - Increase in minimum mesh size for trawl vessel to 6 ½ inch diamond, 6 ½ inch square;
    - Reduced number of rolling closures in the western Gulf of Maine (effective in January 2003, with result there were additional rolling closures in calendar year 2002 compared to calendar year 2001;
    - Cashes Ledge seasonal closure expanded to year-around closure;
    - Increase in Gulf of Maine cod trip limit to 500 lbs. per day/4,000 lbs per trip;
    - Increase in mesh size for large mesh permit category.
- 2002 August, Reduction in allocated days at sea based on past history of use for each permit;
    - Front-loading of days at sea clock prohibited;

    - Additional restrictions on number and deployment of gillnets.
- 2010 Sector management (vessels in sectors are subject to hard TACs).
    - March, all multispecies vessels fishing on a Category A days at sea allowed to use any legal trawl gear in the Western US/CA Area (statistical areas 522, 525) (lifts restrictions adopted November 20, 2009).
    - April, all multispecies vessels fishing on a Category A days at sea allowed to use a flounder trawl net in the Eastern US/CA area.
    - April, Eastern US/CA area (statistical areas 561, 562) closed to multispecies vessels and harvest, possession, and landing of GB yellowtail flounder from entire US/CA area (statistical areas 522, 525, 561, 562) prohibited.
    - May, Implementation of Amendment 16 and Framework 44. Expansion of sector management program to majority of the fishery. Major revisions to common pool measures for permitted vessels not in sectors. Adoption of additional at-sea and dockside monitoring requirements for sector vessels, and new reporting requirements for other vessels. Adoption of new US/CA area TACs. Adoption of annual catch limit (ACL) and accountability measures (AM) for most stocks. No retention of SNE/MA winter flounder, ocean pout, windowpane flounder, Atlantic wolffish. Specific allocations of GOM cod and GOM haddock made to the recreational and commercial groundfish fisheries. Key elements:
    - Sector Management: Vessels in sectors subject to hard TACs for most stocks, increased at-sea monitoring (targeting 38 percent of trips), dockside monitoring; not subject to trip limits, some GOM rolling closures, groundfish DAS limits. Sector vessels required to retain all legal-sized fish (except limited to one Atlantic halibut, and the five species prohibited). Sectors required to stop fishing in a stock area when a quota (Annual Catch Entitlement, or ACE) for a stock in the area is caught.
    - Common pool: Only a small portion of the ACL available to common pool vessels. Major elements of common pool regulations:
    - Days at Sea: Category A days at sea allocations reduced to 27.5 percent of the Amendment 13 baseline allocation. All days at sea charged in 24 hour increments.
    - Possession limits for cod, pollock, yellowtail flounder, white hake, winter flounder, and witch flounder

- Possession of ocean pout, windowpane flounder, Atlantic wollffish, and SNE/MA winter flounder prohibited.
- Restricted Gear Areas: Areas near CAI and off SNE created to reduce flatfish catches; limited to separator/Ruhle trawls, rope trawl, certain gillnets in these areas. Limited to 500 lbs. of flatfish combined in these areas.
- Special Management Programs: US/Canada Area: Opening delayed until August 1 for trawl vessels. Prohibition on discarding legal sized fish. SNE/MA winter flounder SAP suspended. State waters winter flounder exemption eliminated. CAI Hook Gear Haddock SAP expanded to January 31, area increased, no separation between common pool and sector participants.

- 2011 March, Groundfish common pool trip limit changes.
- 2013 May 1, minimum fish size decreased from 14 in (35.6 cm) to 12 in (30.5 cm).
- 2014
  - Framework Adjustment 51 implemented a revised rebuilding plan for American plaice
  - The Gulf of Maine cod Interim Action closed times and areas to commercial and recreational fisheries for groundfish.
  - Large-mesh Accountability Measures implemented for due to overage in 2012 for northern and southern windowpane flounder – requires the use of selective gears to fish in the areas
- 2015
  - Framework Adjustment 53, effective 2015, altered the boundaries of the rolling closures as Gulf of Maine cod protection measures, seasonal closures in specific time and areas to commercial groundfish fishing.
  - Windowpane Flounder Accountability Measure: Due to the overage of the total catch limit in FY 2014
  - Northern windowpane flounder gear restricted area – large in place for 2015 – commercial groundfish vessels fishing on a groundfish trip are required to use approved selective trawl gear (haddock separator trawl, Ruhle trawl, mini-Ruhle trawl or rope separator trawl
- 2017 Amendment 18, effective in 2017, addressed fleet diversity and accumulation limits.

