Operational Assessment of 14 Northeast Groundfish Stocks, Updated Through 2018 *

by the Northeast Fisheries Science Center

* This is a Prepublication Copy of the September 2019 Operational Stock Assessment Report.
The report is "in preparation" for publication by the NEFSC. (~~Oct. 3, 2019~~; latest revision: Jan. 7, 2020)

Contents

1 Executive Summary **2**

1.1 Process . . . 2

1.2 Data . . . 4

1.3 Models . . . 4

1.4 Results . . . 4

1.5 Reviewer Comments: Overview . . . 6

2 Gulf of Maine Atlantic cod **26**

2.1 Reviewer Comments: Gulf of Maine Atlantic cod . . . 30

2.1.1 Review Panel Summary . . . 30

2.1.2 Operational Stock Assessment Terms of Reference . . . 30

3 Georges Bank Atlantic cod **38**

3.1 Reviewer Comments: Georges Bank Atlantic cod . . . 41

3.2 Appendix: Georges Bank Atlantic cod . . . 47

4 Georges Bank haddock **56**

4.1 Reviewer Comments: Georges Bank haddock . . . 60

4.1.1 Review Panel Summary . . . 60

4.1.2 Operational Stock Assessment Terms of Reference . . . 60

5 Gulf of Maine haddock **68**

5.1 Reviewer Comments: Gulf of Maine haddock . . . 72

5.1.1 Review Panel Summary . . . 72

5.1.2 Operational Stock Assessment Terms of Reference . . . 72

6 Cape Cod-Gulf of Maine yellowtail flounder **80**

6.1 Reviewer Comments: Cape Cod-Gulf of Maine yellowtail flounder . . . 84

6.1.1 Review Panel Summary . . . 84

6.1.2 Operational Stock Assessment Terms of Reference . . . 84

7 Southern New England-Mid Atlantic yellowtail flounder **92**

7.1 Reviewer Comments: Southern New England-Mid Atlantic yellowtail flounder 96

7.1.1 Review Panel Summary 96

7.1.2 Operational Stock Assessment Terms of Reference 96

8 Georges Bank Winter Flounder **105**

8.1 Reviewer Comments: Georges Bank Winter Flounder 109

8.1.1 Review Panel Summary 109

8.1.2 Operational Stock Assessment Terms of Reference 109

9 Gulf of Maine-Georges Bank American plaice **118**

9.1 Reviewer Comments: Gulf of Maine-Georges Bank American plaice 122

9.1.1 Review Panel Summary 122

9.1.2 Operational Stock Assessment Terms of Reference 122

10 Witch flounder **130**

10.1 Reviewer Comments: Witch flounder 133

11 White hake **139**

11.1 Reviewer Comments: White hake 143

11.1.1 Review Panel Summary 143

11.1.2 Operational Stock Assessment Terms of Reference 143

12 Pollock **151**

12.1 Reviewer Comments: Pollock 156

12.1.1 Review Panel Summary 156

12.1.2 Operational Stock Assessment Terms of Reference 156

13 Atlantic halibut **164**

13.1 Reviewer Comments: Atlantic halibut 167

14 Gulf of Maine - Georges Bank windowpane flounder **173**

14.1 Reviewer Comments: Gulf of Maine - Georges Bank windowpane flounder 177

14.1.1 Review Panel Summary 177

14.1.2 Operational Stock Assessment Terms of Reference 177

15 Southern New England - mid-Atlantic windowpane flounder **184**

15.1 Reviewer Comments: Southern New England - mid-Atlantic windowpane flounder . 187

16 Appendices **193**

16.1 Generic terms of reference for operational assessments 193

16.2 Summary of Assessment Oversight Panel Meeting 194

16.3 Description of New England and Mid-Atlantic Region Stock Assessment Process . . 200

16.4 Onsite participants 214

1 Executive Summary

1.1 Process

Assessments for 14[1] of the 20 groundfish stocks (Table 1) in the New England Fishery Management Council's (NEFMC) Multispecies Groundfish Fisheries Management Plan were updated and reviewed during September 9-12, 2019 at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA. This represents the sixth assessment of the status of groundfish stocks since 2001. The first three assessments were produced through the Groundfish Assessment Review Meeting (GARM) process (NEFSC 2002, 2005, 2008). Thirteen of the groundfish stocks were updated through the Operational Assessment process in 2012 (NEFSC 2012). All 20 groundfish stocks were updated using operational assessments in 2015 (NEFSC 2015), and 19 of the 20 were updated using operational assessments in 2017 (NEFSC 2017). Operational assessments, first described by the Northeast Regional Coordinating Council (NRCC) in 2011, rely on decisions of previous benchmarks for model formulation and definition of biological reference points (BRPs). The efficiency of the Operational Assessment process increases the frequency of assessments, but reduces the ability to modify model structure either in response to new data or external inputs. In 2019, operational assessments were replaced with management track assessments that allow greater flexibility in model structure while retaining much of the efficiency of operational assessments. Guidelines for management track assessments and their counterpart research track assessments (see Appendix 16.3) were initially developed through collaborative discussions among the NRCC, NEFSC, NEFMC, MAFMC and ASMFC. Under the management track process, increasing the number of changes in model structure in an updated assessment requires increasing levels of peer review. The Assessment Oversight Panel (AOP) decided on the level of review each assessment would receive in its meeting on June 20, 2019. At the June meeting, the AOP reviewed presentations on each assessment prepared by individual analysts, which included any proposed changes to the assessments as well as plans for how scientific advice would be provided for each stock if the primary analytical assessment was not accepted by the peer review panel (sometimes referred to as "Plan B" assessment advice). See Section 16.2 for a summary of the AOP meeting, and the asignment of each stock assessment to a peer review level. Ten of the 14 groundfish stocks were assigned to Level 2 or 3 for subsequent review by the Peer Review Panel (i.e., Panel). Level 1 stocks (Georges Bank cod, Southern windowpane flounder, witch flounder, and Atlantic halibut) were not reviewed by the Panel.

Following the established assessment update process, the NEFSC provided a data-rich dedicated website to supplement the information provided in individual species assessment reports.

The Panel consisted of the following individuals:

- Pat Sullivan (Chair), Cornell University, NEFMC Scientific and Statistical Committee
- Steve Cadrin, University of Massachusetts Dartmouth
- Chris Legault, NEFMC Scientific and Statistical Committee

[1] some stocks have been excluded because they are data updates (wolffish, ocean pout), while others are excluded because they are scheduled for research track assessments soon (Gulf of Maine and Southern New England winter flounders, and redfish), or were already assessed under TRAC (Georges Bank yellowtail flounder).

The Panel was responsible for reviewing each of the stock assessments. Primary and supporting documents for each assessment were available prior to the meeting. Each lead assessment scientist (Table 2) prepared a short presentation to describe the past and updated assessment results and address key sources of uncertainty (see agenda). Following the presentation, the Panel was instructed as follows:

The Peer Review is to determine whether the completed operational assessment is technically sufficient to (a) evaluate stock status and (b) provide scientific advice; (c) successfully address the assessment Terms of Reference. The Peer Review may determine that application of the baseline model in the operational assessment ("Plan A") has not worked; if so, the alternative approach to the assessment ("Plan B") will be implemented, and the stock will be referred to the research track.

For each stock assessment, the reviewer report should address whether each stock assessment TOR was completed successfully. The report should make clear whether the original modeling approach (i.e., Plan A) was accepted, or whether Plan B was recommended. The report should identify major sources of uncertainty in the stock assessment and include qualitative descriptions of stock status based on simple indicators/metrics The report can also make recommendations for improving the assessment in the future.

If an assessment was not considered suitable for estimation of OFL the Panel was responsible for recommending an alternative basis. Additionally, the Peer Review panel was asked to include qualitative descriptions of stock status. NOAA Fisheries has final responsibility for making the stock status determination based on best available scientific information, which in the absence of an accepted quantitative assessment, may be qualitative.

The individual assessment sections within this report are standardized and designed to capture the most relevant information for reviewers and fishery managers. The report structure was developed with, and approved by, a subcommittee of the NRCC, followed by NRCC feedback about the report structure. Each assessment is supported by an online set of companion tables, figures and maps, which provide primary users of the assessment information (e.g., Plan Development Teams, Science and Statistical Committee) with necessary details. The online data portal (SASINF) also contains model inputs and outputs that can be used directly in NOAA Fisheries Toolbox applications.

The meeting was broadcast as a webinar and all sessions were open to the public. The meeting agenda included a daily public comment period. Members of the audience and individuals on the phone were included in the discussions of the panel at the discretion of the Panel chair. However, the tight timeline for completing the assessments required a strong adherence to the terms of reference and the description of the operational assessment process developed by the NRCC. Onsite participants in Woods Hole are listed in Section 16.4.

SAW Process Chairman/Editor's Note: *After the September 2019 peer review, NEFSC staff became aware of two assessment issues. First, there were differences between MRIP datasets used to estimate annual recreational catches of GB cod, GOM cod, GOM haddock, and Pollock. These differences were attributable to a change in a computer program used to extract MRIP data. The magnitude of those differences was minor, and not considered significant enough to redo analytical assessments of GOM cod, GOM haddock, and pollock. Second, the GB cod assessment was affected by the MRIP issue, and also used an incorrect assumed recreational discard mortality rate of 100%, rather than 30%. Given that the GB cod assessment had two issues, which in combination resulted in a small change in catch advice, that assessment was corrected in January 2020 and is now included as an appendix to Chapter 3 "Georges Bank Atlantic cod".*

1.2 Data

The groundfish management track assessments used the following standard procedures for updating data from landings, discards and surveys (Table 3). The US commercial landings are estimated by market category from the area allocation ("AA") tables, which combine dealer and vessel trip reports to determine where fish were caught. The US commercial discards are estimated by gear types using the Standardized Bycatch Reporting Methodology (SBRM), which combines observer data (including at-sea monitors) and dealer landings. The US recreational landings and discards come from the Marine Recreational Information Program (MRIP), including recent revisions to historical data. Both commercial and recreational discards have species-specific discard mortality rates applied to the discarded fish. Catch-at-age is estimated using age-length keys applied to expanded length frequency distributions. For white hake, which is landed headed, the age-length key is applied to predicted lengths based on dorsal fin to caudal fin length. Additional sources of catch for some species come from Canadian or other foreign fishing.

The NEFSC spring and fall bottom trawl surveys are the most common source of information for population trends (Table 3). These surveys are calibrated to "Albatross units" in most cases to allow for the longest time series possible. NOAA ship *Henry B. Bigelow* replaced the *Albatross IV* as the primary bottom trawl survey vessel in spring 2009. In some instances the calibration coefficient varies by length but in others a simple scalar adjustment is applied to all length classes. Other surveys used include the Massachusetts Division of Marine Fisheries spring and fall bottom trawl surveys, the Maine-New Hampshire spring and fall bottom trawl surveys, the Canadian Department of Fisheries and Oceans February survey, and some additional state surveys. Catch per unit effort is not typically used as a source of population trends due in part, to the many regulatory changes that have occurred over time in the Northeast that influence fishing behavior and catch rates, as well as the fact that directed fishing is non-random in time and space, which generally causes bias in estimates of abundance. All updated assessments used a consistent quality assurance criterion (known as TOGA; Politis et al. 2014) for surveys conducted by the NOAA ship *Henry B. Bigelow*.

1.3 Models

Based on previous 2017 operational assessments (Table 4; NEFSC, 2017), there are 11 stocks assessed with an age-based approach. Seven use the statistical catch-at-age model ASAP while 4 others use virtual population analysis (VPA). For the 4 VPA stocks, the 2019 spring survey information was included in the model. The remaining 5 stocks are assessed with a range of model types including index (AIM, FSD, 'plan B smooth'), and direct survey expansion. The reference points for the age- and length-based assessments were derived from stochastic projections of the F_{MSY} (or F_{MSY} proxy) for many years (typically 100), while the other assessment types use stock-specific rules for deriving the reference points. Technical descriptions of the biomass, fishing mortality and reference point estimators used for each stock are shown in Table 4.

1.4 Results

Management track assessments were conducted in 2019 for 14 of the 20 stocks in the Northeast Multispecies Fishery Management Plan (Table 1). The management track assessments replicated

the methods recommended in the most recent benchmark decisions, as modified by any subsequent operational assessments or updates (Table 2). Information supplemental to the assessment report for each stock can be found on the Stock Assessment Support Information (SASINF) website. The reviewers accepted all of the assessments as a scientific basis for management and the assessments provided catch advice for all 14 stocks. Recommended stock status did not change for 10 of the 14 stocks, improved for 2 stocks and declined for 2 stocks (Table 5).

Each of the 14 stock chapters, 10 of which were provided to the Panel for peer review, contains the assessment results followed by a section entitled "Reviewer Comments," which describes final Panel decisions at the conclusion of the peer review. **In this Executive Summary, tables and figures related to stock status from the 2019 review reflect the Panel recommendations** (Tables 5 - 6; Figures 1 - 2; see Legends of those Tables and Figures for key details).

The number of stocks with retrospective adjustments (also called rho adjustments) applied increased from the last assessment from 7 to 8 (Table 7). Decisions to apply a retrospective adjustment to estimates of terminal year biomass and fishing mortality rates were based on whether the rho adjusted value was outside the 90% joint confidence region for the model estimates. This principle was supported by the AOP and was applied to adjust biomass estimates for Georges Bank haddock, Southern New England yellowtail flounder, Cape Cod/Gulf of Maine yellowtail flounder, Georges Bank winter flounder, American plaice, white hake, and pollock (Table 8). Gulf of Maine cod was an exception because of earlier guidance from the SARC 55 review panel. Despite the presence of a significant retrospective pattern at that meeting no adjustments were made; later Operational Assessments panels have followed that precedent.

Stock status recommendations for the 14 groundfish stocks are summarized in Tables 5 and 6. One stock is experiencing overfishing: Gulf of Maine cod. Eight groundfish stocks are overfished (Table 5). Based on these recommendations, the number of overfished stocks and stocks experiencing overfishing has generally decreased since GARM III in 2007 (Figure 3), and the magnitude of overfishing or depletion for several stocks has generally decreased (Figures 1 and 2).

Simultaneous assessments of 14 groundfish stocks allowed a comprehensive examination of trends in spring and fall survey indices (Figures 4 and 5, respectively). For the majority of stocks the average of the most recent 5 years is below the time series mean for that stock.

Estimates of overall (aggregate) groundfish minimum swept area biomass are at or near an all-time high (Figures 6 and 7). However, the current stock diversity of the overall groundfish biomass is less than that seen in the 1960s and 1970s. Current groundfish biomass is dominated by only a few stocks. For example, the combined biomass of the Georges Bank haddock, Gulf of Maine haddock, and pollock stocks currently constitute more than 91% of the overall groundfish biomass observed in NEFSC spring surveys (Figure 6). It is important to note that the minimum swept area biomass estimates assume a common capture efficiency of 1.0 across all years. Actual biomasses, as derived from models, are adjusted for catchability and selectivity estimates and are higher than the minimum swept area estimates. Unfortunately model-based estimates are not available for all stocks over the entire time period of the surveys (i.e. since 1963); the primary limitation is the availability of age information from the commercial catches that would be needed to support full age-based assessments.

For 12 stocks, model-based biomass estimates can be computed from 1985 onward. The striking increase in abundance since 1985 is driven primarily by Georges Bank haddock, and pollock (Figure 8). Pollock biomass from the stock assessment is much higher than the swept area estimates because of a dome-shaped selectivity pattern in both the survey and catch data. This suggests that a substantial fraction of the stock biomass is unavailable to either the fishery or survey gear. The chapter describing the pollock assessment includes a sensitivity run in which the assumption of dome-shaped selectivity is removed, resulting in a biomass estimate that is less than half as large. The increase in model based estimates of overall biomass, with or without pollock, is consistent with the trends revealed in the swept area estimates (Figures 6, 7 and 8).

An advantage of conducting multiple assessments simultaneously is that measures of productivity can be compared over time. Reductions in average weight-at-age, declines in recruitment and shifts in age-at-maturity all influence the estimated biomass at maximum sustainable yield and total MSY. As such, the combined single species stock assessments provide valuable measures of ecosystem productivity, irrespective of the underlying environmental or ecological causes. Reductions in average weights-at-age have occurred for stocks at high abundance, such as Georges Bank haddock, but also for stocks at low abundance, such as witch flounder. Hence, density dependence alone is insufficient to explain this across all stocks. Reductions in recruitment are often associated with declines in stock size but inter-annual variation often masks trends. Aggregate estimates of total B_{MSY} are available for 8 stocks over the past decade. Total B_{MSY} for these stocks increased by 61% between 2008 and 2019 from 483 kt to 1252 kt (Figure 9).

1.5 Reviewer Comments: Overview

The 2019 Groundfish Operational Assessment Review Panel thanks the assessment leads for their work in preparing these assessments and for their responsiveness to Panel requests and questions before and during the review meeting. It was clear that the effort put into the assessments and the presentations was significant and that was much appreciated.

The Panel welcomed and found beneficial the flexibility that has been incorporated into the new Management Track assessment update process. This flexibility allowed the assessment teams to address several issues in a reasonable and progressive way. We also noted greater ease and cordiality among all participants as a result. However, not everyone at the meeting was made aware of the degree of flexibility allowed. The Panel encourages suggestions for future modeling exploration from the leads and other participants when considering future assessment updates. The decision to create levels indicating the degree of review needed was also helpful. The workload for the meeting was significantly reduced by focusing on the Level 2 and 3 stocks and setting aside the Level 1 stocks for a less formal review. The Panel suggests that while having the Level structure is helpful, some flexibility might be allowed on the Panel's discretion to dive deeper, if necessary, during the review. However, it is also important to have consistent expectations for what will be involved in preparing an update prior to the review. Here too, the degree of flexibility allowed in the new process was not fully understood by all assessment leads or panel members before the meeting. As the process evolves, the opportunities for model improvements should be clearly outlined for the assessment leads and the broader community. The Panel appreciated that the schedule allowed more time for discussions and processing the significance of the updated assessments during the meeting and

for drafting the Panel Report. Model exploration in advance of the Operational Assessment Panel meeting will help to identify allowable revisions and the level of review needed.

The Panel found the Data Portal to be helpful for supporting a thorough review of the assessments. The supplemental files allowed for in-depth reviews by the Panel members prior to the review. The list of supplemental files varied among stocks and could be standardized to facilitate pre-meeting reviews, for example in the draft report section, presentation files, model input files, model output files, most recent benchmark assessment report, and the most recent assessment report. Systematic use of naming conventions (i.e., species, area) helped the Panel keep track the multiple files that were provided. Routine use of stock codes can help to avoid confusion among stocks. Some filenames were not consistent for some stocks and further standardization would be useful.

The Panel appreciated the more in-depth explanation of how the chain sweep survey study was considered for use in each assessment. Including reference to these studies and how they are implemented will continue to be useful going forward. Comparisons of model estimates of total biomass to area-swept survey estimates were informative as a confirmation of model scaling. Confidence limits for both the model and the area-swept estimates should be provided for making inferences about scaling. Continued collaboration with the Northeast Trawl Advisory Panel (NTAP) should help in making the best use of the available information on survey efficiency for each stock.

The Panel noted that a holistic overview of the assessments, similarities and differences in stock behavior and status would have been a valuable addition to this review. Considering questions, such as what is driving the consistent retrospective patterns, poor recent recruitment, and declining weights at age across numerous stocks would have benefited the NEFSC assessment teams, the Panel, the SSC, the Council and other stakeholders. The ecosystem overview presented at the 2017 groundfish updates provided important context for some recent multispecies trends (e.g., changes in weight at age, recruitment, as well as natural mortality and retrospective patterns). If time allows, such context should be presented with this report and considered in future management track and research track assessments.

The revised MRIP estimates of recreational catch were included in the updated assessments and affected some assessments. A workshop on MRIP estimates for northeast stocks could help to understand the new estimates and possibly improve their application in assessments.

Misreporting of commercial landings and underestimation of commercial discards was recognized for some groundfish stocks, but not explicitly addressed in these assessments. Measurement error and bias in catch statistics should be considered in research track stock assessments. Electronic monitoring data was not used in these assessments because they currently represent a small portion of the total trips and multispecies landings. However, electronic monitoring may represent a larger portion of catch and effort for some stocks and should be considered in future research track assessments.

References

Politis PJ, Galbraith JK, Kostovick P, Brown RW. 2014. Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow. US Dept Commer, Northeast Fish

Sci Cent Ref Doc. 14-06; 138 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD14-06

Miller, T.J., Richardson, D.E., Politis P.J., Blaylock, J. *In press*. Northeast Fisheries Science Center bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. US Dept Commer, Northeast Fish Sci Cent Ref Doc. *In press*; 26 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-24; 251 p. CRD15-24

Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026 CRD12-06

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. US Dep Commer, NOAA FIsheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii. CRD08-15

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. U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 05-13; 499 p. CRD05-13

Table 1: List of stocks included in the 2019 groundfish operational assessment and the abbreviations used for each in tables and figures in this document.

Stock Abbrev	Stock Name
CODGM	Gulf of Maine cod
CODGB	Georges Bank cod
HADGM	Gulf of Maine haddock
HADGB	Georges Bank haddock
YELCCGM	Cape Cod/Gulf of Maine yellowtail flounder
YELSNEMA	Southern New England/Mid-Atlantic yellowtail flounder
FLWGB	Georges Bank winter flounder
PLAUNIT	American plaice
WITUNIT	Witch flounder
HKWUNIT	White hake
POLUNIT	Pollock
HALUNIT	Atlantic halibut
FLDGMGB	Gulf of Maine/Georges Bank windowpane flounder
FLDSNEMA	Southern New England/Mid-Atlantic windowpane flounder

Table 2: Lead scientist for each stock (current/previous if different), information about last assessment, including: the forum for review of the last assessment (Forum), the type of assessment done (Type), publication year (Pub.), the terminal year of the catch data included (Term. yr.), overfished/overfishing status, rebuilding status, and reference. *Note: Op. Assess = Operational Assessment*

Stock	Lead	Forum	Type	Pub.	Term. yr.	Overfished?	Overfishing?	Rebuild status	Reference
CODGM	Perretti/Palmer	Op. Assess	Update	2017	2016	Yes	Yes	By 2024	CRD17-17
CODGB	Legault	Op. Assess	Update	2017	2016	Yes	Unknown	By 2026	CRD17-17
HADGM	Perretti/Palmer	Op. Assess	Update	2017	2016	No	No	Rebuilt	CRD17-17
HADGB	Brooks	Op. Assess	Update	2017	2016	No	No	Rebuilt	CRD17-17
YELCCGM	Alade	Op. Assess	Update	2017	2016	Yes	Yes	By 2023	CRD17-17
YELSNEMA	Alade	Op. Assess	Update	2017	2016	Yes	Yes	Rebuilt	CRD17-17
FLWGB	Hendrickson	Op. Assess	Update	2017	2016	Yes	Yes	By 2017	CRD17-17
PLAUNIT	Alade/Terceiro	Op. Assess	Update	2017	2016	No	No	By 2024	CRD17-17
WITUNIT	Wigley	SARC 62	Benchmark	2017	2016	Yes	Unknown	By 2017	CRD17-03
HKWUNIT	Sosebee	Op. Assess	Update	2017	2016	No	No	By 2014	CRD17-17
POLUNIT	Linton	Op. Assess	Update	2017	2016	No	No	Rebuilt	CRD17-17
HALUNIT	Hennen	Op. Assess	Update	2017	2016	Yes	No	By 2055	CRD12-06
FLDGMGB	Chute	Op. Assess	Update	2017	2016	Yes	No	By 2017	CRD17-17
FLDSNEMA	Chute	Op. Assess	Update	2017	2016	No	No	Rebuilt	CRD17-17

Table 3: Data used in each assessment. The column heads are US commercial landings (US c-lnd), US commercial discards (US c-dis), US recreational landings (US r-lnd), US recreational discards (US r-dis), Canadian catch (CA cat), Northeast Fisheries Science Center spring, fall and winter surveys (NE S, NE F and NE W), Massachusetts spring and fall surveys (MA S and MA F), Maine/New Hampshire spring and fall surveys (ME/NH S and ME/NH F) and Canadian Department of Fisheries and Oceans February survey (DFO S).

	Catch					**Surveys**							
Stock	US c-lnd	US c-dis	US r-lnd	US r-dis	CA Cat	NE S	NE F	NE W	MA S	MA F	ME/NH S	ME/NH F	DFO S
CODGM	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	No	No	No
CODGB	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
HADGM	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No	No	No	No
HADGB	Yes	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	Yes
YELCCGM	Yes	Yes	No	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No
YELSNEMA	Yes	Yes	No	No	No	Yes	Yes	Yes	No	No	No	No	No
FLWGB	Yes	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	Yes
PLAUNIT	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No	No	No
WITUNIT	Yes	Yes	No	No	No	Yes	Yes	No	No	No	No	No	No
HKWUNIT	Yes	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	No
POLUNIT	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No	No	No	No
HALUNIT	Yes	Yes	No	No	Yes	No	Yes	No	No	No	No	No	No
FLDGMGB	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	No
FLDSNEMA	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	No

Table 4: Assessment type and reference points from 2015 operational assessment CRD15-24. Biomass and yield values are in metric tons. *Note: sp=stochastic projection and surv. B = survey biomass.*

Stock	Assess.	Type	F def.	B def.	F_{MSY} type	F_{MSY} value	B_{MSY} type	B_{MSY} value	MSY type	MSY value
CODGM(M=.2)	ASAP	age-based	F_{Full}	SSB	$F_{40\%SPR}$	0.17	sp	40,604	sp	7,049
CODGM(M_{ramp})	ASAP	age-based	F_{Full}	SSB	$F_{40\%SPR}$	0.17	sp	59,714	sp	10,502
CODGB	empirical	smoothed survey	$\frac{catch}{surv.B}$	surv. B	NA	NA	NA	NA	NA	NA
HADGM	ASAP	age-based	F_{Full}	SSB	$F_{40\%SPR}$	0.455	sp	6,769	sp	1,547
HADGB	VPA	age-based	avg F ages 5-7	SSB	$F_{40\%SPR}$	0.35	sp	104,312	sp	24,400
YELCCGOM	VPA	age-based	avg F ages 4-6	SSB	$F_{40\%SPR}$	0.273	sp	4,640	sp	1,154
YELSNEMA	ASAP	age-based	avg F ages 4-5	SSB	$F_{40\%SPR}$	0.347	sp	1,986	sp	547
FLWGB	VPA	age-based	avg F ages 4-6	SSB	Fmsy	0.522	sp	7,600	sp	3,500
PLAUNIT	VPA	age-based	avg F ages 6-9	SSB	$F_{40\%SPR}$	0.216	sp	13,503	sp	2,942
WITUNIT	empirical	survey expansion	$\frac{catch}{surv.B}$	surv. B	NA	NA	NA	NA	NA	NA
HKWUNIT	ASAP	age-based	F_{Full}	SSB	$F_{40\%SPR}$	0.1839	sp	30,948	sp	4,867
POLUNIT(base)	ASAP	age-based	avg F ages 5-7	SSB	$F_{40\%SPR}$	0.26	sp	105,51	sp	19,42
POLUNIT(flat)	ASAP	age-based	avg F ages 5-7	SSB	$F_{40\%SPR}$	0.24	sp	60,73	sp	11,69
HALUNIT	FSD	empirical	NA	NA	NA	NA	NA	NA	NA	NA
FLDGMGB	AIM	index	$\frac{catch}{surv.B}$	surv. B	replacement ratio	0.340	$\frac{MSY\,proxy}{F_{MSY\,proxy}}$	2.060	median catch 1995-2001	700
FLDSNEMA	AIM	index	$\frac{catch}{surv.B}$	surv. B	replacement ratio	1.918	$\frac{MSY\,proxy}{F_{MSY\,proxy}}$	0.261	median catch 1995-2001	500

Table 5: Synopsis of recommended status by stock from the 2019 peer review. These recommendations will be considered by NMFS in making final status determinations.