- 2018 Omnibus Habitat Amendment 2, effective in 2018, resulted in changes in spatial management, among those changes to protect spawning and habitat was moving the eastern boundary of the Western Gulf of Maine closure to the west and the boundaries of Closed Area 1.
- 2019 Stock assessment indicates American plaice is rebuilt, not overfished, and overfishing is not occurring. NMFS determines the official stock status in the following year as rebuilt.
- 2022 Amendment 23, which would improve monitoring in the commercial groundfish fishery includes a target coverage rate of 100% of all sector trips. Implementation is anticipated later in the year. Target coverage rate in 2022 for the sector portion of the commercial groundfish fishery is 99% of sector trips

APPENDIX C. PARTICIPANTS IN FISHERMEN’S MEETINGS

<u>American Plaice (Dab) Research Track Assessment Meeting with Fishermen</u> (September 24 2021, Gloucester MA)

Larry Alade	NEFSC Population Dynamics
Terry Alexander	Maine Fisherman
Jamie Behan	GMRI
Steve Cadrin	UMass Dartmouth SMAST
Cole Carrano	UMass Dartmouth SMAST
Al Cottone	Gloucester Fisherman
Jynessa Dutka-Gianelli	UMass Amherst Gloucester Marine Station
Libby Etrie	Northeast Sector Service Network
Catherine Foley	NEFSC
Vito Giacalone	Northeast Seafood Coalition
Max Grezlik	UMass Dartmouth SMAST
Keith Hankowsky	UMass Dartmouth SMAST
Alex Hansell	NEFSC Population Dynamics
Amanda Hart	Gulf of Maine Research Institute
Dan Hennen	NEFSC Population Dynamics
Andy Jones	NEFSC Cooperative Research
Lisa Kerr	Gulf of Maine Research Institute
Dave Leveille	Sector Manager
Mackenzie Mazur	Gulf of Maine Research Institute
Rich McBride	NEFSC Population Biology
Lucy McGinnis	UMass Dartmouth SMAST
Chris McGuire	The Nature Conservancy
Paul Nitschke	NEFSC Population Dynamics
Sam Novello	Gloucester Fisheries Commission
Jackie Odell	Northeast Seafood Coalition
Cate O'Keefe	Fishery Applications Consulting
Joe Orlando	Gloucester Fisherman

Tyler Pavlowich	NEFSC Cooperative Research
Mel Sanderson	Cape Cod Commercial Fishermen's Alliance
Clark Sandler	Gloucester Fisherman
Ashley Silver	University of Maryland Eastern Shores
Dave Sullivan	Gulf of Maine Ocean Resource Alliance
Mike Walsh	Boston Fisherman.

Groundfish Catch Rates Meeting (November 3 2021, Portland ME)

Jamie Behan	GMRI
Steve Cadrin	UMass Dartmouth SMAST
David Goethel	F/V Ellen Dianne
Amanda Hart	Gulf of Maine Research Institute
Mary Hudson	Sector Manager
Lisa Kerr	GMRI
Ben Martens	Sector Manager
Cate O'Keefe	Fishery Applications Consulting
David Osier	F/V Paulo Mare
Maggie Raymond	Associated Fisheries of Maine
Willard Viola	F/V Black Beauty

APPENDIX D. CRITERIA FOR ABUNDANCE INDICES

Criteria for available fishery-dependent abundance indices (adapted from ICCAT 2021).

Document	GLM1-WP9	GLM22-WP7	Spatiotemporal-WP8
Index	logbook LPUE	dealer-logbook LPUE	observer-study fleet CPUE
Diagnostics	none reported	conventional model validation	conventional model validation
Appropriateness of data exclusions and classifications (e.g. to identify targeted trips).	all trawl trips that caught plaice	all trawl trips that caught plaice	all groundfish targeted trawl trips
Geographical Coverage	entire fishery GOM-GB-SNE-MA	entire fishery GOM-GB-SNE-MA	portions of WGOM, GB
Catch Fraction to the total catch weight	100%	~50% 1996-2010, >80% 2011-2019	~20%?
Length of Time Series relative to the history of exploitation.	1994-2019	1996-2019	2003-2020
Are other indices available for the same time period?	Yes (surveys)	Yes (surveys)	Yes (surveys)
Does the index standardization account for Known factors that influence catchability/selectivity?	area, quarter, tonnage	area, quarter, tonnage, depth, price	spatial, temporal
Are there conflicts between the catch history and the CPUE response?	No - similar to surveys	No - similar to surveys	No - similar to surveys
annual CVs, unaccounted process error (deviations from VPA, high peaks, multiple stanzas, increasing/decreasing catchability)	No - CVs 0.05-2.5, similar trends as VPA, no peaks or stanzas	No - CVs 0.03-0.11, similar trends as VPA, no peaks or stanzas	No - CVs 0.05-0.14, similar trends as VPA, no peaks or stanzas
Assessment of data quality and adequacy of data for standardization purpose (e.g. sampling design, sample size, factors considered)	census but unverified location and effort, 1994-5 uncertain	census but unverified location and effort	random observer coverage, fixed study fleet
Is this CPUE time series continuous?	Yes	Yes	Yes
Other Comment	conventional GLM standardization low spatio-temporal resolution	refined GLM standardization low spatio-temporal resolution	advanced VAST standardization high spatio-temporal resolution