		Recommended Status	
Stock	Stock Name	Overfishing?	Overfished?
CODGM	Gulf of Maine cod	Yes	Yes
CODGB	Georges Bank cod	Unknown	Yes
HADGM	Gulf of Maine haddock	No	No
HADGB	Georges Bank haddock	No	No
YELCCGM	Cape Cod/Gulf of Maine yellowtail flounder	No	No
YELSNEMA	S. New Eng./Mid-Atl. yellowtail flounder	No	Yes
FLWGB	Georges Bank winter flounder	No	Yes
PLAUNIT	American plaice	No	No
WITUNIT	Witch flounder	Unknown	Yes
HKWUNIT	White hake	No	Yes
POLUNIT	Pollock	No	No
HALUNIT	Halibut	Unknown	Yes
FLDGMGB	Gulf of Maine/Georges Bank windowpane flounder	No	Yes
FLDSNEMA	S. New Eng./Mid-Atl. windowpane flounder	No	No

Table 6: Summary of Operational Assessment estimates of biomasses and fishing mortality rates in 2018 and biological reference points for groundfish stocks. Reference points are not estimable for some stocks. For FLDGMGB, the 2019 review panel recommended using the estimates of the F_{MSY} *proxy* (0.34 kt/(kg/tow)) and the B_{MSY} *proxy* (2.06 kg/tow) from the 2017 stock assessment NEFSC CRD17-17. This table lists the BRP values from the 2019 assessment, which were not used for stock status determination of FLDGMGB.

Stock	Model type	B_{2018} (mt)	B_{MSY} (mt)	$\frac{B_{2018}}{B_{MSY}}$	F_{2018}	F_{MSY}	$\frac{F_{2018}}{F_{MSY}}$	MSY (mt)	ρ adj?	Comments
CODGM	ASAP (M=0.2)	3,752	42,692	0.09	0.188	0.17	1.09	7,580	No	
CODGM	ASAP (M-ramp)	3,838	63,867	0.06	0.198	0.17	1.13	11,420	No	
CODGB	Empirical	3.742			0.119				No	Smoothed survey indices used to estimate biomass
HADGB	VPA	507,130	138,924	3.65	0.061	0.33	0.18	30,489	Yes	
HADGM	ASAP	82,763	7,993	10.35	0.082	0.37	0.22	1,597	Yes	
YELCCGM	VPA	2,125	3,439	0.62	0.092	0.32	0.29	1,138	Yes	
YELSNEMA	ASAP	90	1,779	0.05	0.259	0.35	0.73	492	Yes	
FLWGB	VPA	2,175	8,910	0.24	0.223	0.52	0.43	4,260	Yes	
PLAUNIT	VPA	17,748	15,293	1.16	0.089	0.26	0.34	3,301	Yes	
WITUNIT	Empirical	35,585			0.02				No	Average survey biomass, exploitation ratio used
HKWUNIT	ASAP	15,891	31,828	0.50	0.129	0.17	0.77	4,601	Yes	
POLUNIT	ASAP (base)	212,416	124,639	1.70	0.038	0.27	0.14	19,856	Yes	Flat top selectivity model was used for sensitivity testing
POLUNIT	ASAP (flat top)	71,322	70,721	1.01	0.094	0.26	0.36	12,007	Yes	see above
FLDGMGB	AIM	0.248	3.489	0.07	0.335	0.18	1.81	647	No	Biomass in kg/tow. F values reflect exploitation rate
FLDSNEMA	AIM	0.319	0.187	1.71	1.632	1.78	0.92	333	No	Biomass in kg/tow. F values reflect exploitation rate
HALUNIT	Empirical	0.94			126				No	Biomass is the catch multiplier. F is the recommended catch in mt

Table 7: Comparison of biomass (B) and fishing mortality rate (F) Mohn's rho values (ρ) by stock between the previous assessment (ρ_{last}) and the 2019 (ρ_{2018}) updates. The biomass (B_{2018}) and fishing mortality rate (B_{2018}) point estimates and ρ adjusted values (Adj.) are provided for the 2019 operational assessments. Stocks using ρ adjusted values in the last assessment and the 2019 assessments (ρ adj. vs. pt. est. for those stocks that did not use the ρ adjustment) are identified, along with the type of ρ adjustment used in the 2019 assessment (NAA=numbers at age, SSB=spawning stock biomass applied to all ages), are also provided. Only age-based and length-based stocks that could exhibit retrospective patterns are included in this table. Missing ρ values indicate a minor retrospective pattern was found and no retrospective adjustments were made.

Stock	Model type	ρ_{last}	ρ_{2018}	B_{2018}	Adj.	ρ_{last}	ρ_{2018}	F_{2018}	Adj.	Last Assess.	2019	Proj. Adj.
CODGM	ASAP (M=0.2)	0.53	0.52	3752	2,468	-0.31	-0.29	0.19	0.27	pt. est.	pt. est.	none
CODGM	ASAP (M-ramp)	0.30	0.29	3838	2,976	-0.17	-0.16	0.20	0.24	pt. est.	pt. est.	none
HADGM	ASAP		-0.24	63143	82,763		0.28	0.10	0.08	pt. est.	ρ adj.	none
HADGB	VPA	0.89	0.69	859587	507,130	-0.63	-0.44	0.03	0.06	ρ adj.	ρ adj.	SSB
YELCCGM	VPA	0.76	0.30	2753	2,125	-0.39	-0.15	0.08	0.09	ρ adj.	ρ adj.	NAA
YELSNEMA	ASAP	0.97	0.63	147	90	-0.47	-0.31	0.18	0.26	ρ adj.	ρ adj.	NAA
FLWGB	VPA	0.54	0.55	3372	2,175	-0.31	-0.35	0.14	0.22	ρ adj.	ρ adj.	SSB
PLAUNIT	VPA	0.14	0.27	22490	17,748	-0.32	-0.20	0.07	0.09	ρ adj.	ρ adj.	NAA
HKWUNIT	ASAP	0.20	0.31	20757	15,891	-0.12	-0.17	0.11	0.13	ρ adj.	ρ adj.	NAA
POLUNIT	ASAP (base)	0.23	0.30	276305	212,416	-0.28	-0.29	0.03	0.04	ρ adj.	ρ adj.	NAA
POLUNIT	ASAP (flat)	0.41	0.58	112633	71,322	-0.35	-0.38	0.06	0.09	ρ adj.	ρ adj.	NAA

Table 8: The biomass (B) and exploitation rate (F) values used for status determination were adjusted to account for a retrospective pattern in some stocks. In general, when the B or F values adjusted for restrospective pattern (B_ρ and F_ρ) were outside of the approximate 90% confidence interval (Conf. limits), the ρ adjusted values were used to determine stock status (Adj. = Yes). Only stocks that had both an estimable 7-year Mohn's ρ for B and F and estimable approximate 90% confidence limits on terminal year B and F values are included.

Stock	B_{2018}	B_ρ	Conf. limits	F_{2018}	F_ρ	Conf. limits	Adj?
CODGM(M=0.2)	3,752	2,468	2,582 - 5,071	0.188	0.265	0.113 - 0.263	No
CODGM(M ramp)	3,838	2,976	2,922 - 5,094	0.198	0.236	0.145 - 0.263	No
HADGB	859,587	507,130	614,031 - 1,253,991	0.034	0.061	0.026 - 0.046	Yes
YELSNEMA	147	90	113 - 200	0.178	0.259	0.12 - 0.25	Yes
YELCCGM	2,753	2,125	2,325 - 3,308	0.078	0.092	0.06 - 0.1	Yes
FLWGB	3,372	2,175	2,725 - 4,346	0.145	0.223	0.111 - 0.194	Yes
PLAUNIT	22,490	17,748	19,592 - 26,220	0.071	0.089	0.063 - 0.084	Yes
HKWUNIT	20,757	15,891	17,792 - 24,216	0.107	0.129	0.088 - 0.128	Yes
POLUNIT(base)	276,305	212,416	83,067 - 364,936	0.027	0.038	0.042 - 0.035	Yes

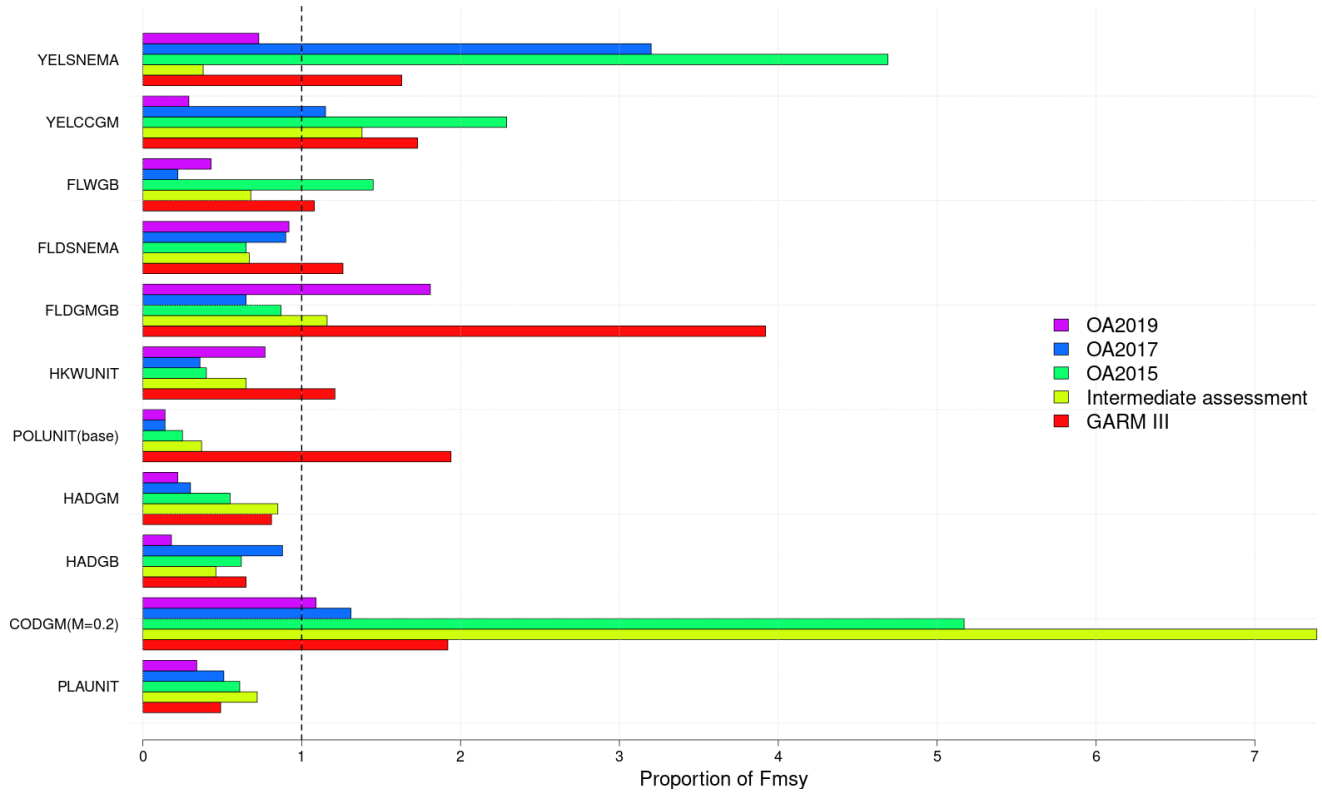

Figure 1: Changes in the ratio of fishing mortality to FMSY proxy from 2007 (GARM III) to 2019 (OA 2019) for selected Northeast Multispecies Fishery Management Plan groundfish stocks. The results from the assessment prior to the OA 2015 assessment are shown for each stock to provide 'Intermediate' values. Stocks on which overfishing is occurring are those where the $\frac{F_{terminal}}{F_{MSYproxy}}$ ratio is greater than 1. *Notes: (1) Because Georges Bank cod, halibut and witch flounder do not currently have quantitative biomass reference points, they were not included in this figure. (2) The plotted value for FLDGMGB (Northern windowpane flounder) in OA2019 reflects the BRPs from the 2019 assessment. The peer review panel recommended using the BRPs from the 2017 assessment for status determination (see Table 6 legend).*

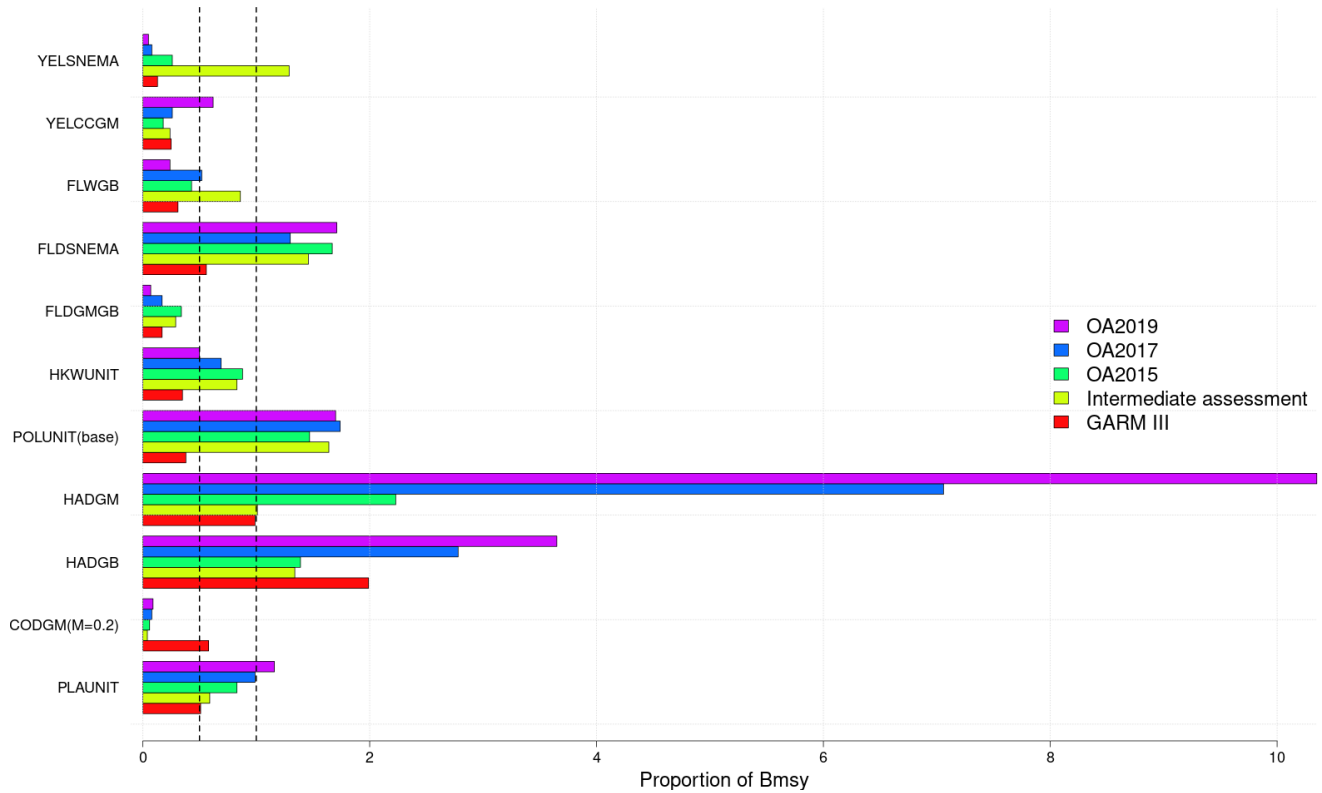

Figure 2: Changes in the ratio of stock biomass to B_{MSY} proxy from 2007 (GARM III) to 2019 (OA 2019) for selected Northeast Multispecies Fishery Management Plan groundfish stocks. The results from the assessment prior to the OA 2015 assessment are shown for each stock to provide 'Intermediate' values. Stocks that are overfished stocks are those where the $\frac{B_{terminal}}{B_{MSYproxy}}$ ratio is less than 0.5. *Notes: (1) Because Georges Bank cod, halibut and witch flounder do not currently have quantitative biomass reference points, they were not included in this figure. (2) The plotted value for FLDGMGB (Northern windowpane flounder) in OA2019 reflects the BRPs from the 2019 assessment. The peer review panel recommended using the BRPs from the 2017 assessment for status determination (see Table 6 legend).*

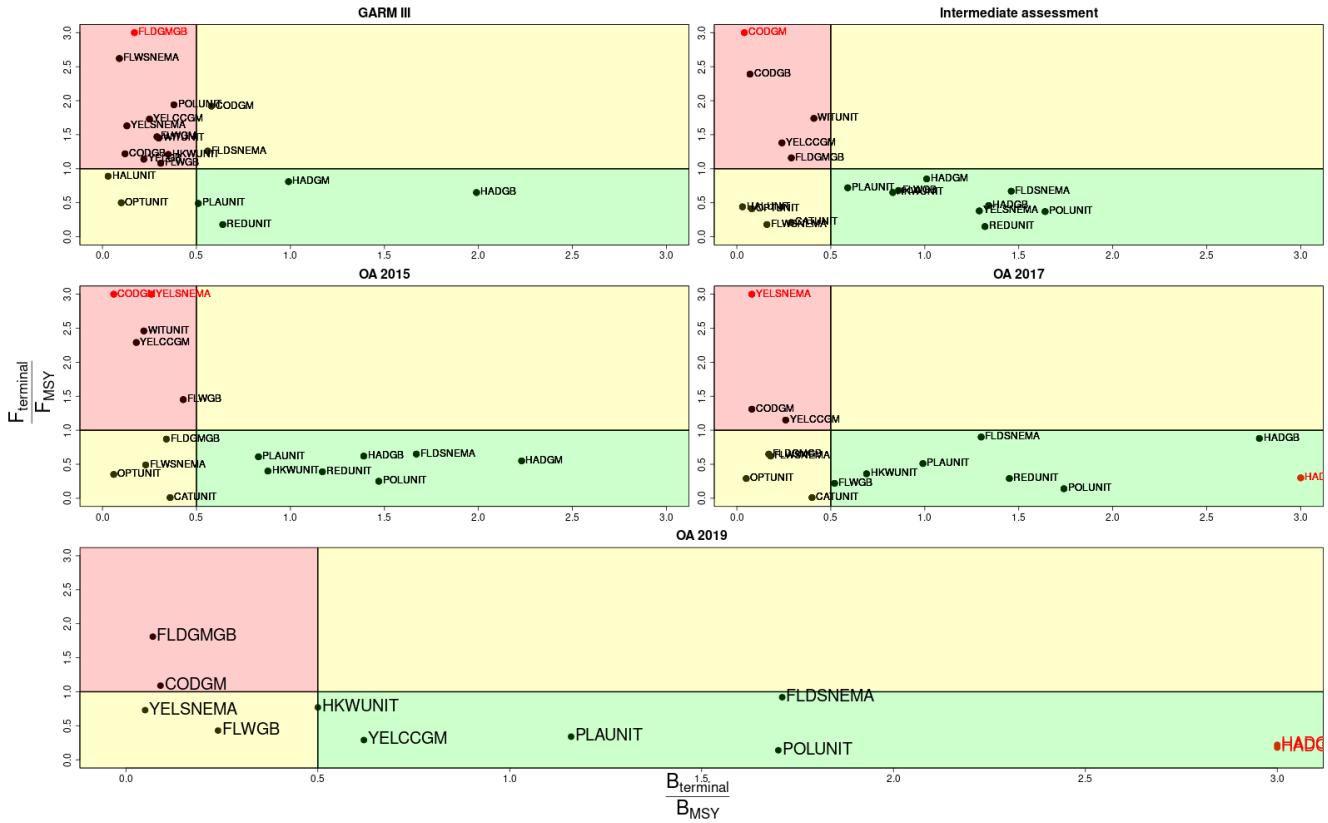

Figure 3: Status of the Northeast Multispecies Fishery Management Plan groundfish stocks in 2007 (GARM III) and 2019 (OA 2019) with respect to the F_{MSY} and B_{MSY} proxies. The 'Intermediate assessment' represents the last stock assessment conducted prior to the OA 2015 assessment (year varies by stock). Stocks on which overfishing is occurring are those where the $\frac{F_{terminal}}{F_{MSYproxy}}$ ratio is greater than 1 and overfished stocks are those where the $\frac{B_{terminal}}{B_{MSYproxy}}$ ratio is less than 0.5. The scale of the axes was restricted to allow for comparison between years. Stocks with red markers and text exceeded these limits (values are available in Table 6). *Notes: (1) Only stocks with known reference points are included in this figure. (2) The plotted value for FLDGMGB (Northern windowpane flounder) in OA2019 reflects the BRPs from the 2019 assessment. The peer review panel recommended using the BRPs from the 2017 assessment for status determination (see Table 6 legend).*

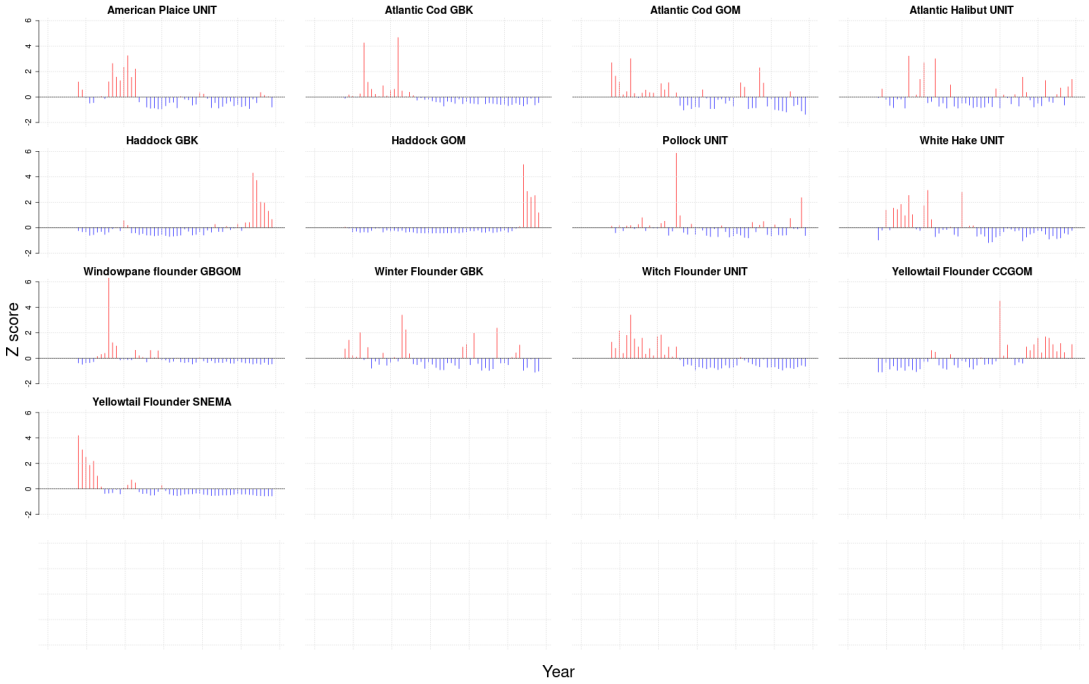

Figure 4: NEFSC spring bottom trawl survey index standardized anomalies (Z-score) for the Northeast Multispecies Fishery Management Plan groundfish stocks from 1968 to 2019. *Note that both the Georges Bank/Gulf of Maine and Southern New England/Mid-Atlantic windowpane flounder stocks are not included since the spring survey is uninformative as an index of abundance and not used in the stock assessment.*

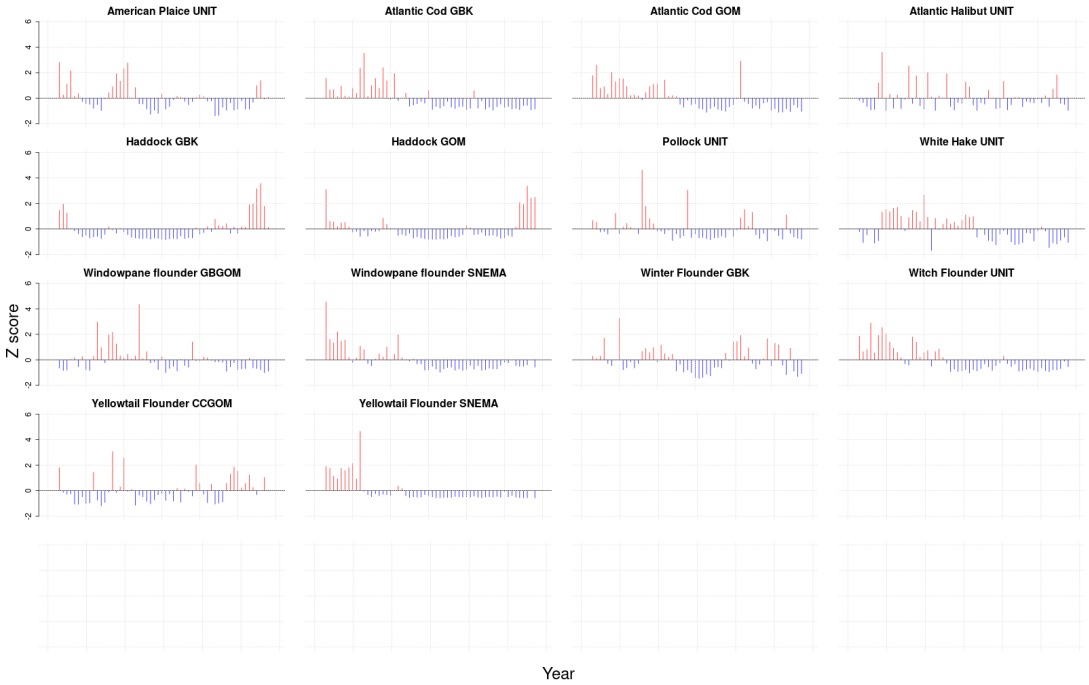

Figure 5: NEFSC fall bottom trawl survey index standardized anomalies (Z-score) for the Northeast Multispecies Fishery Management Plan groundfish stocks from 1963 to 2018. *Note that ocean pout is not included since the fall survey is uninformative as an index of abundance and not used in the stock assessment.*

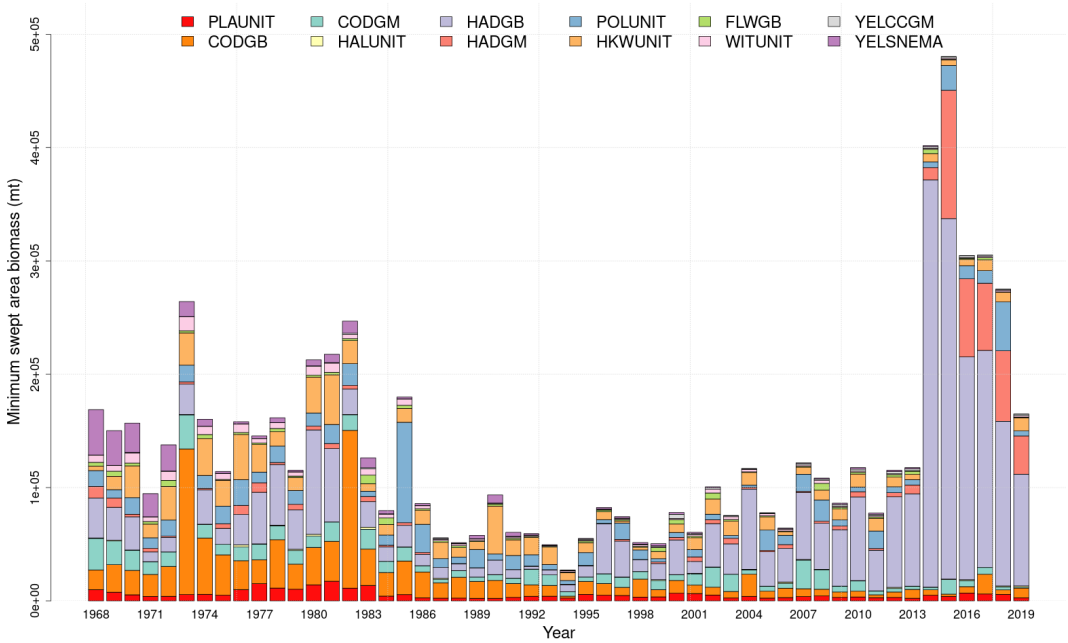

Figure 6: NEFSC spring bottom trawl survey minimum swept area biomass (mt) for the Northeast Multispecies Fishery Management Plan groundfish stocks from 1968 to 2019, by stock. Minimum swept area estimates assume a trawl swept area of 0.0112 nm^2) (0.0384 km^2) based on the wing spread of the trawl net. *Note that both the Georges Bank/Gulf of Maine and Southern New England/Mid-Atlantic windowpane flounder stocks are not included since the spring survey is uninformative as an index of abundance and not used in the stock assessment.*

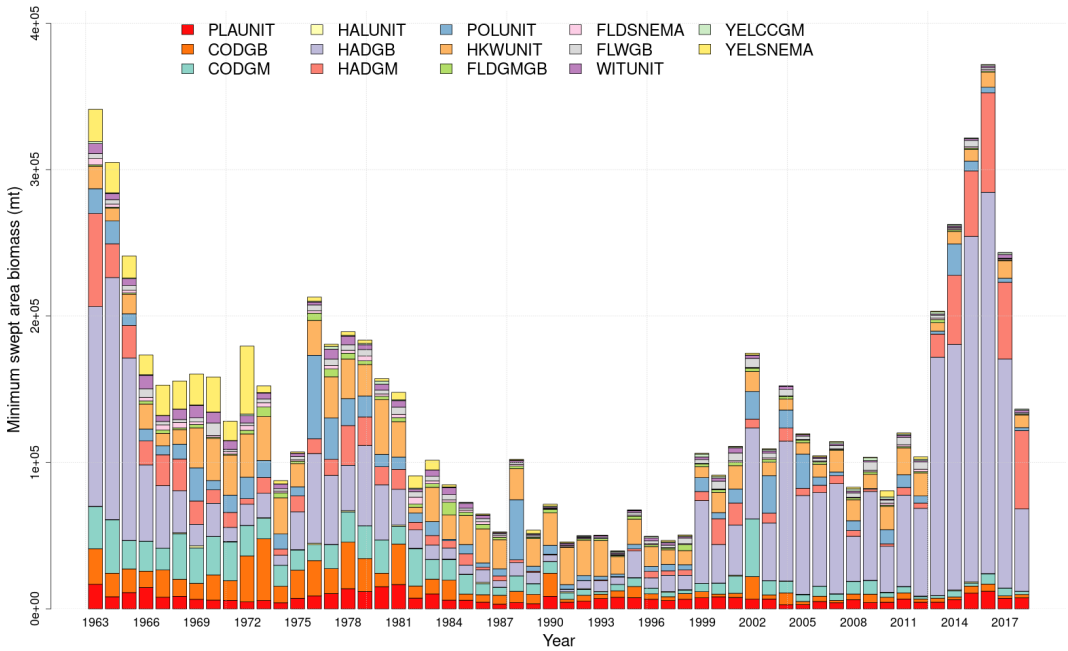

Figure 7: NEFSC fall bottom trawl survey minimum swept area biomass (mt) for for the Northeast Multispecies Fishery Management Plan groundfish stocks from 1963 to 2018, by stock. Minimum swept area estimates assume a trawl swept area of 0.0112 nm^2 (0.0384 km^2) based on the wing spread of the trawl net. *Note that ocean pout is not included since the fall survey is uninformative as an index of abundance and not used in the stock assessment.*

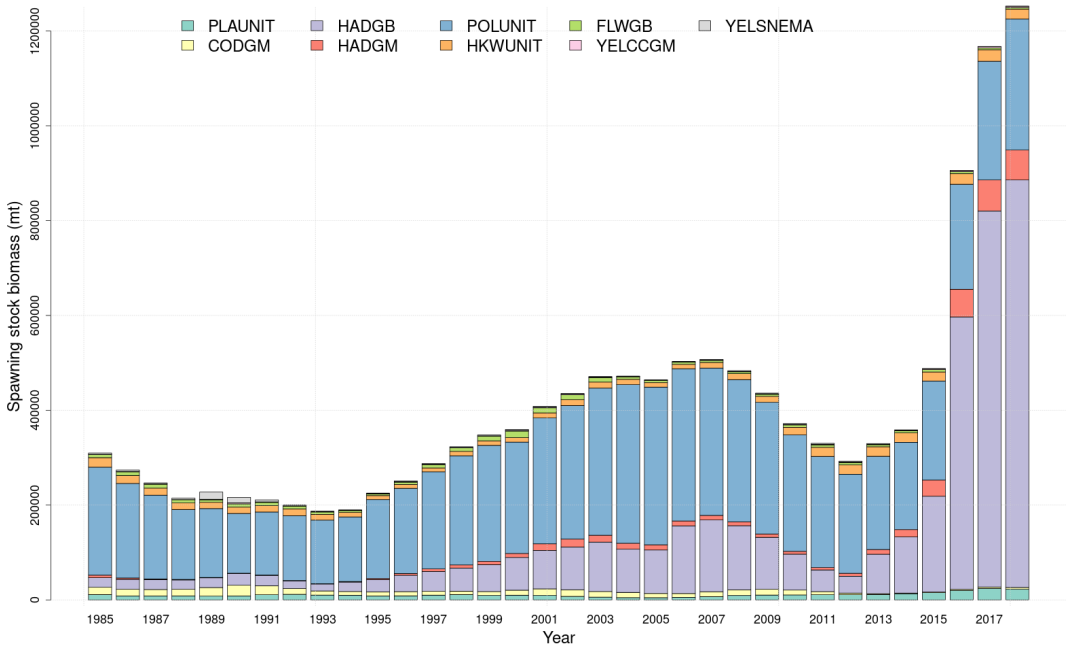

Figure 8: Model-based spawning stock biomass estimates for 11 groundfish stocks, 1985-2018 based on the Operational Assessments in 2019. Models without model-based biomass estimates are excluded.

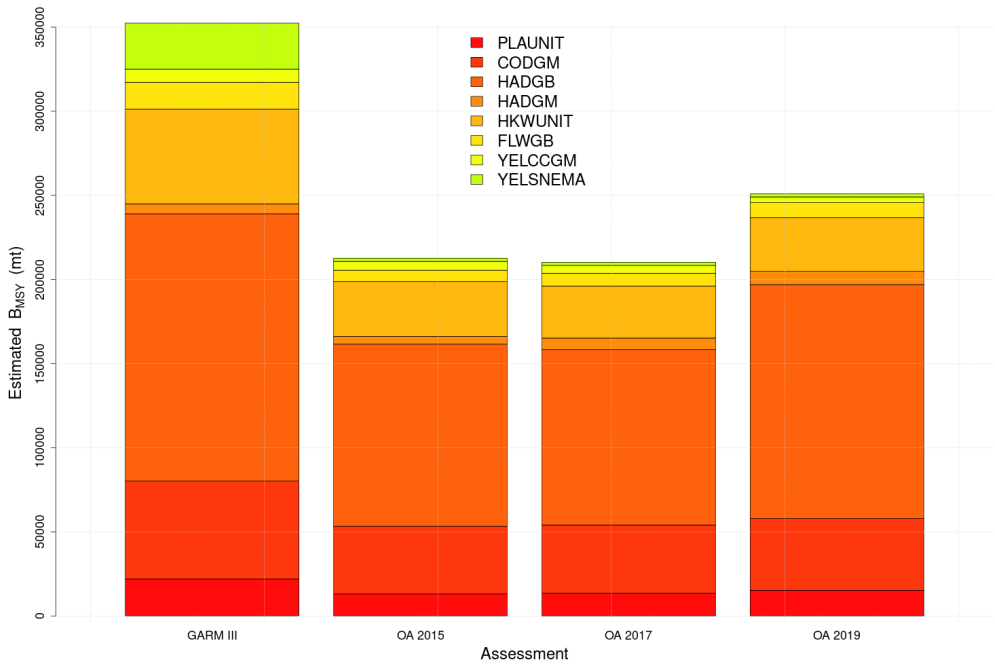

Figure 9: Sum of B_{MSY} estimates for nine stocks which had B_{MSY} estimates in 2008, 2015, 2017, and 2019 assessments. Pollock is not included since biomass targets were not established until 2010 at SARC 50. B_{MSY} estimates for Gulf of Maine winter flounder, witch flounder and Georges Bank yellowtail flounder are not available as both stock assessments are based on swept area expansions. The assessment model for Georges Bank cod was not accepted for catch advice in 2015 and is currently based on smoothed survey estimates.

2 Gulf of Maine Atlantic cod

Charles Perretti

This assessment of the Gulf of Maine Atlantic cod (Gadus morhua) stock is an operational assessment of the existing benchmark assessment (NEFSC 2013). This stock was most recently assessed in 2017 (NEFSC 2017). This assessment updates commercial and recreational fishery catch data, research survey indices of abundance, and the analytical ASAP assessment models through 2018. Additionally, stock projections have been updated through 2022. In what follows, there are two population assessment models brought forward from the most recent benchmark assessment (2012), the M=0.2 (natural mortality = 0.2) and the M-ramp (M ramps from 0.2 to 0.4) assessment models (see NEFSC 2013 for a full description of the model formulations).

State of Stock: Based on this updated assessment, the stock status for the Gulf of Maine Atlantic cod (*Gadus morhua*) stock is overfished and overfishing is occurring (Figures 10-11). Retrospective adjustments were not made to the model results (see Special Comments section of this report). Spawning stock biomass (SSB) in 2018 was estimated to be 3,752 (mt) under the M=0.2 model and 3,838 (mt) under the M-ramp model scenario (Table 9) which is 9% and 6% (respectively) of the biomass target, SSB_{MSY} *proxy* (42,692 (mt) and 63,867 (mt); Figure 10). The 2018 fully selected fishing mortality was estimated to be 0.188 and 0.198 which is 109% and 113% of the F_{MSY} *proxy*($F_{40\%}$; 0.173 and 0.175; Figure 11).

Table 9: Catch and status table for Gulf of Maine Atlantic cod. All weights are in (mt), recruitment is in (000s), and F_{Full} is the fishing mortality on fully selected ages.

	2011	2012	2013	2014	2015	2016	2017	2018
			Data					
Recreational discards	307	103	195	151	168	334	610	326
Recreational landings	2,999	1,245	1,524	796	11	187	170	12
Commercial discards	103	97	54	27	14	8	16	17
Commercial landings	4,598	2,759	951	832	227	320	376	398
Catch for Assessment	8,007	4,204	2,723	1,806	420	850	1,171	753
			Model Results (M=0.2)					
Spawning Stock Biomass	6723	3524	1874	1263	1439	2258	3051	3752
F_{Full}	1.504	1.69	2.178	2.224	0.37	0.459	0.419	0.188
Recruits $age1$	1645	1682	788	2702	1184	758	1845	2767
			Model Results (M-ramp)					
Spawning Stock Biomass	8009	4221	2361	1809	2164	3023	3593	3838
F_{Full}	1.308	1.482	1.859	1.669	0.27	0.374	0.379	0.198
Recruits $age1$	3123	3451	1712	5727	2311	1355	3062	4261

Table 10: Comparison of reference points estimated in an earlier assessment and from the current assessment update. The overfishing threshold is the F_{MSY} *proxy* ($F_{40\%}$). The biomass target, (SSB_{MSY} *proxy*) was based on long-term stochastic projections of fishing at the F_{MSY} *proxy*. Median recruitment reflects the median estimated age-1 recruitment from 1982 - 2016. Intervals shown reflect the 5th and 95th percentiles.

	2017 M=0.2	2017 M-ramp	M=0.2	M-ramp
F_{MSY}	0.174	0.177	0.173	0.175
SSB_{MSY} (mt)	40,604 (27,631 - 58,553)	59,714 (44,732 - 77,611)	42,692 (27,916 - 62,785)	63,867 (46,144 - 84,098)
MSY (mt)	7,049 (4,699 - 10,380)	10,502 (7,734 - 13,822)	7,580 (4,853 - 11,366)	11,420 (8,149 - 15,268)
Median recruits age-1) (000s)	4,377 (1,161 - 14,434)	8,464 (2,353 - 15,934)	4,677 (1,064 - 16,392)	9,249 (2,129 - 18,031)
Overfishing	Yes	Yes	Yes	Yes
Overfished	Yes	Yes	Yes	Yes

Projections: Short term projections of median total fishery yield and spawning stock biomass for Gulf of Maine Atlantic cod were conducted based on a harvest scenario of fishing at the FMSY proxy between 2020 and 2022. Catch in 2019 was estimated at 710 mt. Recruitment was sampled from a cumulative distribution function derived from ASAP estimated age-1 recruitment between 1982 and 2016. The projection recruitment model declines linearly to zero when SSB is below 6.3 kmt under the M=0.2 model and 7.9 kmt under the M-ramp model. The 2019 age-1 recruitment was estimated from the geometric mean of the 2014-2018 ASAP recruitment estimates. No retrospective adjustments were applied in the projections as the retrospective patterns are similar to the 2017 update for which no retrospective adjustments were made. Assumed weights are based on an average of the most recent three years. For the M-ramp model, projections are shown under two assumptions of short-term natural mortality: M=0.2 and M=0.4.

Table 11: Short term projections of total fishery catch and spawning stock biomass for Gulf of Maine Atlantic cod based on a harvest scenario of fishing at the F_{MSY} *proxy* ($F_{40\%}$) between 2020 and 2022. Catch in 2019 has been estimated at 710 (mt).

Year	Catch (mt)	SSB (mt)	F_{Full}	Catch (mt)	SSB (mt)	F_{Full}	Catch (mt)	SSB (mt)	F_{Full}
	M=0.2			*M-ramp(M=0.2)*			*M-ramp(M=0.4)*		
2019	710	4,732	0.148	710	4,326	0.171	710	4,103	0.189

Year	Catch (mt)	SSB (mt)	F_{Full}	Catch (mt)	SSB (mt)	F_{Full}	Catch (mt)	SSB (mt)	F_{Full}
	M=0.2			*M-ramp(M=0.2)*			*M-ramp(M=0.4)*		
2020	1,102	6,276	0.173	1,027	6,112	0.175	758	4,719	0.175
2021	1,440	8,064	0.173	1,469	8,547	0.175	893	5,461	0.175
2022	1,813	10,673	0.173	1,995	11,927	0.175	1,010	6,415	0.175

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  An important source of uncertainty is the estimate of natural mortality. Past investigations into changes in natural mortality over time have been inconclusive (NEFSC 2013). Different assumptions about natural mortality affect the scale of the biomass, recruitment, and fishing mortality estimates, though terminal estimates (2018) of biomass, fishing mortality and recruitment are similar under both models. Other areas of uncertainty include the retrospective error in the M=0.2 model, stock structure, and the veracity of fishery catch data.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}).

  The M=0.2 model has a major retrospective pattern (7-year Mohn's rho SSB=0.52, F=-0.29) and the M-ramp model has a minor retrospective pattern (7-year Mohn's rho SSB=0.29, F=-0.16). The 7-year Mohn's rho values from the current assessment are similar to those from the 2017 assessment (M=0.2: SSB=0.53, F=-0.31; M-ramp: SSB=0.30, F=-0.17) where the M=0.2 model had a major retrospective pattern and the M-ramp model had a minor pattern. No retrospective adjustments have been applied to the terminal model results or in the base catch projections following the recommendations of the SARC 55 (NEFSC 2013) and 2014 assessment review panels (Palmer 2014). The 2017 assessment review panel (NEFSC 2017) supported this decision, noting that the most recent retrospective 'peel' suggested that an adjustment using the 7-year average may not be appropriate. However, the 2017 review panel highlighted the retrospective error in the M=0.2 model as a source of uncertainty. Should the retrospective patterns continue then the models may have overestimated spawning stock size and underestimated fishing mortality.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Gulf of Maine Atlantic cod are reasonably well determined, though the projected biomasses from the last assessment did not fall within the confidence bounds of the biomass estimated in the current assessment. Multiple factors likely contributed to this including overstimation of the initial stock size and underestimation of F in the projection bridge year (2017). This stock is not on target to rebuild by 2024.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Recreational catch estimates were re-estimated in this update by using the re-calibrated MRIP data. In general, inclusion of the re-calibrated data resulted in an increase in SSB, F, and recruitment. Prior to 2004, there is no length information for recreational releases, and there are several years with either limited or no length information for recreational harvest. However, proportions-at-age are similar between the pre- and post-calibrated data. Therefore, recreational catch-at-age prior to 2004 was calculated by applying the historical

proportions-at-age to the new total catch numbers estimated in the re-calibrated MRIP procedure.

Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

The experimental catchability data were not applicable to the Gulf of Maine Atlantic cod stock assessment because the catchability experiments were focused on flatfish species.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  There has been no change in stock status since the 2017 update assessment.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The Gulf of Maine Atlantic cod shows a truncated size and age structure, consistent with a population experiencing high mortality. Additionally, there are only limited signs of incoming recruitment, continued low survey indices, and the current spatial distribution of the stock is considerably less than its historical range within the Gulf of Maine.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Gulf of Maine Atlantic cod assessment could be improved with additional studies on natural mortality, stock structure, a characterization of the overall uncertainty and possible biases in the fishery catch estimates, and research into potential causes of low stock productivity (i.e., low recruitment).

- Are there other important issues?

  When setting catch advice, careful attention should be given to the retrospective error present in both models, particularly given the poor performance of previous stock projections.

2.1 Reviewer Comments: Gulf of Maine Atlantic cod

2.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference. Two models (M=0.2, M-ramp) are updated as per the previous benchmark to help characterize the overall uncertainty in natural mortality. Results are generally similar from the two models. The M=0.2 model exhibited a major retrospective pattern, while the M-ramp had minor retrospective pattern. Recent low recruitment compromises the rebuilding potential of the stock.

2.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The updated model run incorporates two additional years of data (2017-18) into the assessment time series and incorporates re-calibrated MRIP data for the recreational fishery. MRIP landings estimates are much higher than anticipated given the restrictive measures for recreational season (i.e. short seasons and complete closures; 1 fish limit or no possession). Recreational discard increases were expected, but landings were much higher than expected, especially in recent years.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used. Note that fishery selectivity differences exist between M-ramp and M-0.2. The stock-recruitment curve uses a hockey stick formulation. The Panel supports the use of M=0.2 in the calculation of reference points for the M-ramp model based on the lack of time under the new M in the M-ramp model for cod to have adjusted their biological characteristics to compute reference points under the higher M value and because the additional natural mortality may in fact be missing catch in the M-ramp model formulation (see Legault and Palmer 2016 CJFAS 73: 349-357 for more details about what to do when the natural mortality rate changes within an assessment).

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is considered overfished and overfishing is occurring

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

The survey indices continue to show low biomass and truncated age structure. There are no signs of incoming recruitment except for the Massachusetts Division of Marine Fisheries survey showing a high value of age zero fish in 2019. The NEFSC spring survey in 2019 did not indicate large recruitment, and age zero recruitment signals in the past have not carried through to older ages.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at FMSY or at an FMSY proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

The SSBMSY proxies were updated and are based on 100 year projections run at the FMSY proxy. The projection model samples from a distribution of recruitment estimates for 1982-2016. When SSB is below a hinge point recruitment is assumed to decline linearly to zero. M=0.2: 6,300 mt, M-ramp: 7,900 mt. A set of four short term projections were considered including the two benchmark methods and two reasonable alternatives: M=0.2 model with and without rho-adjustments and the M-ramp model with M=0.2 and M=0.4. The Panel agreed there is no basis to support the M-ramp model with M=0.2 projection because there is no reason to believe the natural mortality rate would immediately return to 0.2 and doing so artificially increases the rebuilding rate of the population. The Panel did not reach agreement over which M=0.2 model projection is favored: not rho-adjusting is the benchmark formulation and allows the M-ramp model to account for uncertainty associated with the retrospective pattern while rho-adjusting follows the standard procedure for dealing with retrospective patterns.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

Regarding the use of the M=0.2 vs. Mramp model some consideration should be given to evaluating whether we need to continue with both or come up with some alternative approach.

The Gulf of Maine longline survey could be considered for inclusion in the next assessment.

Updating fishery CPUE would help to document fishery perceptions.

Consideration of a new approach for providing catch advice for stocks that are at extremely low biomass should be considered.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2013. 55th Northeast Regional Stock Assessment Workshop (55th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 13-11; 41 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD13-11

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD15-24

Palmer MC. 2014. 2014 Assessment update report of the Gulf of Maine Atlantic cod stock. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-14; 119 p. Available from: National Marine Fisheries Service,166 Water Street, Woods Hole, MA 02543-1026. CRD14-14

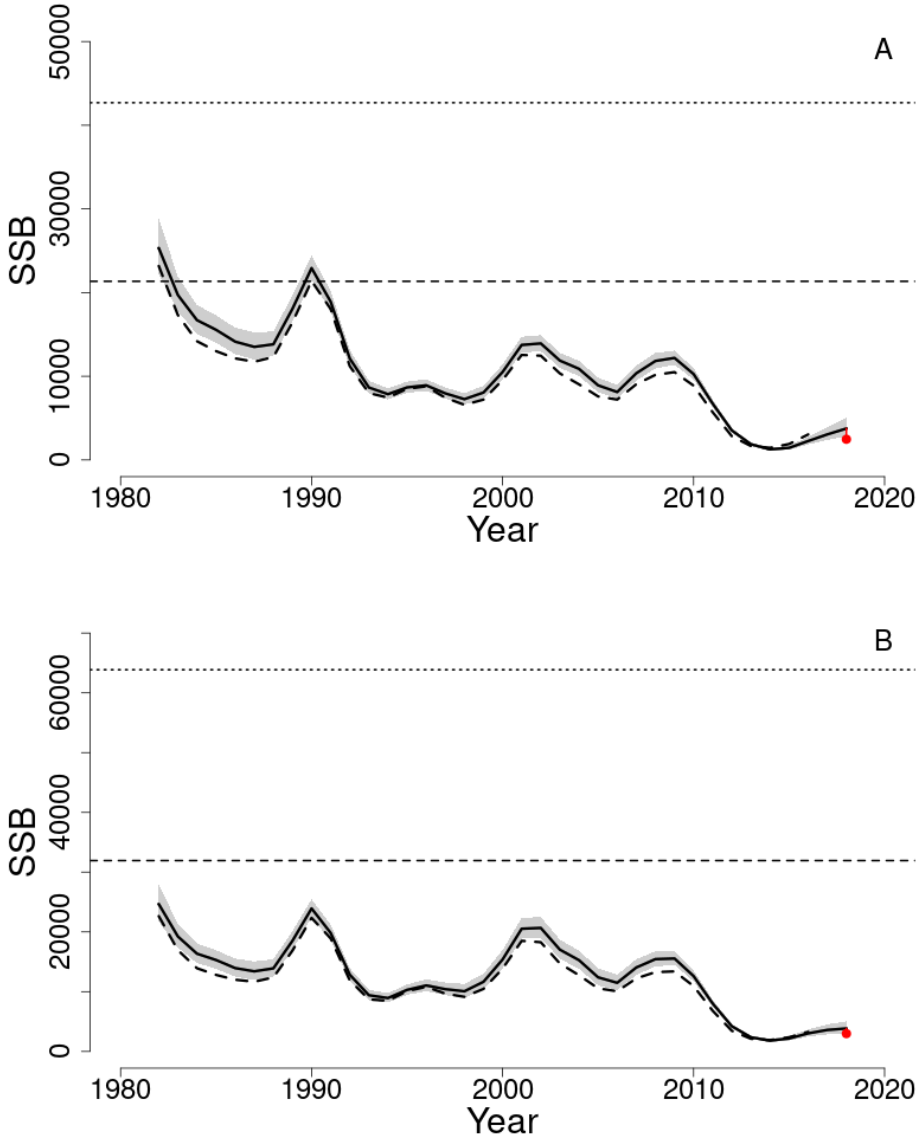

Figure 10: Estimated trends in the spawning stock biomass (SSB) of Gulf of Maine Atlantic cod between 1982 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY}; horizontal dashed line) as well as SSB_{Target} SSB_{MSY}; horizontal dotted line) based on the 2019 M=0.2 (A) and M-ramp (B) assessment models. The 90% lognormal confidence intervals are shown. The red dot indicates the rho-adjusted SSB values that would have resulted had a retrospective adjusment been made to either model (see Special Comments section).

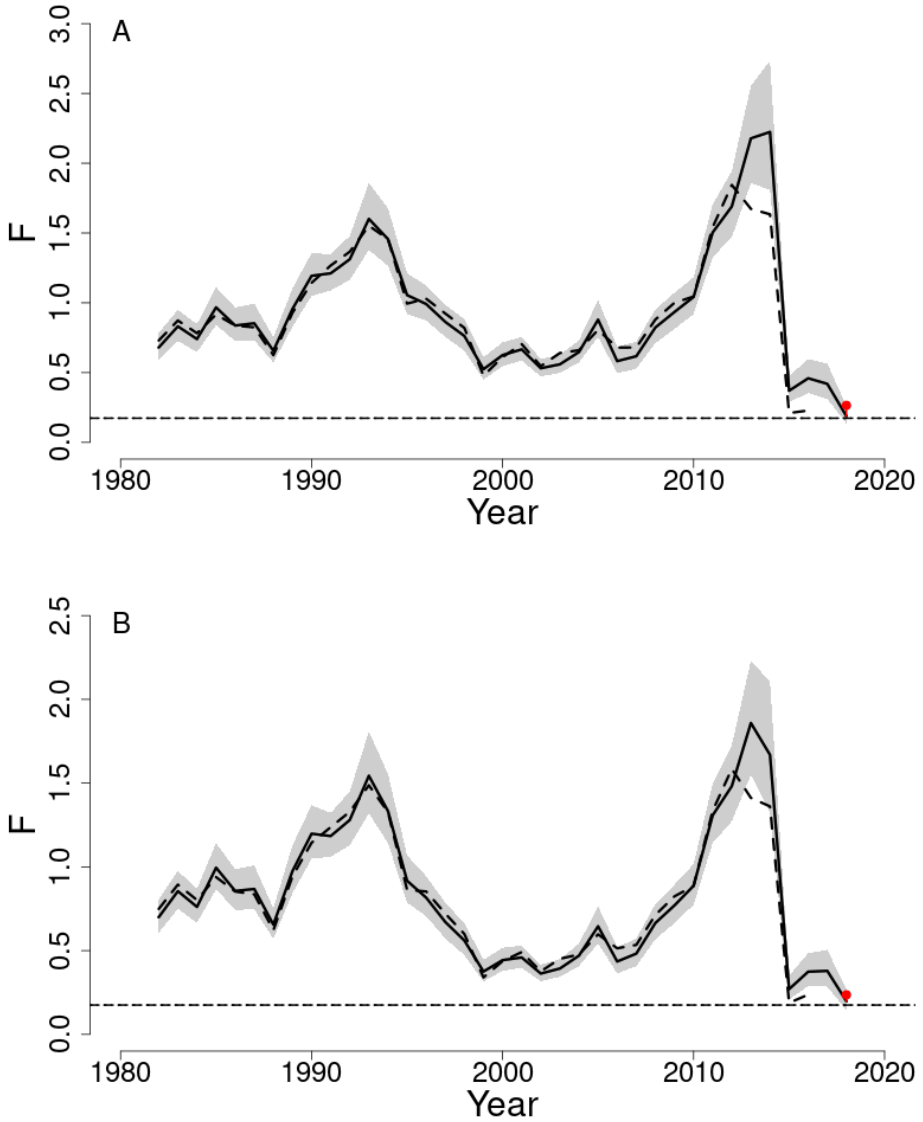

Figure 11: Estimated trends in the fully selected fishing mortality (F) of Gulf of Maine Atlantic cod between 1982 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (0.173 (M=0.2), 0.175 (M-ramp); dashed line) based on the 2019 M=0.2 (A) and M-ramp (B) assessment models. The 90% lognormal confidence intervals are shown. The red dot indicates the rho-adjusted F values that would have resulted had a retrospective adjusment been made to either model (see Special Comments section).

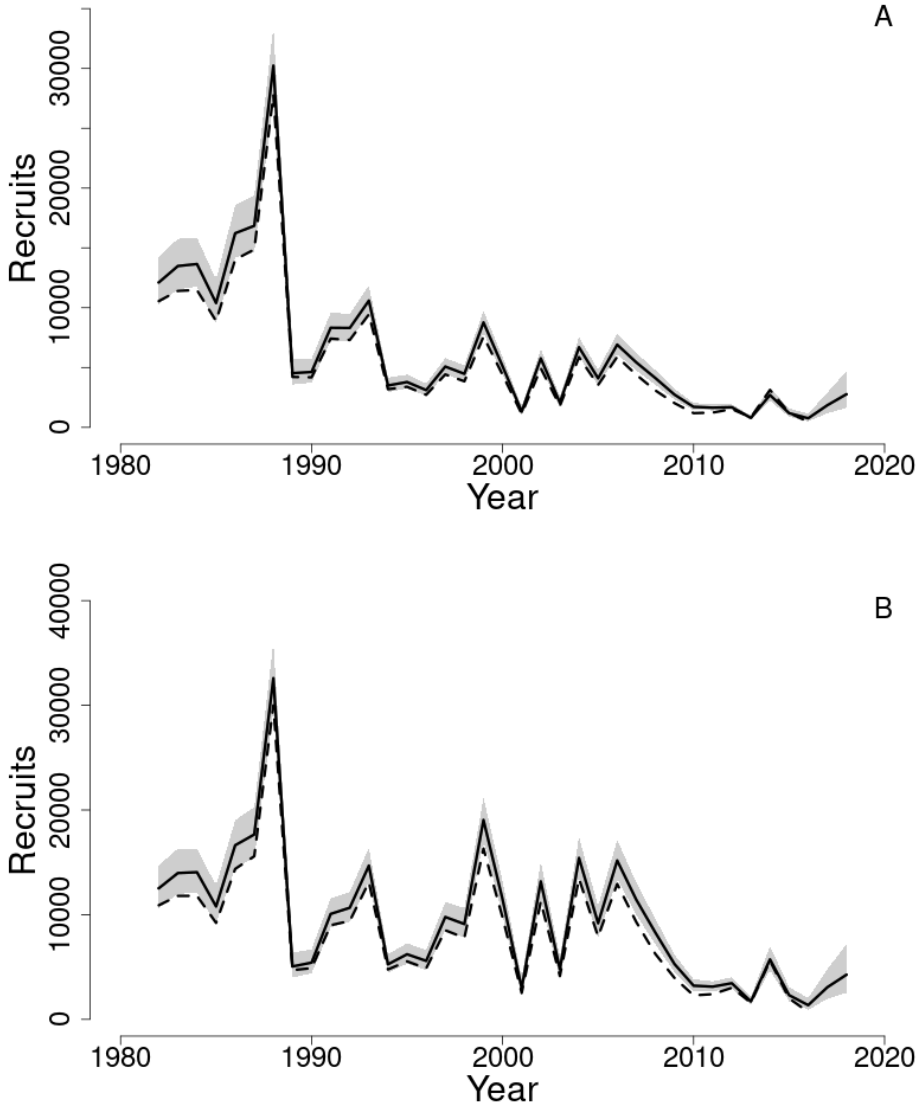

Figure 12: Estimated trends in age-1 recruitment (000s) of Gulf of Maine Atlantic cod between 1982 and 2018 from the current (solid line) and previous (dashed line) M=0.2 (A) and M-ramp (B) assessment models. The 90% lognormal confidence intervals are shown.

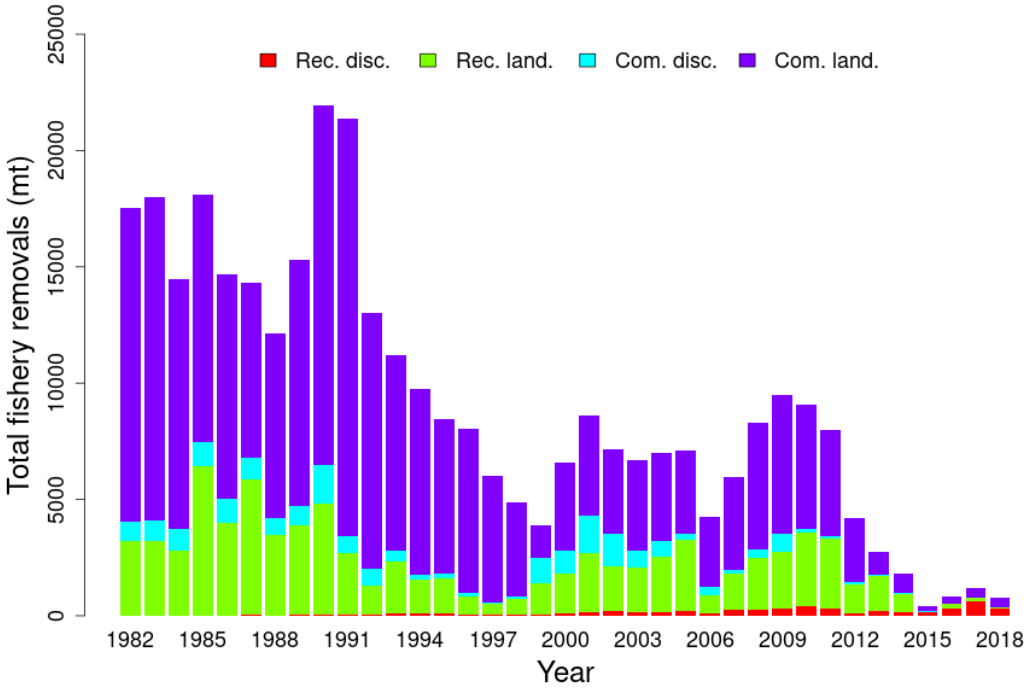

Figure 13: Total catch of Gulf of Maine Atlantic cod between 1982 and 2018 by fleet (commercial and recreational) and disposition (landings and discards).

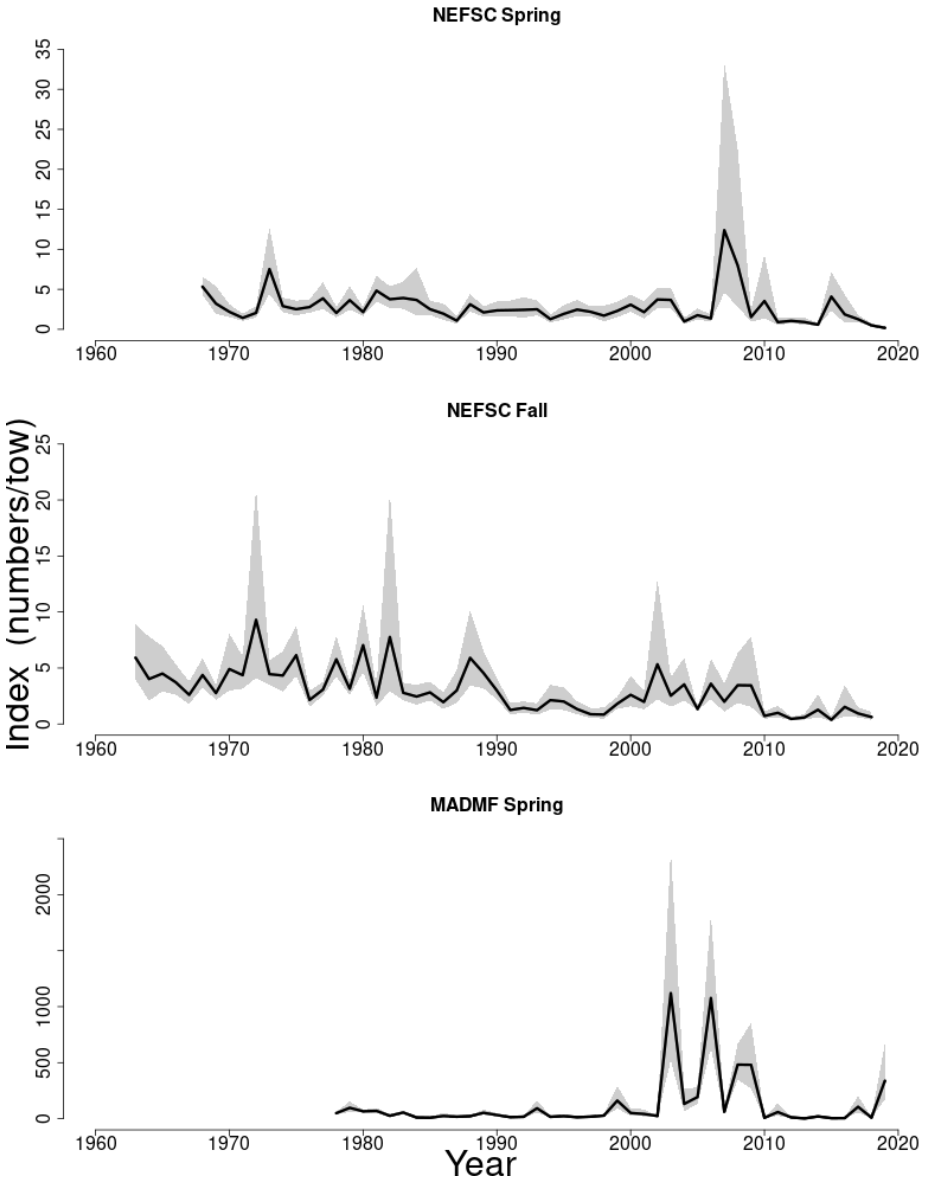

Figure 14: Indices of biomass for the Gulf of Maine Atlantic cod between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys and Massachusetts Division of Marine Fisheries (MADMF) spring bottom trawl survey. The 90% lognormal confidence intervals are shown.

3 Georges Bank Atlantic cod

SAW Process Chairman/Editor's Note: *As described in the Executive Summary, Section 1.1, the GB cod assessment was affected by an MRIP data issue as well as having used an incorrect assumed recreational discard mortality rate of 100%, rather than an assumed rate of 30%. In January 2020 the GB cod assessment was corrected to account for both issues, and is now included as an appendix to Chapter 3 "Georges Bank Atlantic cod". The appendix is appropriate for use by fishery managers. The original chapter that was peer reviewed in September 2019 is included here for documentation.*

Chris Legault

This assessment of the Georges Bank Atlantic cod (Gadus morhua) stock is an operational assessment of the existing 2017 operational update assessment (NEFSC 2017). Based on the previous assessment the stock status could not be quantitatively determined but was qualitatively determined to be overfished based on poor stock condition, while overfishing status remained unknown (see Table 13 Legend). This 2019 assessment updates commercial fishery catch data through 2018 (Table 12, Figure 17) and updates research survey indices of abundance and the PlanBsmooth assessment model through 2019 (Figure 18).

State of Stock: Based on this updated assessment, the Georges Bank Atlantic cod (*Gadus morhua*) stock status cannot be quantitatively determined due to a lack of biological reference points associated with the PlanBsmooth approach but is recommended to be overfished due to poor stock condition, while recommended overfishing status is unknown (Table 13). Retrospective adjustments were not made to the model results. The survey biomass in 2019 (the arithmetic average of the 2019 NEFSC spring and 2018 NEFSC fall surveys smoothed using a loess) was estimated to be 3.742 (kg/tow) (Figure 15). The 2018 relative exploitation rate (2018 catch divided by 2018 smoothed survey biomass) was estimated to be 0.119 (Figure 16).

Table 12: Catch and model results for Georges Bank Atlantic cod. Catch weights are in (mt), Biomass is the average survey biomass in (kg/tow) smoothed using a loess, and Rel. Exploit. Rate is the relative exploitation rate (catch/smoothed survey). Model results are from the PlanBsmooth assessment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
					Data					
Commercial landings	2,999	2,688	3,387	2,007	1,312	1,514	1,300	1,109	464	574
Commercial discards	385	253	122	120	83	19	31	33	20	13
Recreational landings	142	195	142	81	7	257	486	1,075	785	66
Recreational discards	9	27	25	3	2	19	71	32	25	6
CA landings	1,003	748	702	395	384	430	472	428	474	510
CA discards	206	94	43	75	39	28	20	12	14	7
Catch for Assessment	4,744	4,005	4,421	2,681	1,828	2,267	2,380	2,690	1,782	1,176
					Model Results					
Biomass	3.227	3.107	3.13	3.175	3.022	2.428	2.919	4	4.27	4.256
Rel. Exploit. Rate	0.633	0.555	0.609	0.364	0.261	0.402	0.351	0.29	0.18	0.119

Table 13: Comparison of reference points estimated in the previous assessment and from the current assessment update. Note: based on NOAA's policy, the Agency decided after the 2015 assessment that the stock status would remain as overfishing occurring and overfished based on an earlier benchmark assessment.

	2017	2019
F_{MSY} *proxy*	NA	NA
SSB_{MSY} (kg/tow)	NA	NA
MSY (mt)	NA	NA
Overfishing	Unknown	Unknown
Overfished	Yes	Yes

Projections: Short term projections cannot be computed using the PlanBsmooth approach. The PlanBsmooth approach estimates the rate of change in the recent three years of the smoothed survey biomass to be 0.936. This multiplier is applied to the average of the recent three years of catch (1,882 mt) to produce the catch advice for 2020 of 1,762 mt. The PlanBsmooth approach is fully described in NEFSC (2015) and available as an R package. A Shiny app demonstrating the performance of the PlanBsmooth approach is also available.

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The major source of uncertainty is the cause of the retrospective pattern that led to the analytical assessment of this stock not being accepted during the 2015 operational update meeting.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}).

  No retrospective adjustment of spawning stock biomass or fishing mortality was required.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Georges Bank Atlantic cod are not computed. Catch advice is derived from applying an estimate of recent change in the smoothed survey biomass to the average of the recent three years of catch and thus is influenced by uncertainty in survey estimates. The smoothed survey biomass is decreasing, but without a biomass reference point it is not known if rebuilding is on schedule.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

The experimental catchability data were not applicable to the Georges Bank Atlantic cod stock assessment because the catchability experiments were focused on flatfish species. Additionally, the restrictor cable used during the experiment may have impacted catchability of cod in both nets.

The Marine Recreational Information Program data were changed from the previous assessment. For example, the 2017 assessment used years 2014-2016 to determine the average catch for use in the PlanBsmooth approach for catch advice. The old recreational catch (landings plus dead discards) for these years were 90, 140, and 399 mt, respectively. The new recreational catch for these years are 276, 557, and 1,107 mt. This results in the average total catch (US commercial, US recreational, and Canadian) increasing from 2,008 mt to 2,445 mt, a 22% increase. Thus, the catch advice in the 2017 assessment would have been 22% higher (3,710 mt instead of 3,047 mt) had the new MRIP data been used. Comparisons of the old and new recreational catch data are provided in the tables and figures files located in the data portal.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status for Georges Bank Atlantic cod remains overfished based on a qualitative evaluation of poor stock condition.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The Georges Bank Atlantic cod continues to show a truncated age structure. The most recent survey values remain below the mean of their time series. The 2013 year class was larger than recent year classes, but has not continued to be large as it ages and is below the average from the 1970s at every age in both surveys.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Georges Bank Atlantic cod assessment could be improved with additional studies on natural mortality, the potential for missing catch, and other possible sources of retrospective patterns in analytical assessments.

- Are there other important issues?

  The differences in modeling approaches between the full Georges Bank cod assessment (reported here) and the TRAC cod assessment of eastern Georges Bank (a portion of the whole bank) remain a potential problem.

3.1 Reviewer Comments: Georges Bank Atlantic cod

The Georges Bank Atlantic cod stock assessment was not reviewed by the 2019 Review Panel because it was determined to be a level 1 assessment at the AOP meeting in June of 2019 (Appendix 16.2), according to the stock assessment process adopted for this and future management track assessments (Appendix 16.3).

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. Northeast Fisheries Science Center Reference Document 17-17; 259 p. doi: 10.7289/V5/RD-NEFSC-17-17 CRD17-17

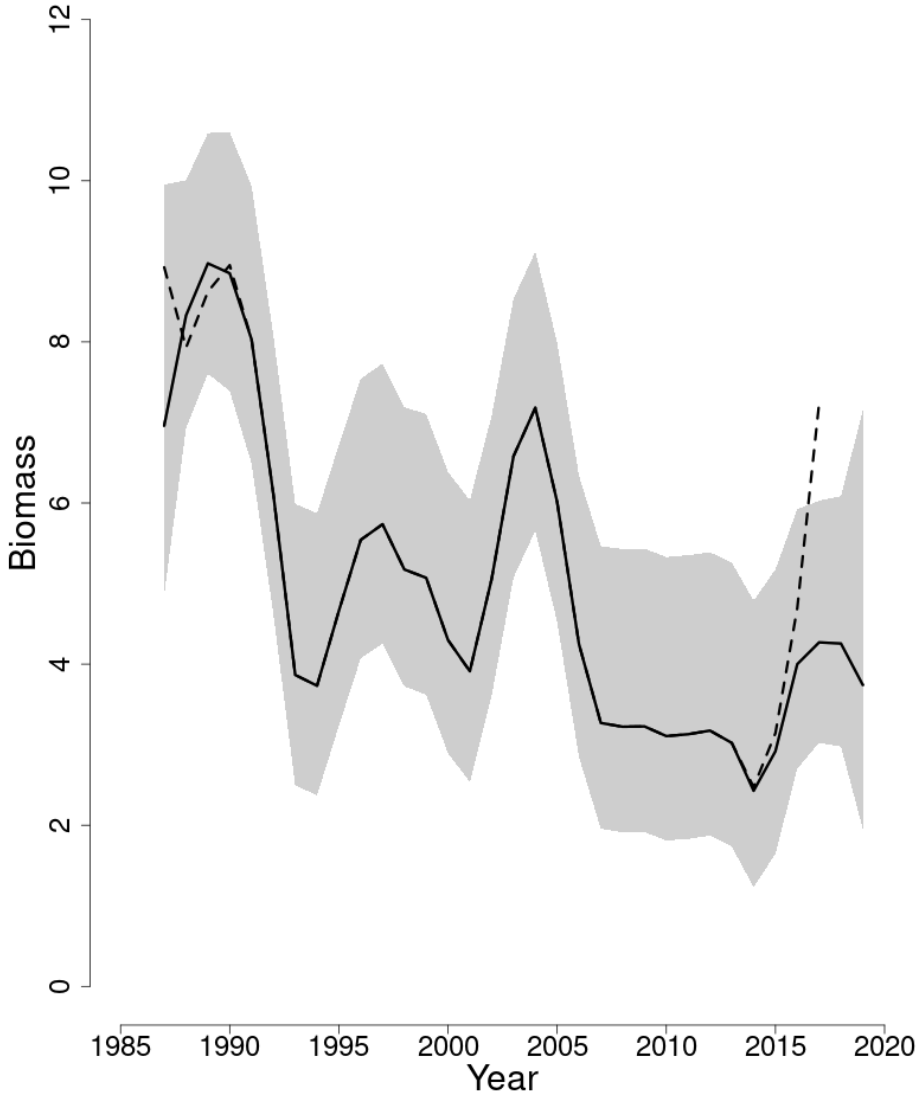

Figure 15: Trends in smoothed survey biomass (kg/tow) of Georges Bank Atlantic cod between 1987 and 2019 from the current (solid line) and previous (dashed line) assessment based on the 2019 assessment. The approximate 90% lognormal confidence intervals are shown.

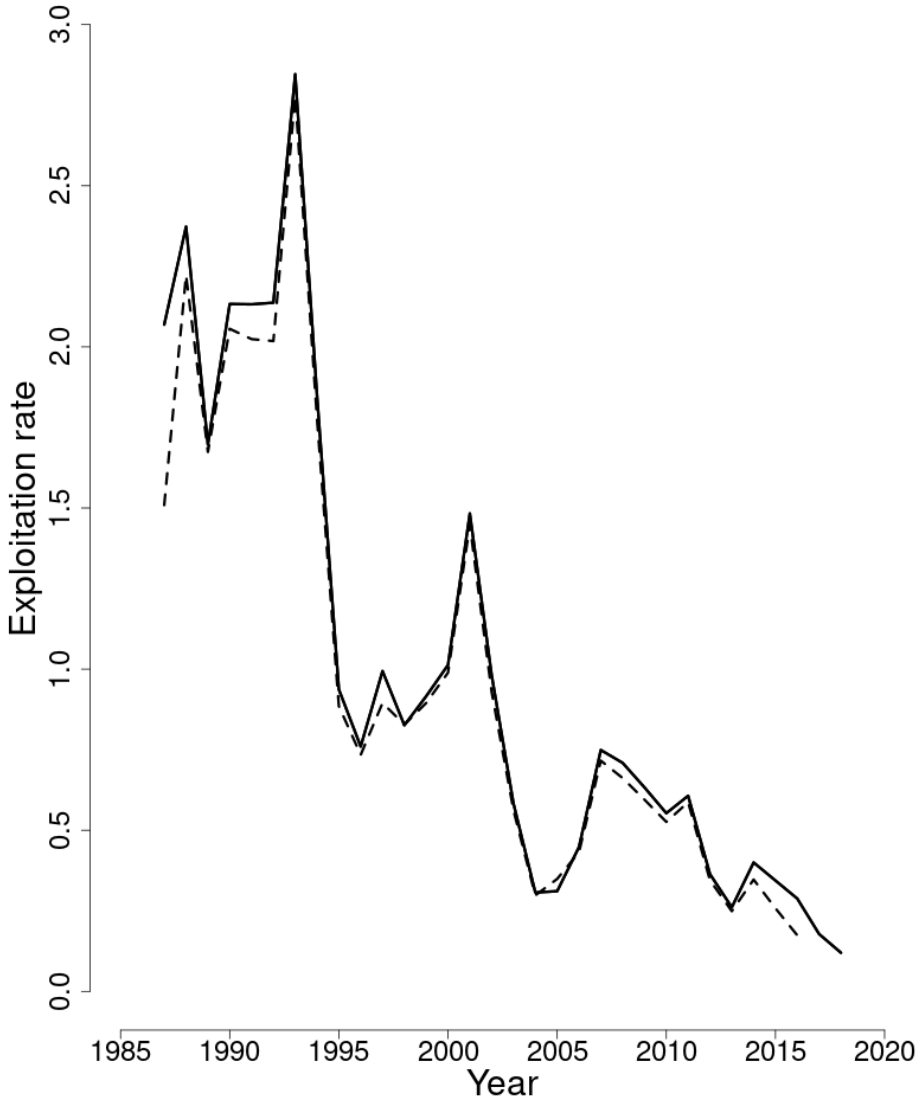

Figure 16: Trends in the relative exploitation rate (catch/smoothed survey) of Georges Bank Atlantic cod between 1987 and 2018 from the current (solid line) and previous (dashed line) assessment based on the 2019 assessment.

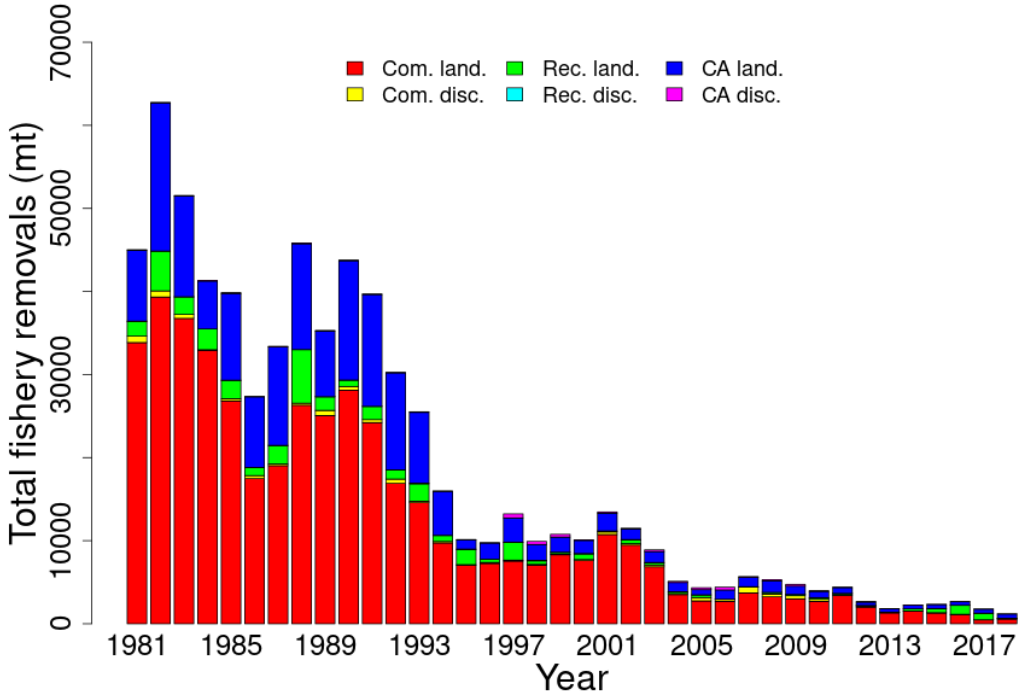

Figure 17: Total catch of Georges Bank Atlantic cod between 1981 and 2018 by fleet (US commercial, US recreational, or Canadian) and disposition (landings and discards).

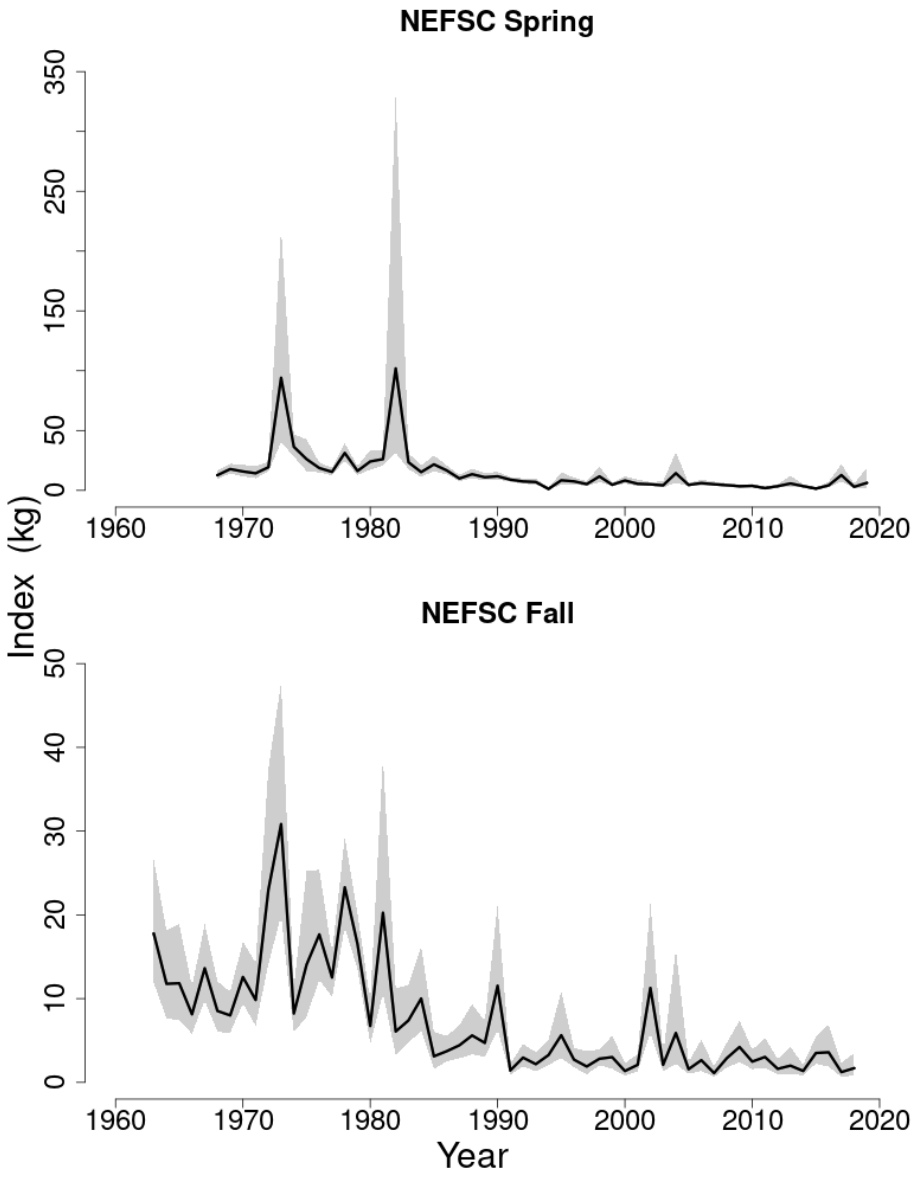

Figure 18: Indices of biomass for the Georges Bank Atlantic cod between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall trawl surveys. The approximate 90% lognormal confidence intervals are shown.

This marks the end of the original Georges Bank Atlantic cod chapter from September 2019, and the start of the appendix with the revisions to the Georges Bank Atlantic cod chapter.

3.2 Appendix: Georges Bank Atlantic cod

SAW Process Chairman/Editor's Note: *As described in the Executive Summary, Section 1.1, the GB cod assessment was affected by an MRIP data issue as well as having applied an incorrect assumed recreational discard mortality rate of 100%, rather than an assumed rate of 30%. This new appendix was added in January 2020. It correctly accounts for both issues and is suitable for use by fishery managers.*

Chris Legault

This assessment of the Georges Bank Atlantic cod (Gadus morhua) stock is an operational assessment of the existing 2017 operational update assessment (NEFSC 2017). Based on the previous assessment the stock status could not be quantitatively determined but was qualitatively determined to be overfished based on poor stock condition, while overfishing status remained unknown (see Table 15 Legend). This 2019 assessment updates commercial fishery catch data through 2018 (Table 14, Figure 21) and updates research survey indices of abundance and the PlanBsmooth assessment model through 2019 (Figure 22).

State of Stock: Based on this updated assessment, the Georges Bank Atlantic cod (*Gadus morhua*) stock status cannot be quantitatively determined due to a lack of biological reference points associated with the PlanBsmooth approach but is recommended to be overfished due to poor stock condition, while recommended overfishing status is unknown (Table 15). Retrospective adjustments were not made to the model results. The survey biomass in 2019 (the arithmetic average of the 2019 NEFSC spring and 2018 NEFSC fall surveys smoothed using a loess) was estimated to be 3.742 (kg/tow) (Figure 19). The 2018 relative exploitation rate (2018 catch divided by 2018 smoothed survey biomass) was estimated to be 0.12 (Figure 20).

Table 14: Catch and model results for Georges Bank Atlantic cod. Catch weights are in (mt), Biomass is the average survey biomass in (kg/tow) smoothed using a loess, and Rel. Exploit. Rate is the relative exploitation rate (catch/smoothed survey). Model results are from the PlanBsmooth assessment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
				Data						
Commercial landings	2,999	2,688	3,387	2,007	1,312	1,514	1,300	1,109	464	574
Commercial discards	385	253	122	120	83	19	31	33	20	13
Recreational landings	142	195	142	81	7	257	486	1,075	786	77
Recreational discards	3	8	8	1	1	5	21	10	8	2
CA landings	1,003	748	702	395	384	430	472	428	474	510
CA discards	206	94	43	75	39	28	20	12	14	7
Catch for Assessment	4,738	3,986	4,404	2,679	1,827	2,253	2,330	2,667	1,765	1,183
				Model Results						
Biomass	3.227	3.107	3.13	3.175	3.022	2.428	2.919	4	4.27	4.256
Rel. Exploit. Rate	0.633	0.553	0.607	0.364	0.261	0.4	0.344	0.288	0.178	0.12

Table 15: Comparison of reference points estimated in the previous assessment and from the current assessment update. Note: based on NOAA's policy, the Agency decided after the 2015 assessment that the stock status would remain as overfishing occurring and overfished based on an earlier benchmark assessment.

	2017	2019
F_{MSY} *proxy*	NA	NA
SSB_{MSY} (kg/tow)	NA	NA
MSY (mt)	NA	NA
Overfishing	Unknown	Unknown
Overfished	Yes	Yes

Projections: Short term projections cannot be computed using the PlanBsmooth approach. The PlanBsmooth approach estimates the rate of change in the recent three years of the smoothed survey biomass to be 0.936. This multiplier is applied to the average of the recent three years of catch (1,872 mt) to produce the catch advice for 2020 of 1,752 mt. The PlanBsmooth approach is fully described in NEFSC (2015) and available as an R package. A Shiny app demonstrating the performance of the PlanBsmooth approach is also available.

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The major source of uncertainty is the cause of the retrospective pattern that led to the analytical assessment of this stock not being accepted during the 2015 operational update meeting.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}).

  No retrospective adjustment of spawning stock biomass or fishing mortality was required.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Georges Bank Atlantic cod are not computed. Catch advice is derived from applying an estimate of recent change in the smoothed survey biomass to the average of the recent three years of catch and thus is influenced by uncertainty in survey estimates. The smoothed survey biomass is decreasing, but without a biomass reference point it is not known if rebuilding is on schedule.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  The experimental catchability data were not applicable to the Georges Bank Atlantic cod stock assessment because the catchability experiments were focused on flatfish species. Additionally, the restrictor cable used during the experiment may have impacted catchability of cod in both nets.

  The Marine Recreational Information Program data were changed from the previous assessment. For example, the 2017 assessment used years 2014-2016 to determine the average catch for use in the PlanBsmooth approach for catch advice. The old recreational catch (landings plus dead discards) for these years were 90, 140, and 399 mt, respectively. The new recreational catch for these years are 262, 507, and 1,085 mt. This results in the average total catch (US commercial, US recreational, and Canadian) increasing from 2,008 mt to 2,417 mt, a 20% increase. Thus, the catch advice in the 2017 assessment would have been 20% higher (3,667 mt instead of 3,047 mt) had the new MRIP data been used. Comparisons of the old and new recreational catch data are provided in the tables and figures files located in the data portal.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status for Georges Bank Atlantic cod remains overfished based on a qualitative evaluation of poor stock condition.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

The Georges Bank Atlantic cod continues to show a truncated age structure. The most recent survey values remain below the mean of their time series. The 2013 year class was larger than recent year classes, but has not continued to be large as it ages and is below the average from the 1970s at every age in both surveys.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Georges Bank Atlantic cod assessment could be improved with additional studies on natural mortality, the potential for missing catch, and other possible sources of retrospective patterns in analytical assessments.

- Are there other important issues?

  The differences in modeling approaches between the full Georges Bank cod assessment (reported here) and the TRAC cod assessment of eastern Georges Bank (a portion of the whole bank) remain a potential problem.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. Northeast Fisheries Science Center Reference Document 17-17; 259 p. doi: 10.7289/V5/RD-NEFSC-17-17 CRD17-17

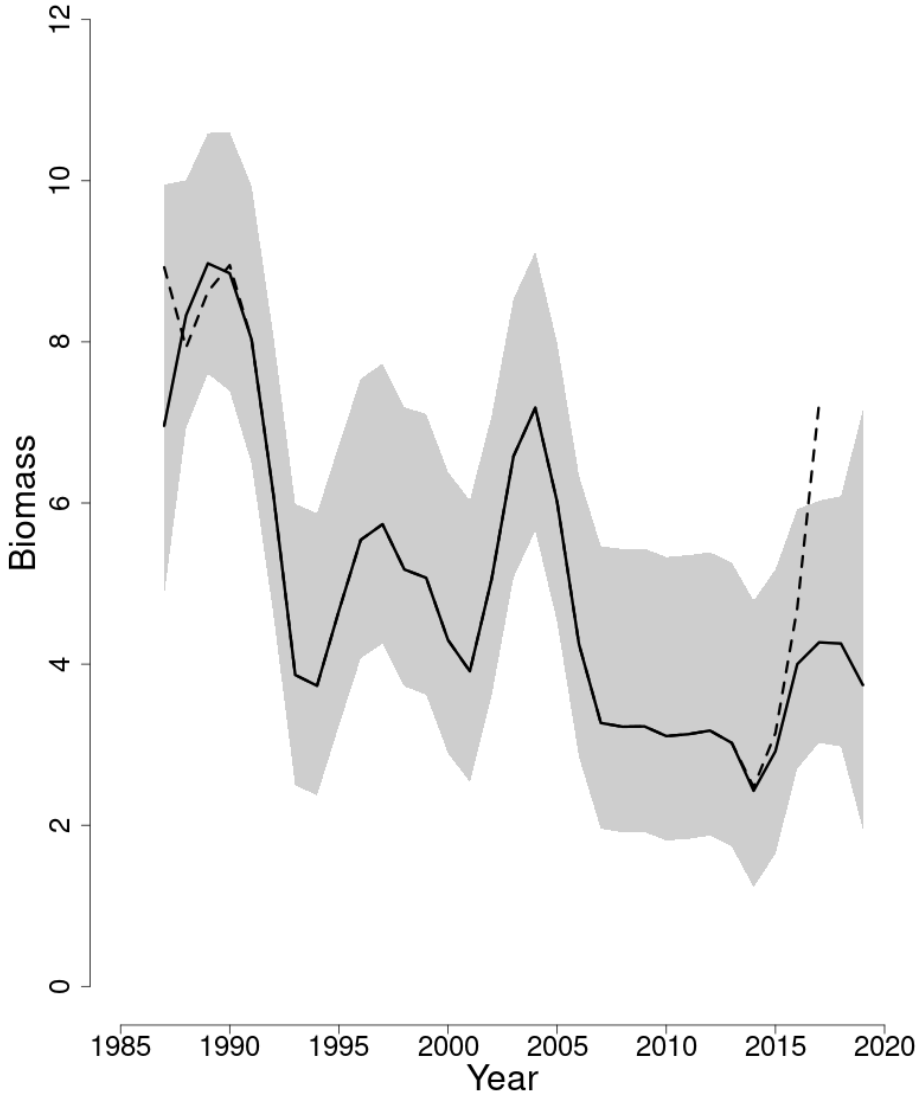

Figure 19: Trends in smoothed survey biomass (kg/tow) of Georges Bank Atlantic cod between 1987 and 2019 from the current (solid line) and previous (dashed line) assessment based on the 2019 assessment. The approximate 90% lognormal confidence intervals are shown.

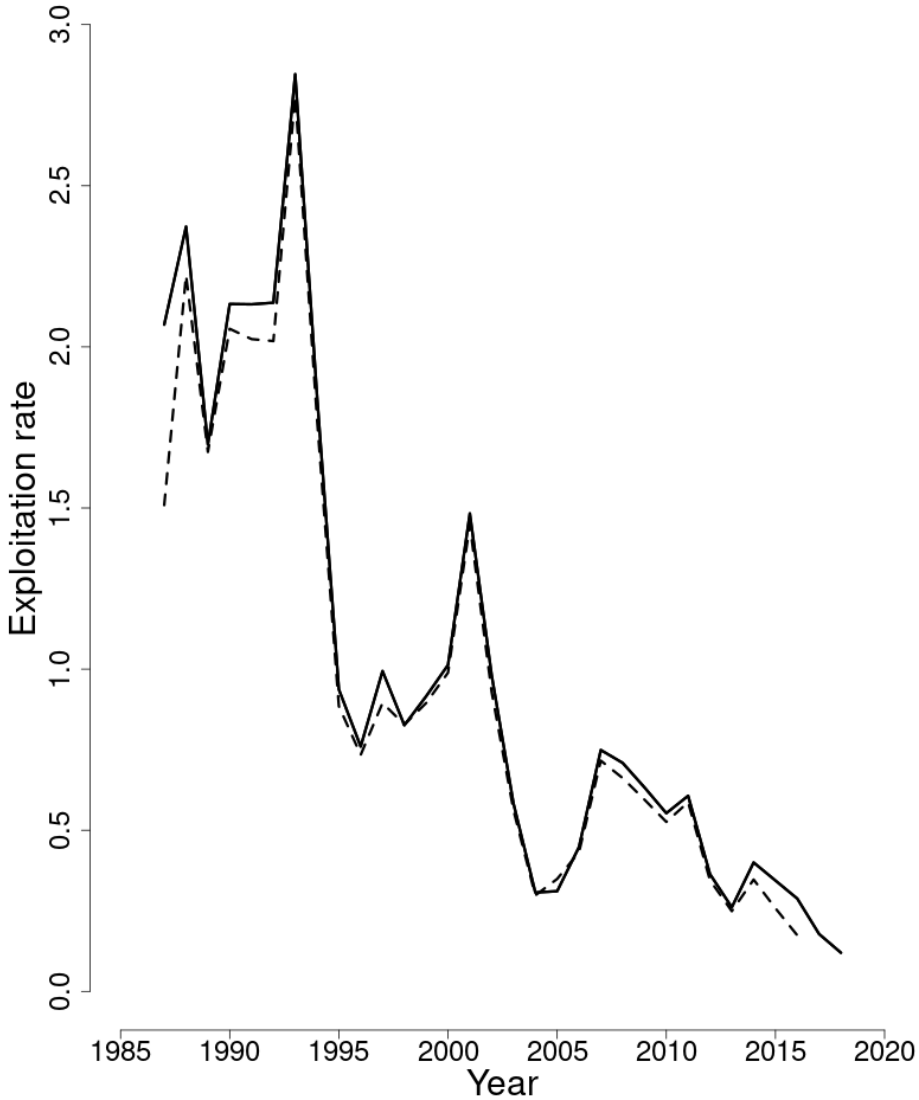

Figure 20: Trends in the relative exploitation rate (catch/smoothed survey) of Georges Bank Atlantic cod between 1987 and 2018 from the current (solid line) and previous (dashed line) assessment based on the 2019 assessment.

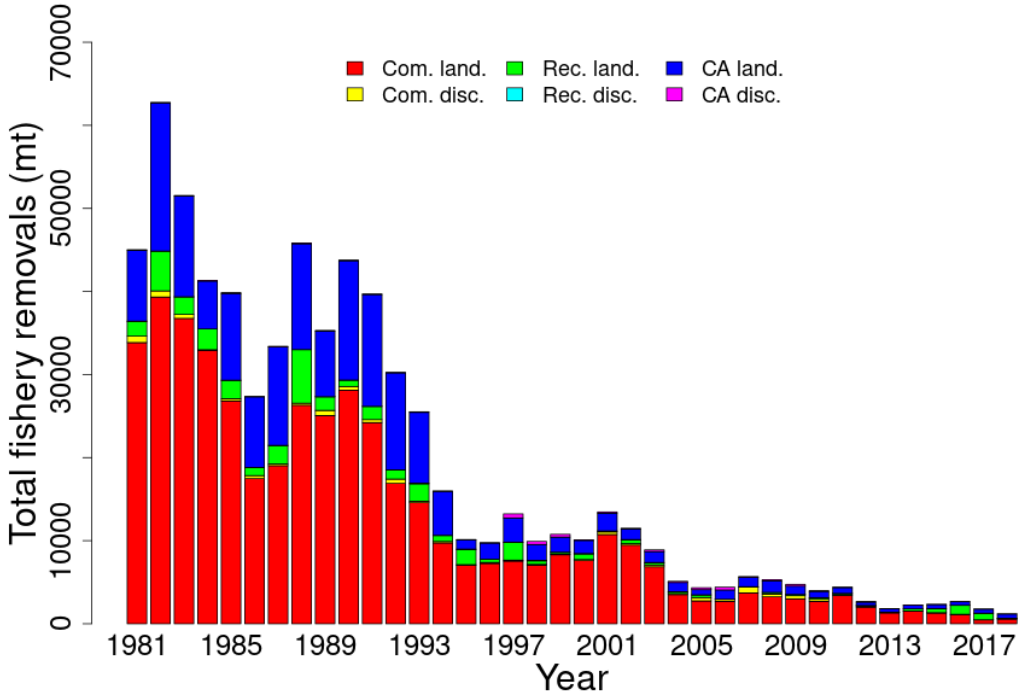

Figure 21: Total catch of Georges Bank Atlantic cod between 1981 and 2018 by fleet (US commercial, US recreational, or Canadian) and disposition (landings and discards).

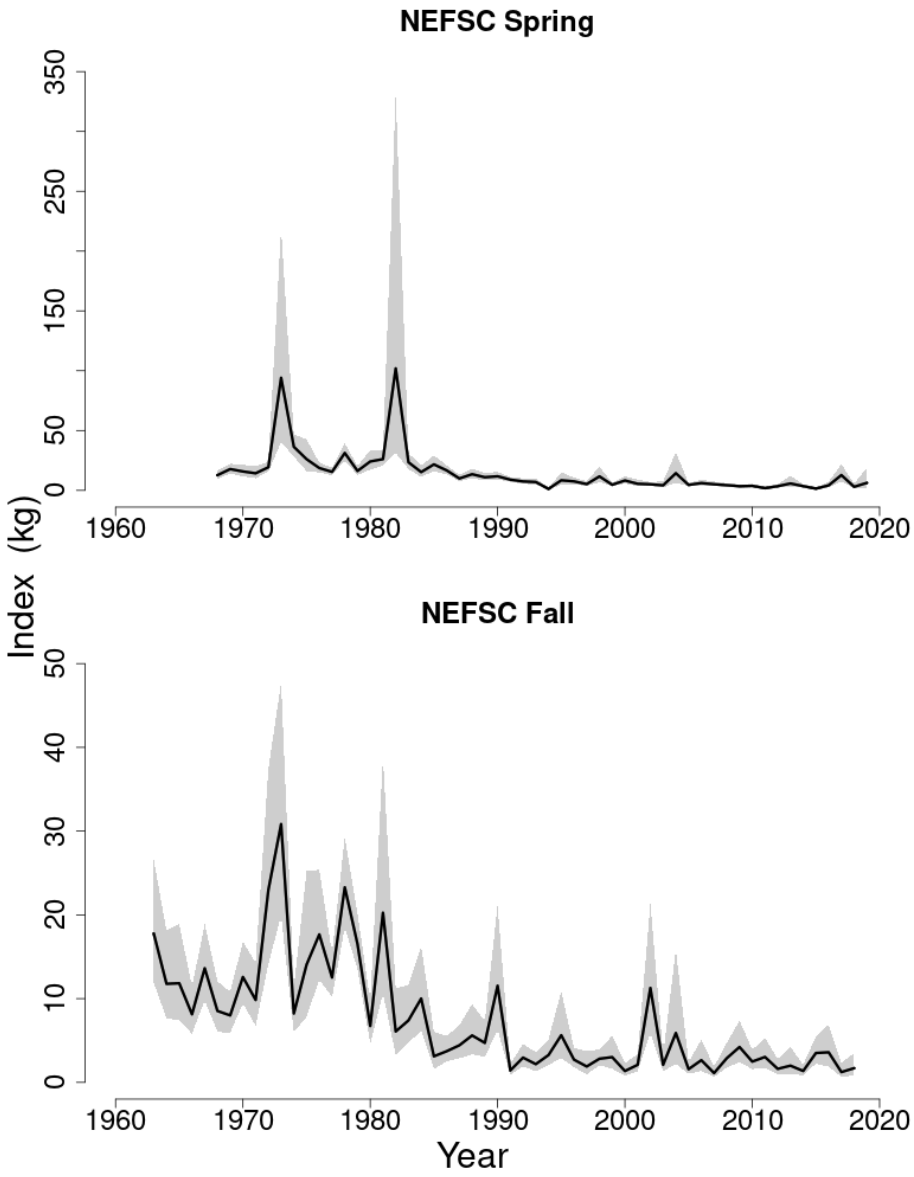

Figure 22: Indices of biomass for the Georges Bank Atlantic cod between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall trawl surveys. The approximate 90% lognormal confidence intervals are shown.

4 Georges Bank haddock

Liz Brooks

This assessment of the Georges Bank haddock (Melanogrammus aeglefinus) stock is a Level-2 operational update of the existing 2017 update VPA assessment (NEFSC, 2017). The last benchmark for this stock was in 2008 (Brooks et al., 2008). Based on the previous assessment in 2017, the stock was not overfished, and overfishing was not ocurring. This assessment updates commercial fishery catch data, research survey indices of abundance, weights and maturity at age, and the analytical VPA assessment model and reference points through 2018. Stock projections have been updated through 2022. This report reflects decisions made during the Peer Review September 9-12, 2019.

State of Stock: Based on this updated assessment, the Georges Bank haddock (*Melanogrammus aeglefinus*) stock is not overfished, and overfishing is not occurring (Figures 23-24). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 507,130 (mt) which is 365% of the biomass target (SSB_{MSY} *proxy* = 138,924; Figure 23). The 2018 average fishing mortality on ages 5-7 was estimated to be 0.061 which is 18% of the overfishing threshold proxy (F_{MSY} *proxy* = 0.33; Figure 24). The F_{MSY} *proxy* is expressed as the average F on ages 5-7 for compariability with the VPA estimated F.

Table 16: Catch and status table for Georges Bank haddock. All weights are in (mt), recruitment is in (000s), and $\bar{F}_{5-7}$ is the average fishing mortality on ages 5 to 7. Model results are from the current updated VPA assessment. A rho adjustment was not applied to values in this Table.

	2011	2012	2013	2014	2015	2016	2017	2018
			Data					
US Commercial discards	212	321	538	1,409	1,552	1,880	786	408
US Commercial landings	5,210	1,550	1,659	4,240	4,762	3,682	3,217	4,017
Canadian Catch	11,248	5,064	4,631	12,953	14,374	11,713	13,384	12,222
Catch for Assessment	16,670	6,935	6,828	18,601	20,687	17,274	17,387	16,647
			Model Results					
Spawning Stock Biomass	45,624	35,501	83,187	118,415	202,052	574,481	793,125	859,587
$\bar{F}_{5-7}$	0.425	0.522	0.45	0.447	0.332	0.23	0.068	0.034
Recruits (age 1)	207,156	38,754	29,515	2,267,641	55,083	154,684	546,138	79,974

Table 17: Comparison of reference points estimated in an earlier assessment and from the current assessment update. An $F_{40\%}$ proxy was used for the overfishing threshold (simple average for the current assessment, numbers weighted average for the previous assessment). The medians and 90% probability intervals are reported for MSY, SSBMSY, and RMSY, based on long-term stochastic projections with fishing mortality fixed at $F_{40\%}$.

	2017	2019
F_{MSY} *proxy*	0.35	0.33
SSB_{MSY} (mt)	104,312	138,924 (67,347 - 511,852)
MSY (mt)	24,400	30,489 (14,894 - 111,258)
Median recruits (age 1) (000s)	52,249	59,143 (2,780 - 394,017)
Overfishing	No	No
Overfished	No	No

Projections: Short term projections of biomass were derived by sampling from a cumulative distribution function (cdf) of recruitment estimates from ADAPT VPA (corresponding to SSB>75,000 mt and dropping the two most recent year class estimates for 2017 and 2018). The extremely large 1963, 2003, 2010, 2013, and 2016 year classes were included in the cdf. The annual fishery selectivity was a recent 5 year average except for the 2013 year class, which was assigned the same selectivity at age as the 2010 year class. The 2010 and 2013 year classes have demonstrated the slowest growth of any observed year classes in the time series. The maturity ogive was a recent 5 year average. Mean weights at age were a recent 2 year average, except for the 2010 and 2013 year classes, where recent trends in growth were assumed to continue. Retrospective adjustments were applied to the starting numbers at ages (2019) in the projections (each age was multiplied by 0.59).

Table 18: Short term projections of total fishery catch and spawning stock biomass for Georges Bank haddock based on a harvest scenario of fishing at F_{MSY} *proxy* between 2020 and 2022. Catch in 2019 was assumed to be 19,445 mt (estimate provided by the Groundfish Plan Development Team).

Year	Catch (mt)	SSB (mt)	$\bar{F}_{5-7}$
2019	19,445	605,990 (443,224 - 853,0145)	0.052 (0.036 - 0.072)

Year	Catch (mt)	SSB (mt)	$\bar{F}_{5-7}$
2020	184,822 (131,096 - 271,319)	581,672 (429,415 - 810,119)	0.332
2021	106,805 (79,085 - 148,763)	503,812 (363,623 - 755,210)	0.332
2022	100,009 (73,029 - 145,801)	412,276 (289,733 - 718,407)	0.332

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Sources of uncertainty include the retrospective bias, and future assumptions about

weights and selectivity at age. The 2013 year class accounts for a substantial portion of catch and SSB in projections (approximately 80% of catch and 60% of SSB in 2019 and 2020). The ρ-adjusted projections reduce all starting numbers at age to 59% of unadjusted values (i.e., all 2019 numbers at age are multiplied by 1/(1 + ρ[SSB]) = 0.59). The assumed values for selectivity and weights at age in the 2017 update were fairly accurate when compared to the observed weights and estimated selectivity for 2017 and 2018, and may indicate less uncertainty for these parameters compared to previous projections. This update has retained the assumptions used to derive those values in the current projections, but it is unknown if growth and selectivity patterns will change if abundance increases further. The magnitude of the 2016 year class is another source of uncertainty. It is currently estimated to be more than twice the size of the current estimate of the 2010 year class, and accounts for about 20% of projected SSB and 10-30% of projected catch in 2020-2022. The 2018 year class is also estimated to be large (1.66 times the current estimate of the 2010 year class) and highly uncertain (CV=138%), however, its contribution to projections is negligible until 2021 for SSB (18%) and 2022 for catch (11%).

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or $\bar{F}_{5-7}$ lies outside of the approximate joint confidence region for SSB and $\bar{F}_{5-7}$).

  The 7-year Mohn's ρ, relative to SSB, was 0.89 in the 2017 assessment and was 0.70 in 2018. The 7-year Mohn's ρ, relative to F, was -0.55 in the 2017 assessment and was -0.44 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimates of 2018 SSB (SSB_ρ=507,130) and 2018 F (F_ρ=0.061) were outside the approximate 90% confidence regions around SSB (614,031 - 1,253,991) and F (0.026 - 0.046). A retrospective adjustment was made for both the determination of stock status and for projections of catch in 2020. The retrospective adjustment changed the 2018 SSB from 859,587 to 507,130 and the 2018 $\bar{F}_{5-7}$ from 0.034 to 0.061.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  As noted in (1) above, population projections for Georges Bank haddock are uncertain due to the retrospective bias, assumed future values of selectivity and weights at age, and magnitude of incoming 2016 and 2018 year classes. This stock is not in a rebuilding plan.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  No changes, other than the incorporation of new data, were made to the Georges Bank haddock assessment for this update. However, recent years where the DFO survey did not sample the full Georges Bank strata (2012, 2013, 2015, 2017, 2018) were dropped from the VPA analysis.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was

used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

The experimental catchability data were not applicable to the Georges Bank haddock stock assessment because the catchability experiments were focused on flatfish species.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status of Georges Bank haddock has not changed.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The Georges Bank haddock shows a broad age structure, and broad spatial distribution. This stock has produced several exceptionally strong year classes in the last 15 years, leading to record high SSB in recent years. Catches in recent years have been well below the total quota (US+Canada). All survey indices of abundance support the finding that this stock is at an all-time high. Weights at age have been declining since the large 2003 year class, and show further declines with the most recent data.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  Projection advice and reference points for Georges Bank haddock are strongly dependent on recruitment. A decade ago, extremely large year classes were considered anomalies (e.g., 1963 and 2003). However, since 2003, there have been four more extremely large year classes (2010, 2013, 2016, and 2018). Future work could focus on recruitment forecasting and providing robust catch advice. Assumptions about weights at age and selectivity are very influential in short term projections. As multiple large year classes move through the population, it is difficult to predict how strong the density dependent response will be, but future work could continue examining performance of projected values with realized values. For this assessment, reference points are estimated with a recent 5 year average for selectivity, maturity, and weights at age, whereas short-term projections use year-specific decisions to deal with the current large year classes. Considering that estimated population abundance at MSY is much less than the current population abundance, recent average biological and fishery parameters may not reflect MSY conditions. Calculating per recruit statistics on an annual basis demonstrates the dynamic range of reference points in response to density dependent changes in growth (see Model Results pdf).

- Are there other important issues?

  The Georges Bank haddock assessment has developed a major retrospective pattern in recent years. This stock assessment has historically performed very consistently. This should continue to be monitored. Density-dependent responses in growth should also continue to be monitored. On an annual basis, known research removals account for 0-0.7% of annual catch removals by weight, and 0-4.6% of annual catch removals by number; this level is insufficient to cause the observed retrospective pattern.

4.1 Reviewer Comments: Georges Bank haddock

4.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

The Georges Bank haddock stock was assessed using VPA including a retrospective adjustment for status determination and catch projections. A subset of this stock is also assessed by the Transboundary Resources Assessment Committee (TRAC) for the eastern portion of the stock only. Both assessments assume a closed population, which cannot be true for both. Previous research on stock identification and the current resource distribution suggests that the entire Bank should be considered a unit stock. The eastern Georges Bank haddock VPA was rejected in 2019 by TRAC due to worsening retrospective patterns and other diagnostics, poor tracking of survey trends and catchability being greater than 1. These concerns are not as applicable to the HADGB VPA, leading to the continued use of the VPA for HADGB.

Further, while the metrics used for setting ABCs are accurate, the scale of the assessment is uncertain in terms of total biomass, because this stock appears to be much larger than ever observed.

4.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. Similar to the 2017 update assessment, Canada DFO surveys that did not sample the entire Bank were excluded from the VPA calibration, and this revision from the benchmark method is well justified.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used. There was a 33% increase in the estimate of SSB_{MSY}, primarily from increased recent recruitment.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is not overfished and overfishing is not taking place.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

Multiple surveys indicate this stock has been much larger in recent years than observed in the past, age data indicate low total mortality rates, and survey data indicate expanded area occupied by the stock. The retrospective pattern in the VPA is getting better.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at FMSY or at an FMSY proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the VPA with retrospective adjustments.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

During the upcoming research track in 2021, statistical catch-at-age or state-space modeling approaches should be considered to allow improved tracking of survey indices and allow for uncertainty in catch at age (particularly for dominant year classes) and more control over fishery and survey selectivity estimation.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Brooks, E.N, M.L. Traver, S.J. Sutherland, L. Van Eeckhaute, and L. Col. 2008. In. 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. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii. http://www.nefsc.noaa.gov/publications/crd/crd0815/

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. doi: 10.7289/V5/RD-NEFSC-17-17

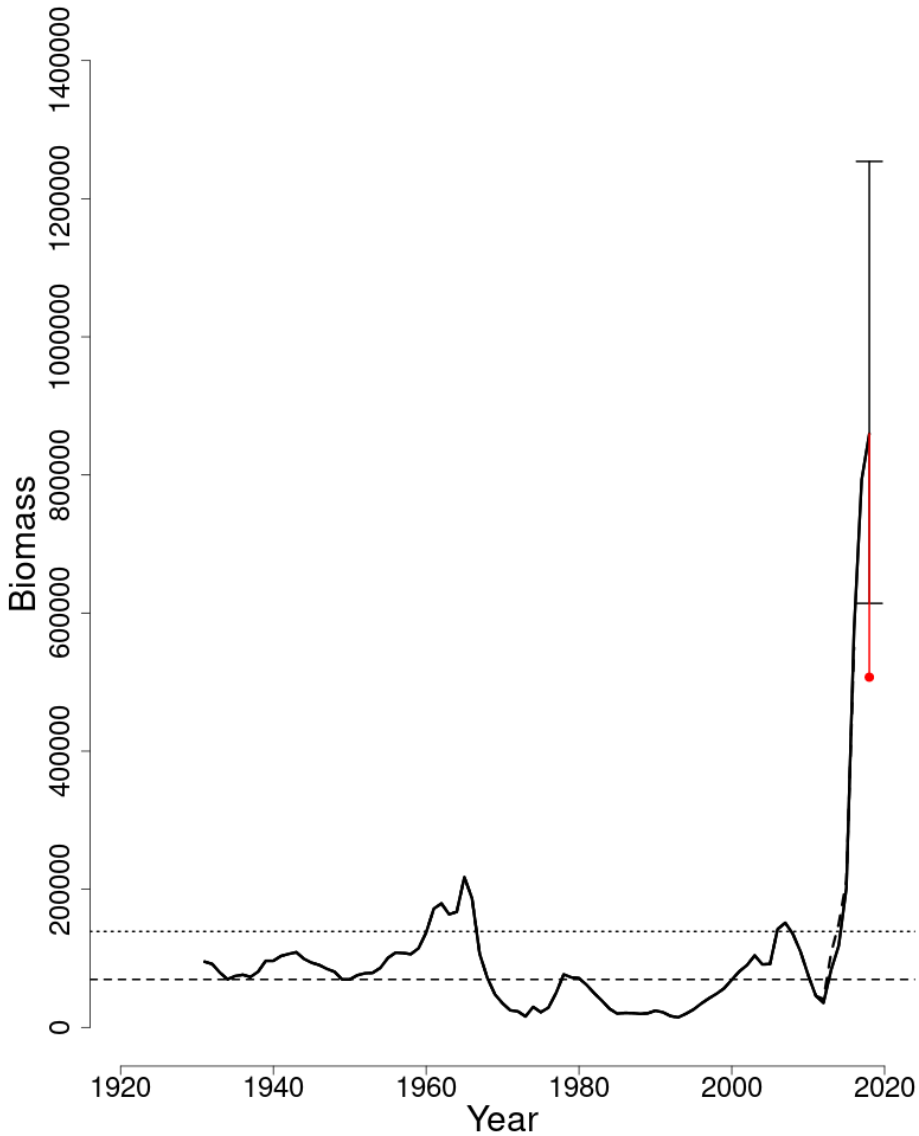

Figure 23: Trends in spawning stock biomass of Georges Bank haddock between 1931 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The 90% bootstrap probability intervals are shown.

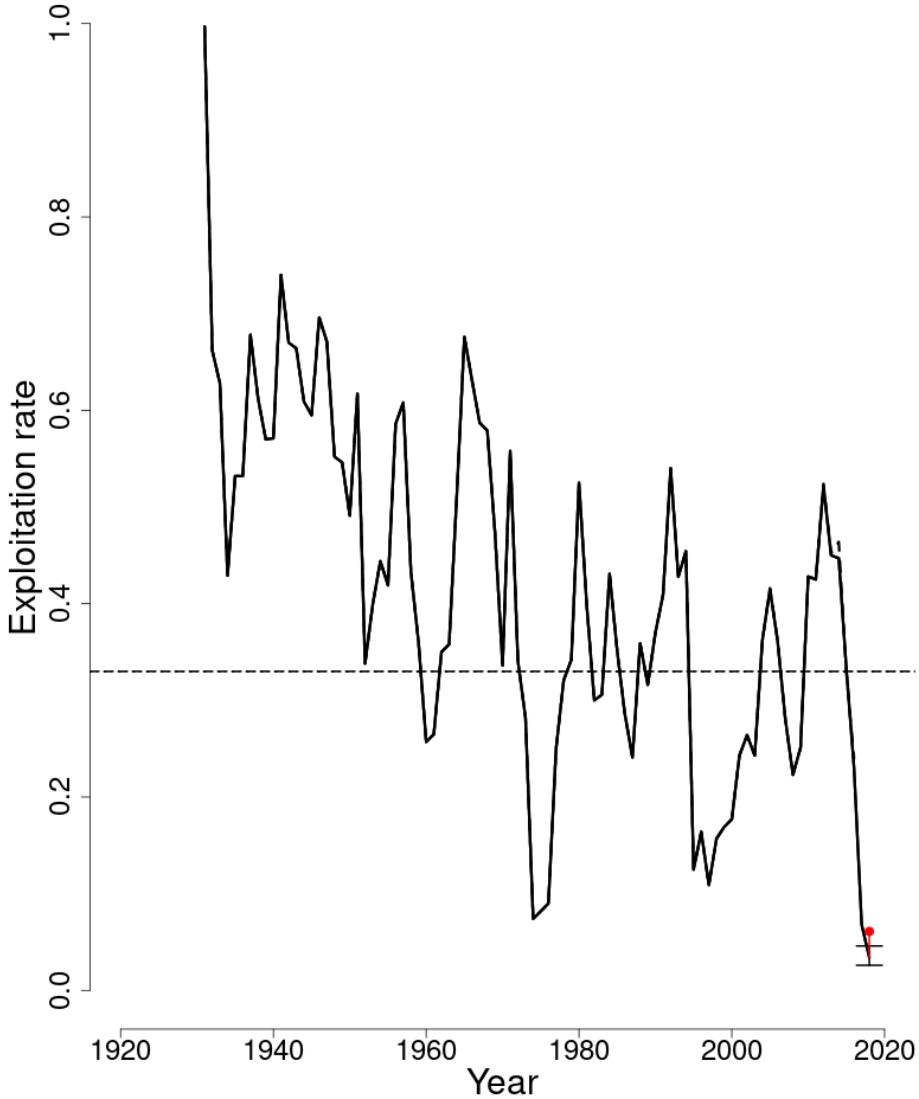

Figure 24: Trends in the average fishing mortality ($\bar{F}_{5-7}$) of Georges Bank haddock between 1931 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.33; horizontal dashed line) based on the 2019 assessment. $\bar{F}_{5-7}$ was adjusted for a retrospective pattern and the adjustment is shown in red. The 90% bootstrap probability intervals are shown.

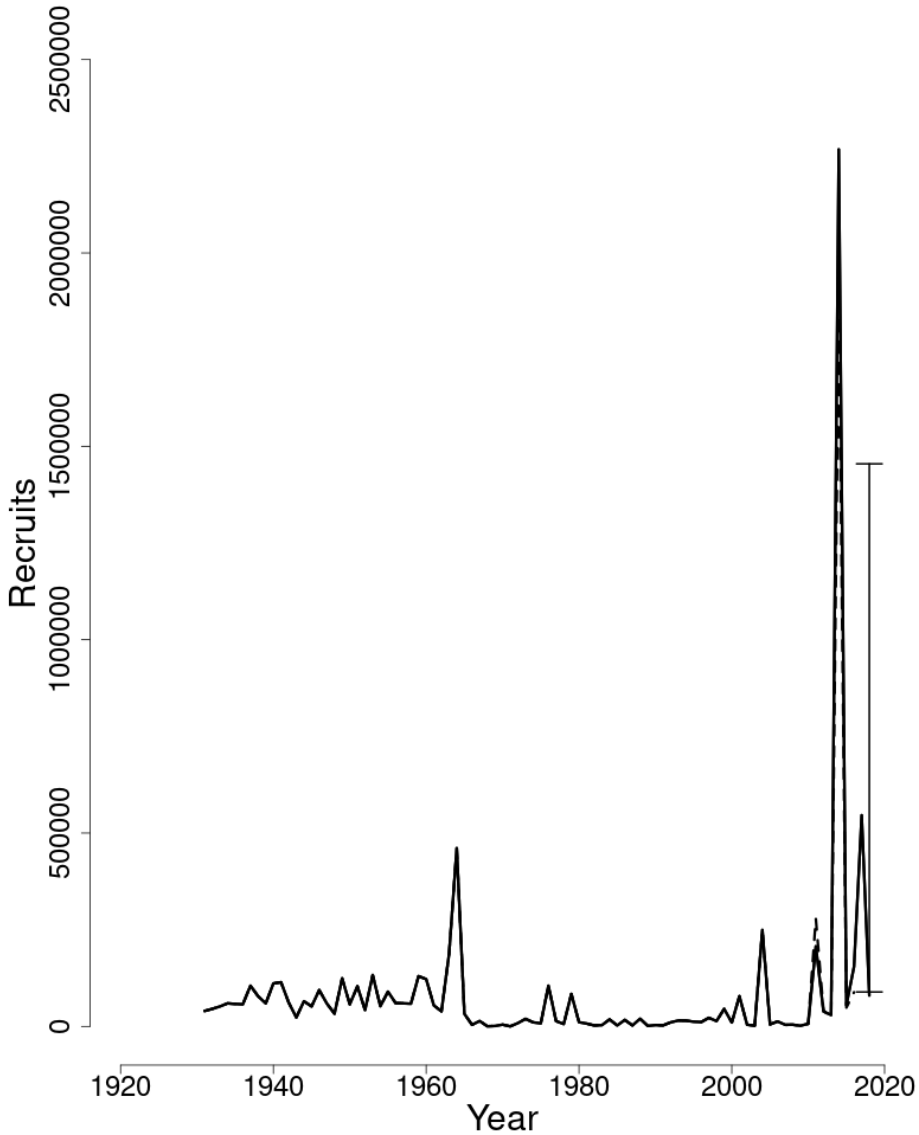

Figure 25: Trends in Recruits (age 1) (000s) of Georges Bank haddock between 1931 and 2018 from the current (solid line) and previous (dashed line) assessment. The 90% bootstrap probability intervals are shown.

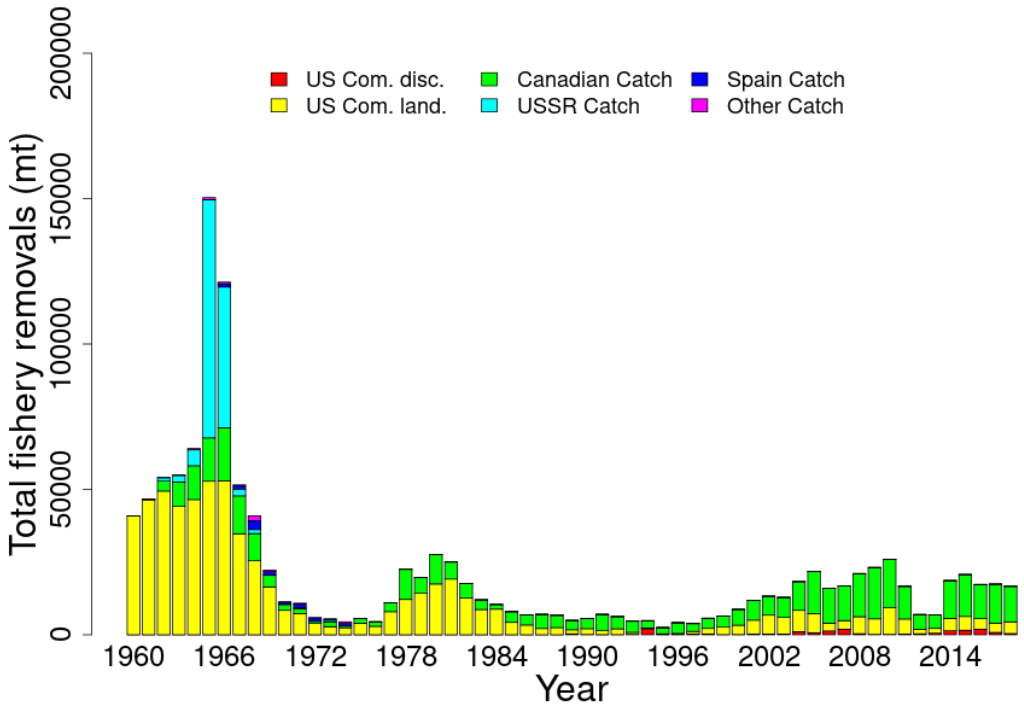

Figure 26: Total catch of Georges Bank haddock between 1931 and 2018 by fleet (US Commercial, Canadian, or foreign fleet) and disposition (landings and discards).

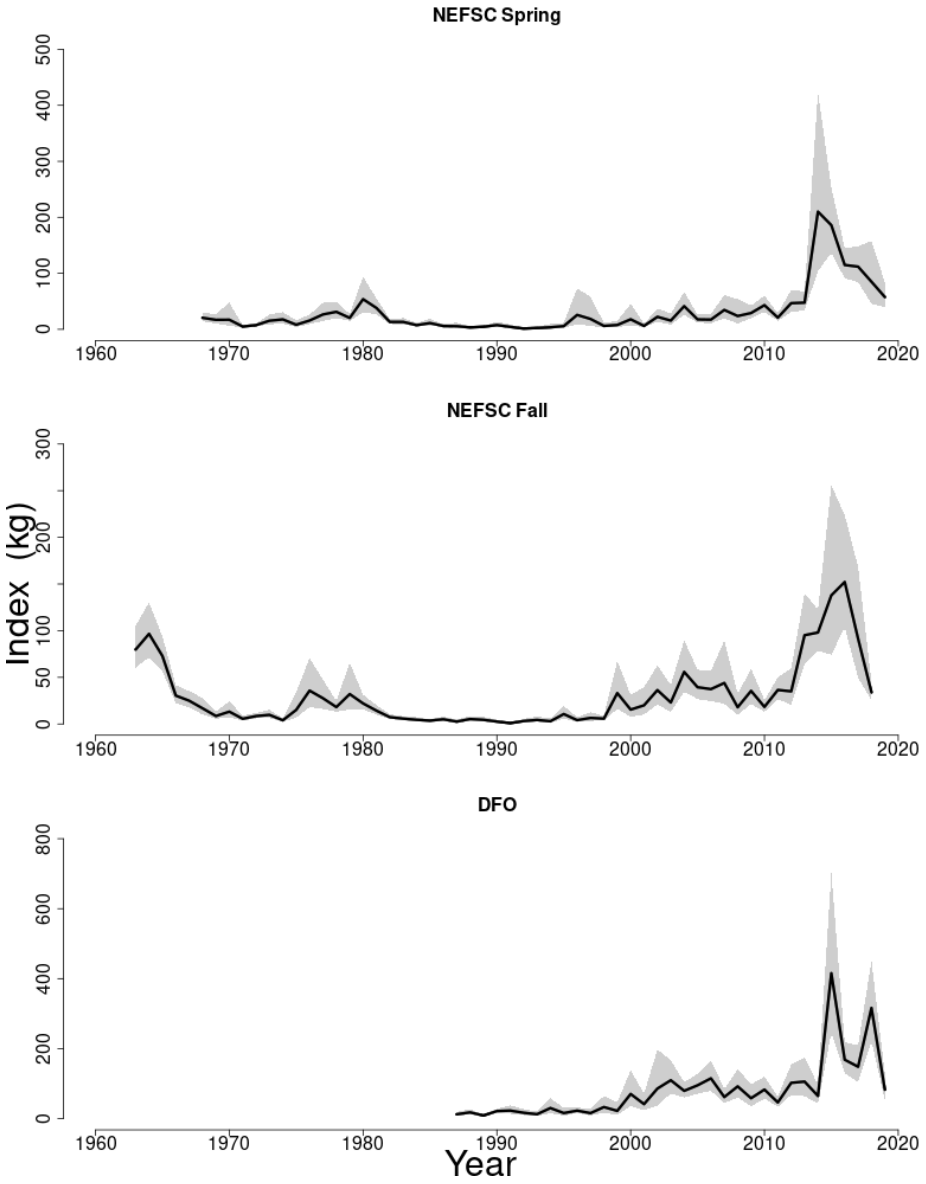

Figure 27: Indices of biomass (Mean kg/tow) for the Georges Bank haddock stock between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys and the DFO winter bottom trawl survey. The approximate 90% lognormal confidence intervals are shown.

5 Gulf of Maine haddock

Charles Perretti

This assessment of the Gulf of Maine haddock (Melanogrammus aeglefinus) stock is an operational assessment of the existing benchmark assessment (NEFSC 2014). Based on the previous assessment (NEFSC 2017), the stock was not overfished, and overfishing was not occurring. This assessment updates commercial and recreational fishery catch data, research survey indices of abundance, and the analytical ASAP assessment model and reference points through 2018. Additionally, stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, the stock status for the Gulf of Maine haddock (*Melanogrammus aeglefinus*) stock is not overfished and overfishing is not occurring (Figures 28-29). Retrospective adjustments were made to the model results (see Special Comments section of this report). Spawning stock biomass (SSB) in 2018 was estimated to be 82,763 (mt) which is 1035% of the biomass target (SSB_{MSY} *proxy* = 7,993; Figure 28). The 2018 fully selected fishing mortality was estimated to be 0.082 which is 22% of the overfishing threshold proxy (F_{MSY} *proxy* = $F_{40\%}$ = 0.369; Figure 29).

Table 19: Catch and status table for Gulf of Maine haddock. All weights are in (mt) recruitment is in (000s) and F_{Full} is the fully selected fishing mortality. Model results below are from the current updated ASAP assessment **without** retrospective adjustment.

	2011	2012	2013	2014	2015	2016	2017	2018
			Data					
Recreational discards	21	158	504	618	526	966	733	319
Recreational landings	400	467	528	457	295	1,026	1,747	817
Commercial discards	6	18	32	22	42	72	91	54
Commercial landings	499	417	212	314	650	1,342	2,273	2,542
Foreign landings	0	0	0	0	0	0	0	0
Catch for Assessment	926	1,060	1,277	1,412	1,513	3,406	4,843	3,731
			Model Results					
Spawning Stock Biomass	5,019	6,215	9,963	15,575	34,226	58,404	65,397	63,143
F_{Full}	0.266	0.383	0.349	0.254	0.144	0.18	0.167	0.105
Recruits (age 1)	17,611	5,800	24,849	140,737	7,962	7,502	12,480	3,246

Table 20: Comparison of reference points estimated in an earlier assessment and from the current operational assessment. The overfishing threshold is the F_{MSY} *proxy* ($F_{40\%}$). The biomass target, (SSB_{MSY} *proxy*) was based on long-term stochastic projections of fishing at the F_{MSY} *proxy*. Median recruitment reflects the median estimated age-1 recruitment from 1977 - 2016. Intervals shown reflect the 5th and 95th percentiles.

	2017	2019
F_{MSY} *proxy*	0.455 (0.380 - 0.538)	0.369 (0.307 - 0.447)
SSB_{MSY} (mt)	6,769 (2,525 - 27,545)	7,993 (3,218 - 34,191)
MSY (mt)	1,547 (584 - 6,160)	1,597 (651 - 6,797)
Median recruits (age 1) (000s)	1,498 (275 - 17,307)	1,789 (285 - 17,883)
Overfishing	No	No
Overfished	No	No

Projections: Short term projections of median total fishery yield and spawning stock biomass for Gulf of Maine haddock were conducted based on a harvest scenario of fishing at the F_{MSY} *proxy* between 2020 and 2022. Catch in 2019 has been estimated at 5,239 mt. Recruitment was sampled from a cumulative distribution function of model estimated age-1 recruitment from 1977-2016. The age-1 estimate in 2019 was generated from the geometric mean of the 1977-2018 recruitment series. The annual fishery selectivity, maturity ogive, and mean weights at age used in the projections were estimated from the most recent 5 year averages. Retrospective adjustments were applied in the projections.

Table 21: Short term projections of total fishery catch and spawning stock biomass for Gulf of Maine haddock based on a harvest scenario of fishing at F_{MSY} *proxy* ($F_{40\%}$) between 2020 and 2022. Catch in 2019 was assumed to be 5,239 (mt).

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	5,239	103,670	0.075

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	24,803	91,167	0.369
2021	19,536	65,929	0.369
2022	12,563	50,468	0.369

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The strength of terminal year classes had been a large source of uncertainty in previous assessments. The 2012 and 2013 year classes are now reasonably well estimated and the relative size of more recent year classes is expected to be near average and unlikely to have

much impact on the terminal estimates of stock size or in the performance of stock projections. Future reference point values will be sensitive to whether future recruitment events are similar to the 2012 and 2013 year classes, or to the historical average. In addition, the reliability of fishery catch data remains an important source of uncertainty for this stock.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lie outside of the approximate joint confidence region for SSB and F_{Full}).

  The terminal year rho-adjusted SSB is greater than the upper bound of the confidence interval for SSB, therefore this assessment meets the criteria for a major retrospective pattern. The 7-year Mohn's rho value for SSB is -0.24, and for F is 0.29. Retrospective adjustments were made to terminal year F and SSB.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Gulf of Maine haddock are reasonably well determined. The projected 2018 biomass from the last assessment is within the confidence interval of the 2018 biomass estimated in the current assessment. This stock is not currently in a rebuilding plan.

- Describe any changes that were made to the current stock assessment beyond incorporating additional years of data, and the effect these changes had on the assessment and stock status.

  Recreational catch estimates were re-estimated in this update by using the re-calibrated MRIP data. In general, inclusion of the re-calibrated data resulted in an increase in SSB, F, and recruitment. Prior to 2004, there is no length information for recreational releases, and there are several years with either limited or no length information for recreational harvest. However, proportions-at-age are similar between the pre- and post-calibrated data. Therefore, recreational catch-at-age prior to 2004 was calculated by applying the historical proportions-at-age to the new total catch numbers estimated in the re-calibrated MRIP procedure.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  The experimental catchability data were not applicable to the Gulf of Maine haddock stock assessment because the catchability experiments were focused on flatfish species.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  There has been no change in stock status since the previous assessment (2017).

- Provide qualitative statements describing the condition of the stock that relate to stock status.

The Gulf of Maine haddock has experienced several large recruitment events since 2010. The population biomass is currently near an all time high and overall, the population is experiencing low mortatity.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  A better understanding of recruitment processes may help to improve recruitment forecasting.

- Are there other important issues?

  None.

5.1 Reviewer Comments: Gulf of Maine haddock

5.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference. The ASAP model was used to provide updated estimates for Gulf of Maine haddock. Following protocols an adjustment was made to the terminal F and SSB estimates to account for the retrospective pattern coming out of the model analysis. This adjustment results in a lower F and higher SSB, which is the opposite of what is happening to Georges Bank haddock and most other groundfish species in this region.

5.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. Two additional years of data (2017-18) were incorporated into the assessment time series. Recalibrated recreational landings and discards from the MRIP survey were also included. The new MRIP data had little impact on the assessment.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use with the retrospective adjustment.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used. The updated F_{MSY} proxies ($F_{40\%}$) were calculated using the most recent 5-year average weights at age. The change in the reference points, F_{MSY} proxy decreased while SSB_{MSY} proxy increased relative to the last assessment, is due to mostly decreases in weight at age and increased recruitment.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is not overfished and overfishing is not occurring.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

Recent recruitment has been strong and the stock has recently been at record high levels. The stock is expected to decline towards SSB_{MSY} without further large recruitment events.

The abundance estimates of recent year classes have historically been an area of uncertainty, but are not expected to be a major source of uncertainty in this update. The 2012 and 2013 year classes are reasonably well estimated and the 2014-2017 year classes are near average and unlikely to have much impact on terminal estimates of stock size or projections.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the ASAP model with retrospective adjustment. Despite the direction of the retrospective pattern being in the opposite direction typically seen (the adjustment increased the starting population for projections), the Panel found it appropriate to make these adjustments to account for the retrospective pattern as a matter of protocol. The Panel suggests that the PDT present both retrospective adjusted and not retrospective adjusted projections to the full SSC to demonstrate the impact of this decision.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

The contrast in the direction of retrospective patterns between haddock stocks is worth examining. The patterns may reflect environmental change, catch misreporting (e.g. catch being assigned to the wrong stat area), differences in natural mortality or some other factor. More broadly, it is suggested that there be a broader examination of retrospective patterns and their potential sources across all stocks in this system. For the commercial fishery, the accuracy of fishery removal estimates is still in question. There may exist some stock-area reporting errors (Palmer 2017 CRD) and dealer misreporting.

The Panel recommends attempting to split the snapper and small market categories in the next assessment to better characterize the age composition of the fishery catch.

The Panel suggests exploring the utility of the complex season and length specific discard mortality rates used in this stock compared to the standard single value applied across season and length in the next assessment.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2014. 59th Northeast Regional Stock Assessment Workshop (59th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-09; 782 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD14-09

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD15-24

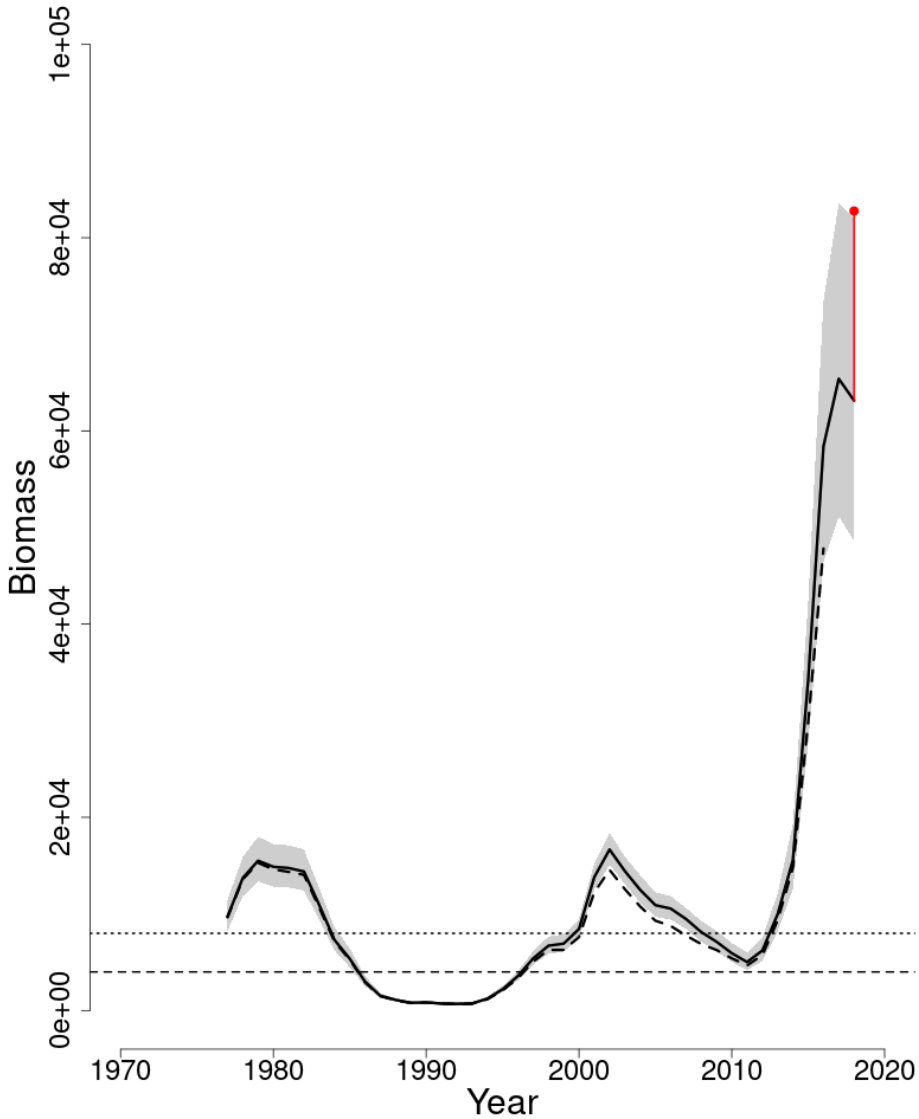

Figure 28: Trends in spawning stock biomass (SSB) of Gulf of Maine haddock between 1977 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. SSB was adjusted for a retrospective pattern and the adjustment is shown in red based on the 2019 assessment. The approximate 90% lognormal confidence intervals are shown.

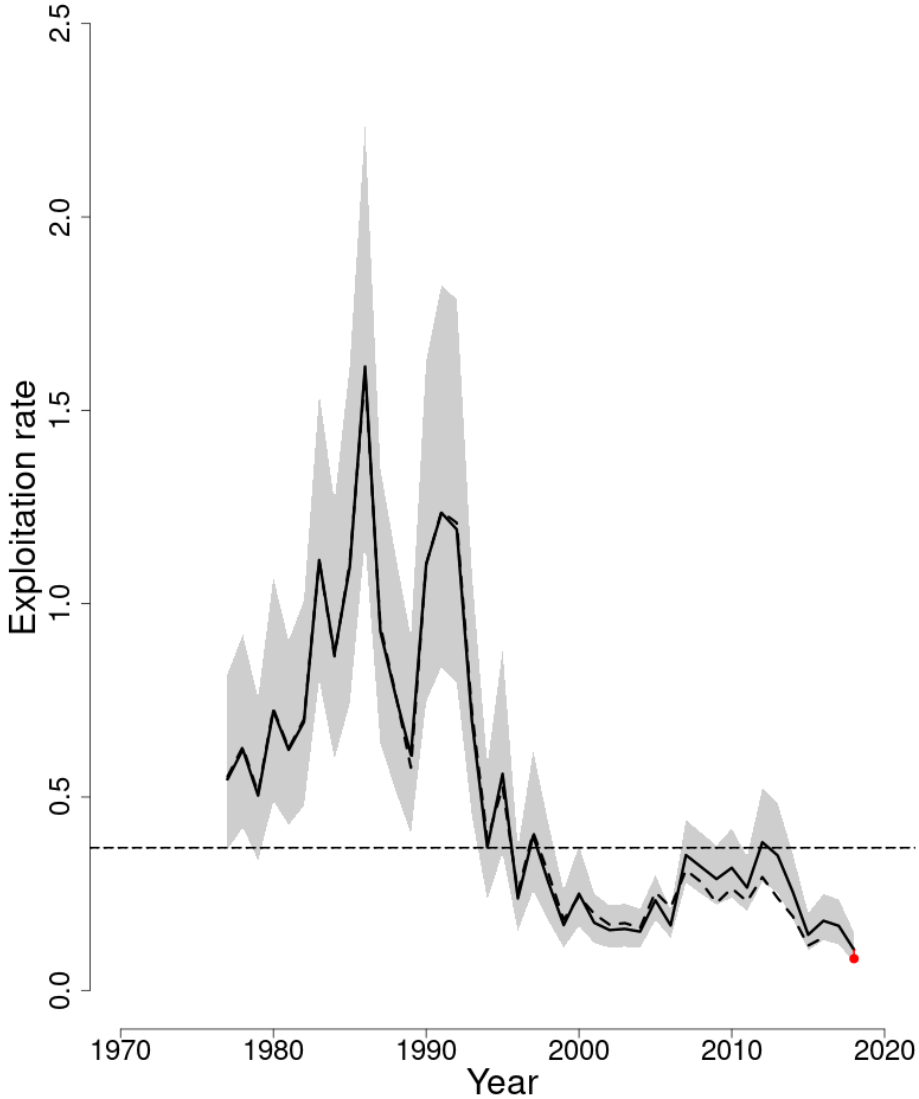

Figure 29: Trends in the fully selected fishing mortality (F) of Gulf of Maine haddock between 1977 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.369; horizontal dashed line) from the 2019 assessment model.F in 2019 was adjusted for a retrospective pattern and the adjustment is shown in red based on the 2019 assessment. The approximate 90% lognormal confidence intervals are shown.

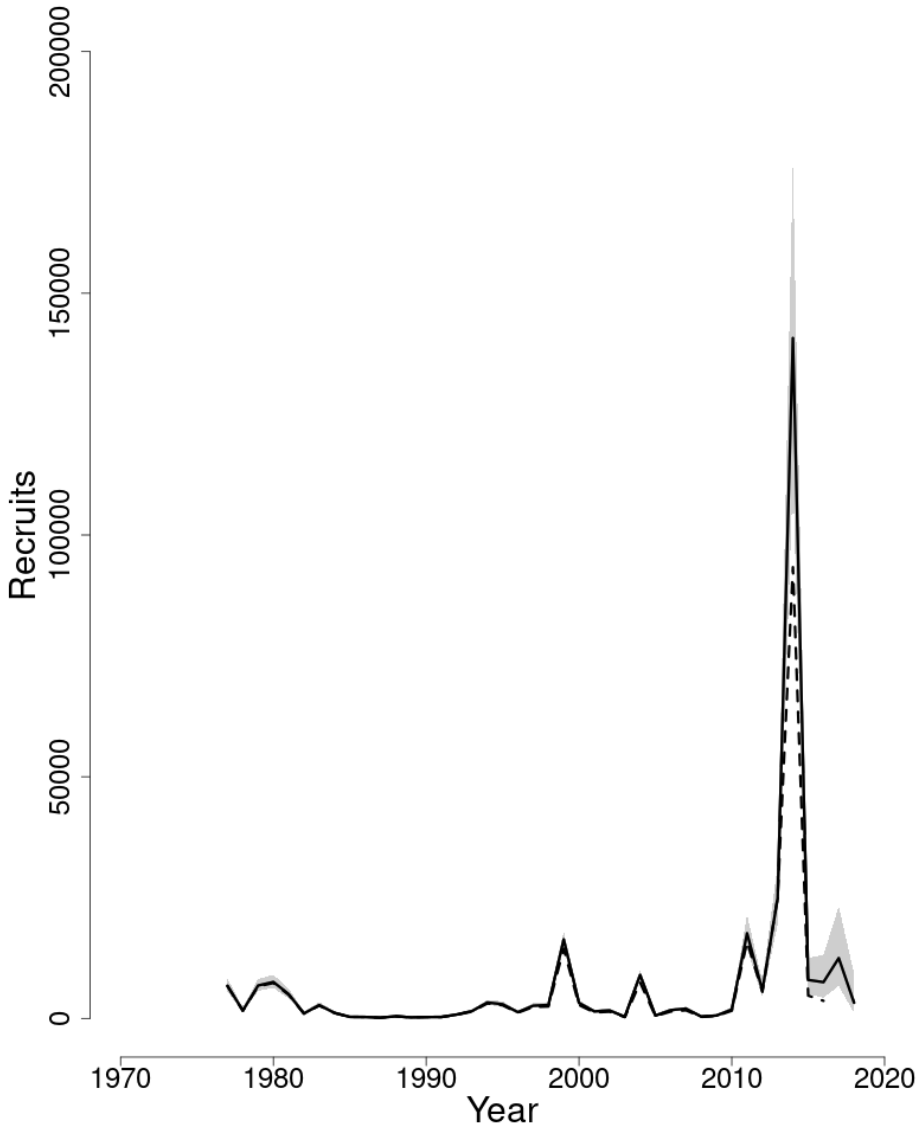

Figure 30: Trends in Recruits (age 1) (000s) of Gulf of Maine haddock between 1977 and 2018 from the current (solid line) and previous (dashed line) assessment. The approximate 90% lognormal confidence intervals are shown.

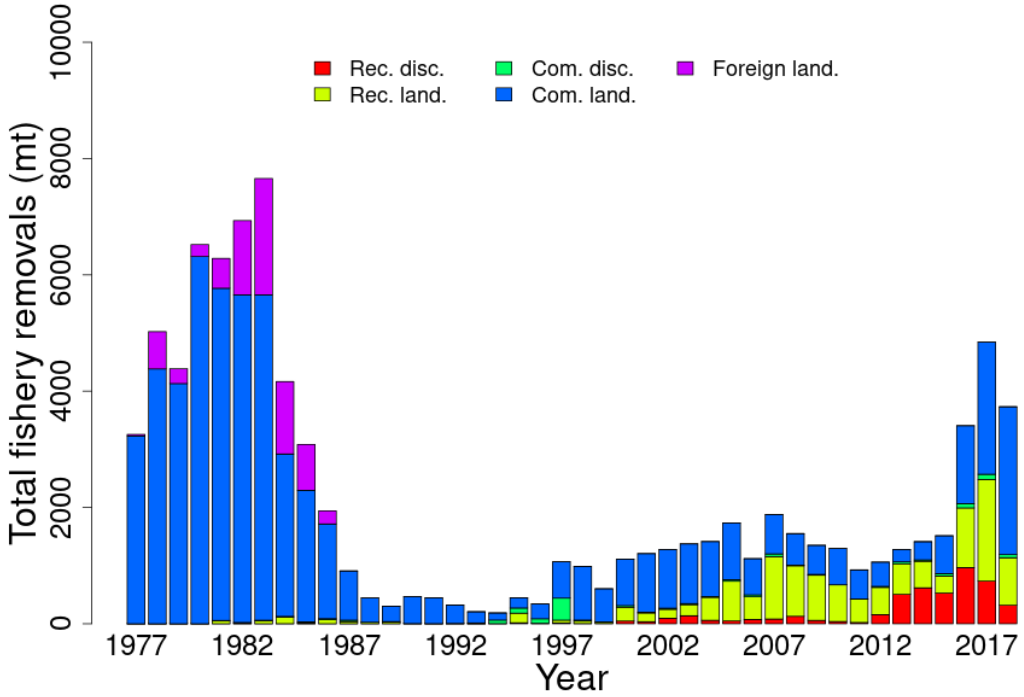

Figure 31: Total catch of Gulf of Maine haddock between 1977 and 2018 by fleet (commercial, recreational, or foreign) and disposition (landings and discards).

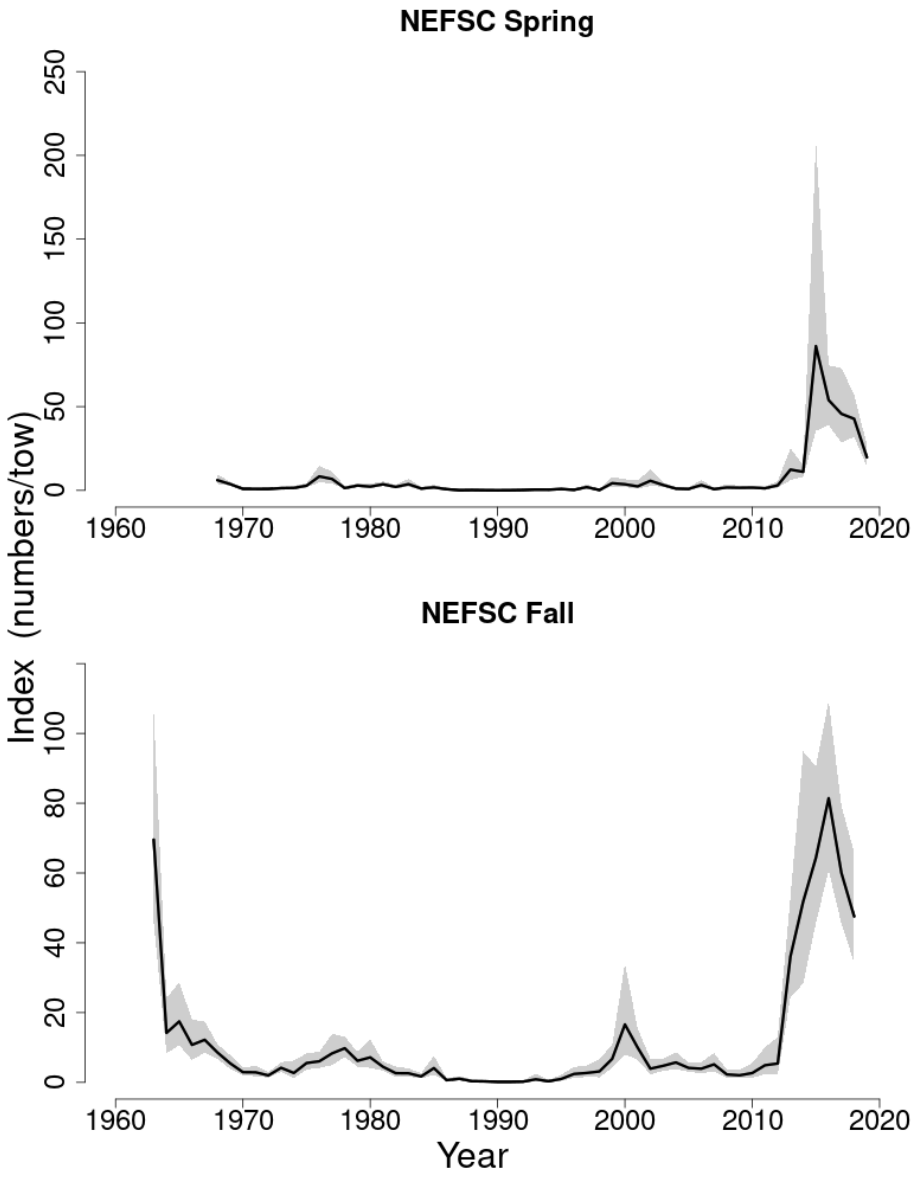

Figure 32: Indices of abundance for the Gulf of Maine haddock between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys. The approximate 90% lognormal confidence intervals are shown.

6 Cape Cod-Gulf of Maine yellowtail flounder

Larry Alade

This assessment of the Cape Cod-Gulf of Maine yellowtail flounder (Limanda ferruginea) stock is an operational assessment of the existing 2017 VPA assessment (Alade 2017). The last benchmark for this stock was in 2008 (Legault et al., 2008). Based on the previous assessment the stock was overfished, and overfishing was occurring. This 2019 assessment updates commercial fishery catch data, research survey indices of abundance, weights at age, and the analytical VPA assessment model and reference points through 2018. Additionally, stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, Cape Cod-Gulf of Maine yellowtail flounder (*Limanda ferruginea*) stock is not overfished and overfishing is not occurring (Figures 33-34). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 2,125 (mt) which is 62% of the biomass target (SSB_{MSY} *proxy* = 3,439; Figure 33). The 2018 fully selected fishing mortality was estimated to be 0.092 which is 29% of the overfishing threshold proxy (F_{MSY} *proxy* = 0.32; Figure 34).

Table 22: Catch and model results for Cape Cod-Gulf of Maine yellowtail flounder. All weights are in (mt), recruitment is in (000s) and F_{Full} is the average fishing mortality on ages (ages 4 and 5). Model results below are from the current updated VPA assessment **without** any retrospective adjustment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
				Data						
Commercial discards	175	87	74	146	86	54	45	66	50	45
Commercial landings	464	546	684	946	590	421	306	302	314	226
Total Catch for Assessment	639	633	758	1,092	676	475	351	368	365	271
				Model Results						
Spawning Stock Biomass	935	1,232	1,391	1,117	903	1,066	1,725	2,307	2,857	2,753
F_{Full}	0.754	0.501	0.669	1.062	1.015	0.44	0.204	0.133	0.118	0.078
Recruits (age 1)	4,005	3,321	3,232	3,086	5,614	5,241	5,784	5,719	7,524	5,537

Table 23: Comparison of reference points estimated in the previous assessment and from the current assessment update. An $F_{40\%}$ proxy was used for the overfishing threshold and SSB_{MSY} *proxy* was based on long-term stochastic projections.

	2017	2019
F_{MSY} *proxy*	0.273	0.32
SSB_{MSY} (mt)	4,640	3,439 (2,593 - 4,794)
MSY (mt)	1,154	1,138 (860 - 1,582)
Median recruits (age 1) (000s)	6,186	5,781
Overfishing	Yes	No
Overfished	Yes	No

Projections: Short term projections of biomass were derived by sampling an empirical cumulative distribution function of 30 recruitment estimates from the VPA model results. Hindcasted age-1 recruitment estimates for years 1977-1984 were excluded from the projections due to the poor linear relationship between the VPA age-1 estimates and the NEFSC age-1 autumn survey. The most recent two years (2018 and 2019) were also not included in the series of recruitment values due to high uncertainty in these estimates. The annual fishery selectivity, maturity ogive, and mean weights at age used in projection are the most recent 5 year averages. Retrospective adjustments were applied in the projections.

Table 24: Short term projections of total fishery catch and spawning stock biomass for Cape Cod-Gulf of Maine yellowtail flounder based on a harvest scenario of fishing at F_{MSY} *proxy* between 2021 and 2022. Catch in 2019 was assumed to be 271 (mt).

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	271	3,408 (2,807 - 4,104)	0.076

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	1173 (938 - 1401)	3,466 (2,853 - 4,274)	0.320
2021	998 (837 - 1210)	3,018 (2,534 - 3,649)	0.320
2022	1000 (800 - 1,276)	3,039 (2,429 - 3,861)	0.320

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Retrospective patterns remain a source of uncertainty in the assessment. This has persisted for a number of years causing a decrease in estimates of adult biomass and recruitment and an increase in estimates of fishing mortality when more years of data are added. However, the magnitude of these retrospective biases in this assessment were notably reduced by approximately 61% for both fishing mortality and adult biomass when compared to the previous 2017 operational assessment. Despite the improvement in retrospective bias, rho adjusted projections were still conducted, which reduced starting numbers at age by an average of 39% (Note that the rho adjustments for the projections are based on numbers at age with Mohns rho ranging from 8-92 %). The spring 2019 aging data from the Massachusetts state inshore survey was not available in time for this assessment update to derive the 2019 indices at age for the inshore state surveys (Massachusetts DMF and Maine-New Hampshire) used in this assessment. Instead, the NEFSC spring survey age-length key was applied as an alternative and is a potential source of uncertainty in the assessment. The impact and the magnitude of borrowing aging data from offshore to derive inshore indices at age is unknown. However, if there are age and size dependent spatial differences in the availability yellowtail to the surveys this could potentially result in biased age distribution for the terminal year in the model.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}; see Table 8).

  The 7-year Mohn's ρ, relative to SSB, was 0.76 in the 2017 assessment and was 0.29 in 2018. The 7-year Mohn's ρ, relative to F, was -0.38 in the 2017 assessment and was -0.15 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimates of 2018 SSB (SSB_{ρ}=2,125) was outside of the approximate 90% confidence region around SSB (2,325 - 3,308). The 2018 F (F_{ρ}=0.092) however was within the approximate 90% confidence region around F(0.06 - 0.1). A retrospective adjustment was still made for both the determination of stock status and for projections of catch in 2020. The retrospective adjustment changed the 2018 SSB from 2,753 to 2,125 and the 2018 F_{Full} from 0.078 to 0.092.

- Based on this stock assessment, are population projections well determined or uncertain?

  Population projections for Cape Cod-Gulf of Maine yellowtail flounder are uncertain for reasons associated with the retrospective bias in this updated assessment. The 2019 estimates of SSB and yield from this assessment are not within the bounds of projected values in the 2017 operational assessment. The stock is in a rebuilding plan with a rebuilding date of 2023. Based on the 2019 assessment, estimated $S\bar{S}B$ in 2018 is above $SSB_{Threshold}$ but below the SSB_{Target}.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Major changes to the assessment include a revision to the inshore spring and autumn MADMF State surveys to address data inconsistencies. The revision resulted in a time series average difference of 10% and 16% in the spring and autumn respectively. Another major change is the exclusion of hindcasted recruitment estimates for years 1977-1984 used in the projections. The increasingly poor linear relationship between the NEFSC fall survey age-1 and the VPA estimates of age-1 justified a departure from this approach. Instead, the VPA age-1 abundance from 1985-2017 were used in sampling a cumulative distribution function in the projections which provides sufficient historical context and contrast in the time series. The exclusion of the hindcast recruitments resulted in a difference of approximately 2% less age-1 fish, due to exclusion of the the second highest recruitment value generated from the hindcasted estimate in 1980. The impact of excluding the hindcasted recruitment is likely to be inconsequential to the projections.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  In this Cape Cod-Gulf of Maine yellowtail flounder assessment, the model derived

catchability estimate was directly compared with the experimental estimate for use as a diagnostic. Averages of the NEFSC spring and fall survey values were calulated to account for inter-survey variation and also to provide an estimate that could be considered for the start of the calendar year. The catchability corrected average survey biomass for January 2018 (14,110 mt) is approximately 140% higher than that predicted from the VPA model (5,888mt).

The differences in scaling of the January-1 biomass estmates between the catch efficiency experiment and the VPA model could not be fully reconciled due to limitaions in the existing modeling framework used for stock status determination. However, alternative modeling approaches will be developed in the next research track assessment to fully examine the scaling issue and the feasibility of incorporating the catchability estimate directly in the model.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status for Cape Cod-Gulf of Maine yellowtail flounder has changed from being overfished and overfishing occuring to NOT overfished and overfishing NOT occuring. The change in status is supported by an above average estimated 2016 incoming year class coupled with very low exploitatin of the fishery resource. The estimated 2018 catch was the lowest in the time series and approximately 46% of the 2018 ACL.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Cape Cod-Gulf of Maine yellowtail flounder shows no truncation in age structure. There has been some moderate expansion in the older age groups in the catch, as well as the surveys. There is an above average estimated 2016 incoming year class which has contributed to the increase in total biomass. As indicated previously, estimates of commercial landings and discards continue to decline which is consistent with the recent low fishing mortality in the stock. The reductions in fishing mortality and above average 2016 year class has resulted in the stock biomass to increase. However, SSB is projected to decrease in the short-term if fished at $F_{40\%}$.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Cape Cod-Gulf of Maine yellowtail flounder assessment could potentially benefit from updated growth and maturity studies. The current values are based on GARM III estimates (NEFSC 2008) which are apprximately 10 years old. Future modeling efforts should consider forward-projecting statistical catch-at-age models to account for uncertainty in the data inputs. Additionally, investigations to characterize spatial dynamics in age and size dependent distribution of yellowtail any potential implications it may have on the survey catch.

- Are there other important issues?

  None.

6.1 Reviewer Comments: Cape Cod-Gulf of Maine yellowtail flounder

6.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

Although there are major diagnostic problems with the assessment (e.g., major retrospective pattern, apparent problems with estimates of scale, residua patterns), the updated assessment has some improvements from the 2017 update assessment. The assessment shows that the VPA is getting closer to the Bigelow swept-area biomass time series in the most recent 2 years. The retrospective pattern is also improving.

We note that there are uncertainties that are not captured by the current model and that a more generalizable statistical catch-at-age model should be considered in the f 2024 research track assessment. Until then, the uncertainties in the assessment justify a level-3 review for management track updates.

6.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. It is important to recognize that there are other stocks that may be constraining the ability to catch the full quota for this yellowtail stock. This most recent fishing year, only 42% of commercial ACL was caught.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The current assessment indicate that Cape Cod / Gulf of Maine yellowtail flounder is not overfished and overfishing is not occurring. This is a change in status from the previous 2017 assessment update which concluded the stock was both overfished and that overfishing was occurring. The change in status is due to recent low catches allowing the stock to increase with recruitments closer to the time series mean than the low recruitments seen previously.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

Catch has been declining since 2011 and is currently the lowest estimated in the time series. There is some moderate expansion in the catch at age.

The persistence of the retrospective pattern continues to be a source of uncertainty. However, retrospective diagnostics improved by 61% for both F and SSB compared to the 2017 assessment.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the VPA with retrospective adjustments. The Panel supports removal of the hindcast recruitments from the set of recruitments included in the projection series for reasons associated with the increasing poor relationship between VPA estimates of age-1 and the NEFSC age-1 fall survey used in deriving hindcast recruitments values.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

We are expecting that the remaining VPA stocks, which all have research tracks coming up, will be moving away from the VPA. A Lorenzen M is being used for SNE yellowtail, and it may be that a similar M should be considered for this stock. During the research track for all three yellowtail stocks in 2024, a consistent approach to determining natural mortality should be applied across yellowtail flounder stocks.

The length-to-weight conversion should be examined in the future to determine if it has changed over time.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Legault, C, L. Alade, S. Cadrin, J. King, and S. Sherman. 2008. In. 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. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii. CRD08-15

Legault, C, L. Alade, S. Emery, J. King, and S. Sherman. 2012. In. Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 12-06.; 789 p. CRD12-06

Alade, L. 2015. In Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci CentRef Doc. 15-24; 251 p.

Alade, L. 2017. In Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci CentRef Doc. 15-17; 259 p.

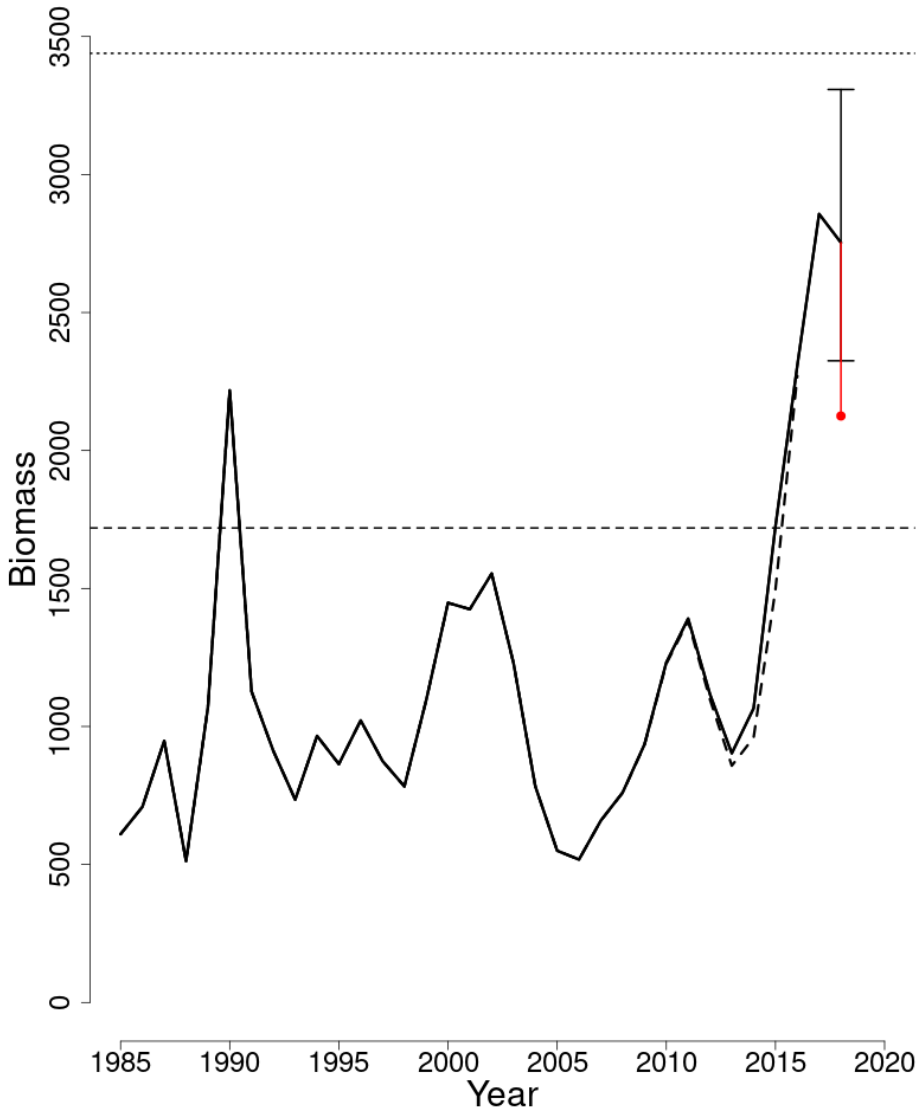

Figure 33: Trends in spawning stock biomass of Cape Cod-Gulf of Maine yellowtail flounder between 1985 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The 90% bootstrap probability intervals are shown.

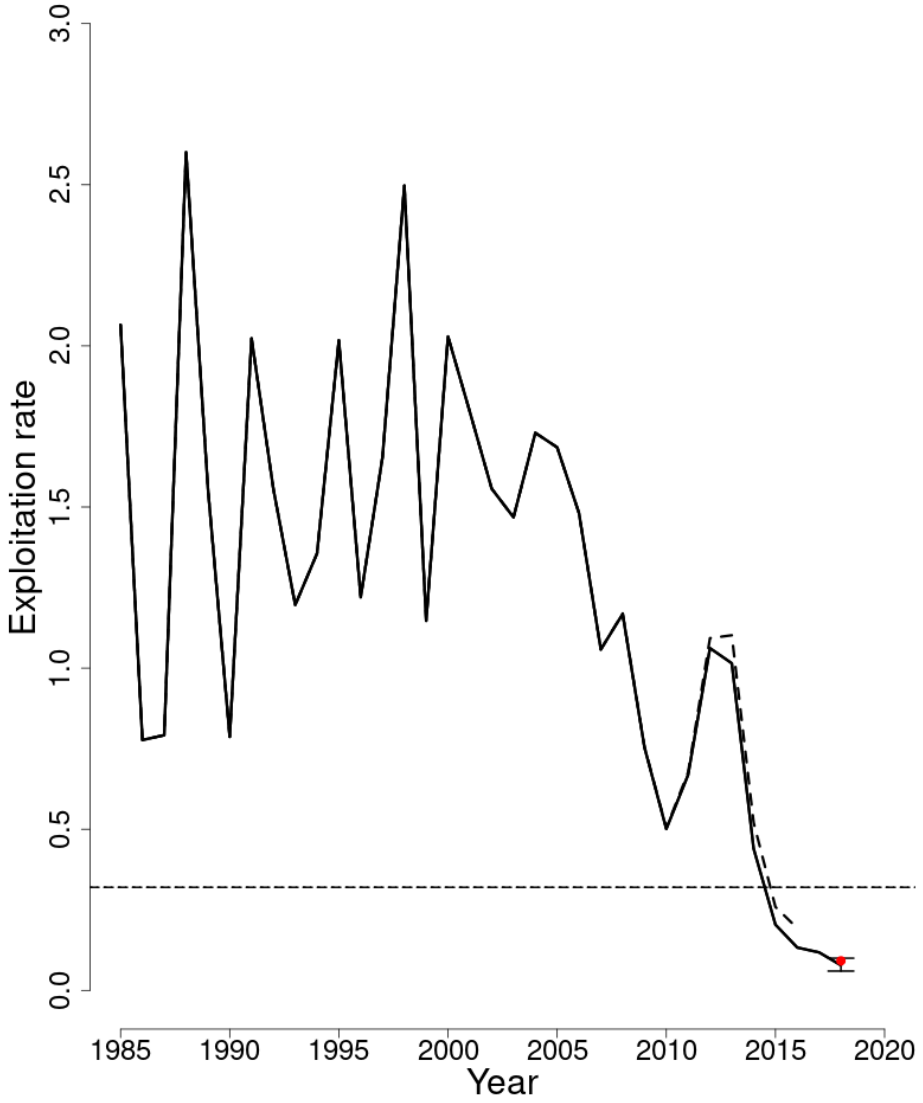

Figure 34: Trends in the fully selected fishing mortality (F_{Full}) of Cape Cod-Gulf of Maine yellowtail flounder between 1985 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.32; horizontal dashed line). F_{Full} was adjusted for a retrospective pattern and the adjustment is shown in red based on the 2019 assessment. The 90% bootstrap probability intervals are shown.

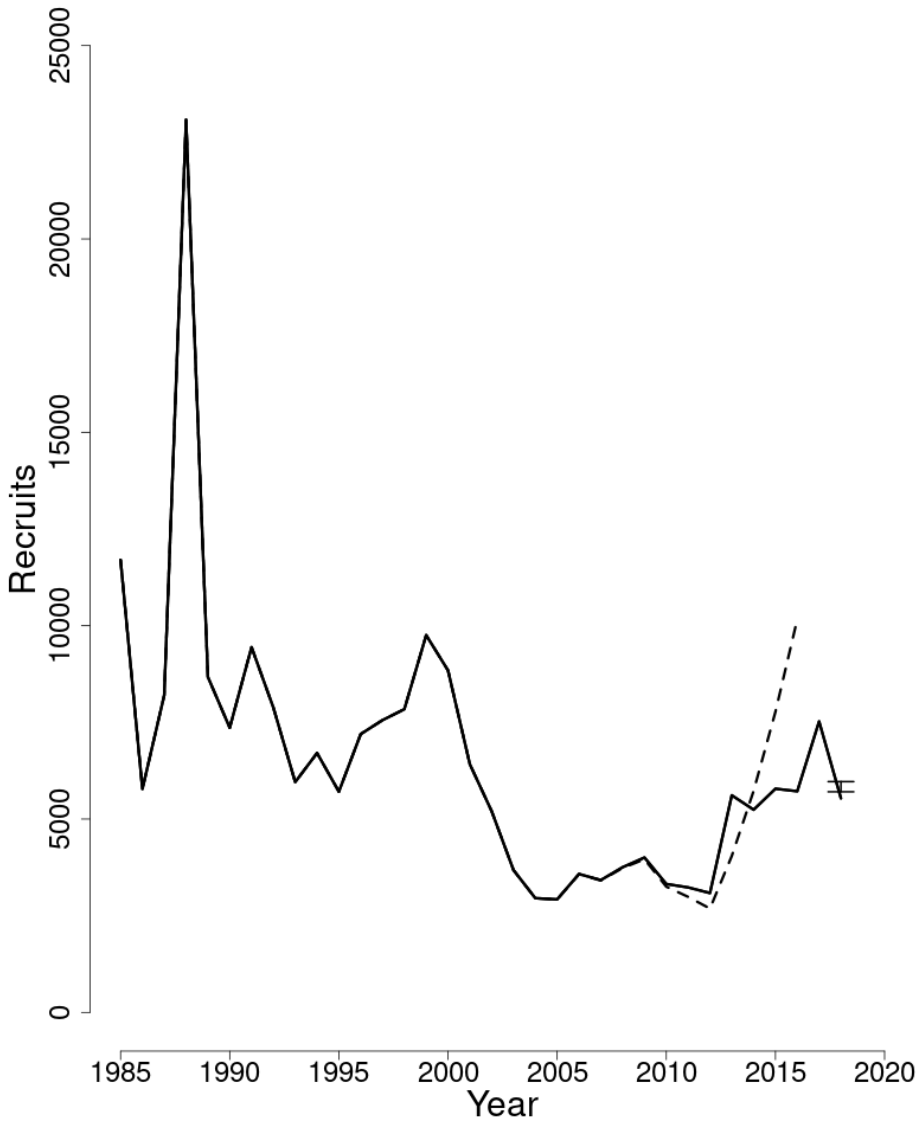

Figure 35: Trends in Recruits (age 1) (000s) of Cape Cod-Gulf of Maine yellowtail flounder between 1985 and 2018 from the current (solid line) and previous (dashed line) assessment. The 90% bootstrap probability intervals are shown.

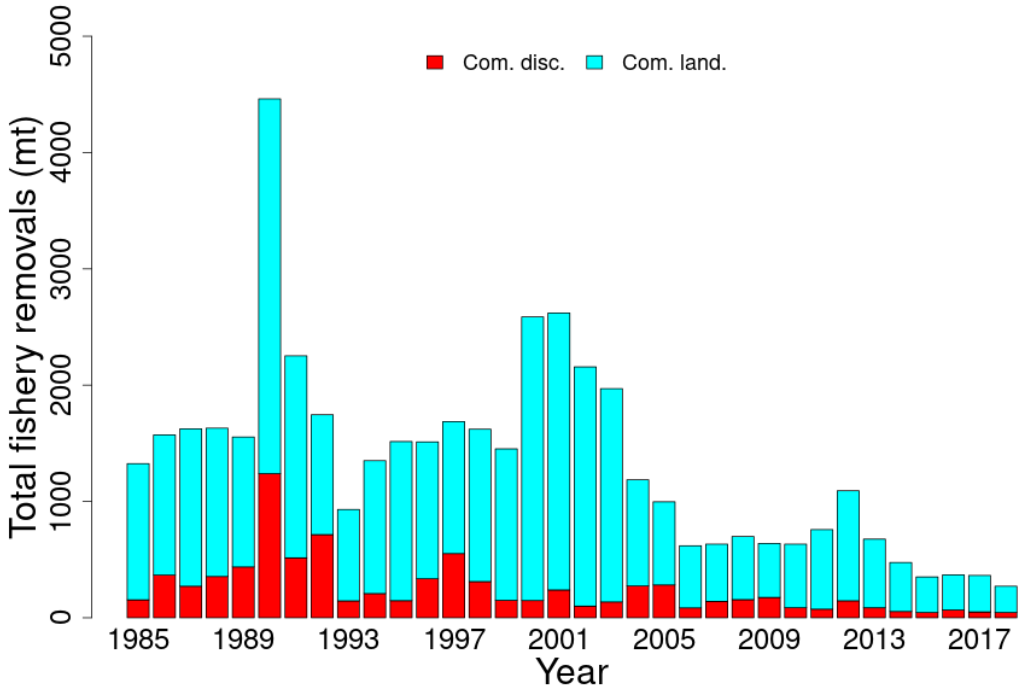

Figure 36: Total catch of Cape Cod-Gulf of Maine yellowtail flounder between 1985 and 2018 by disposition (landings and discards).

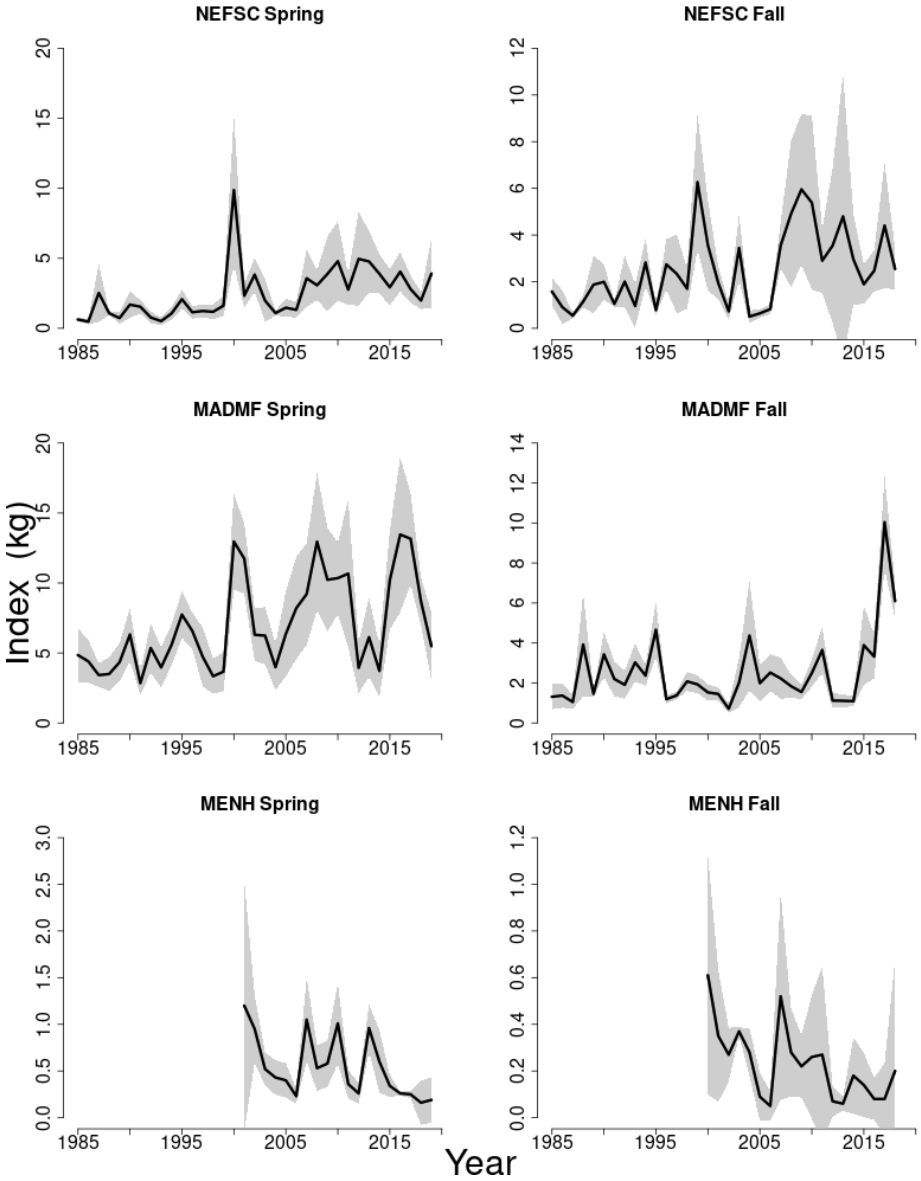

Figure 37: Indices of biomass for the Cape Cod-Gulf of Maine yellowtail flounder between 1985 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys, Massachusetts Division of Marine Fisheries (MADMF) inshore state spring and fall bottom trawl surveys,and the Maine New Hampshire inshore state spring and fall state surveys. The 90% bootstrap probability intervals are shown.

7 Southern New England-Mid Atlantic yellowtail flounder

Larry Alade

This assessment of the Southern New England-Mid Atlantic yellowtail flounder (Limanda ferruginea) stock is an operational assessment update of the existing 2012 benchmark assessment (NEFSC 2012). Based on the last operational assessment (Alade 2017), the stock was overfished and overfishing was occurring. This assessment updates commercial fishery catch data, research survey indices of abundance, weights at age and the analytical ASAP assessment model and reference points through 2018. Additionally, stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, Southern New England-Mid Atlantic yellowtail flounder (*Limanda ferruginea*) stock is overfished and overfishing is not occurring (Figures 38-39). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 90 (mt) which is 5% of the biomass target (SSB_{MSY} *proxy* = 1,779; Figure 38). The 2018 fully selected fishing mortality was estimated to be 0.259 which is 73% of the overfishing threshold proxy (F_{MSY} *proxy* = 0.355; Figure 39).

Table 25: Catch and model results for Southern New England-Mid Atlantic yellowtail flounder. All weights are in (mt) recruitment is in (000s) and F_{Full} is the average fishing mortality on ages (ages 4 and 5). Model results are from the current updated ASAP assessment. Note: Terminal year estimates of SSB and F reflect the unadjusted values for retrospective error.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
				Data						
Commercial discards	268	177	145	221	185	109	53	26	16	8
Commercial landings	185	113	243	342	461	516	284	126	48	11
Total Catch for Assessment	453	291	388	563	646	625	337	152	64	19
				Model Results						
Spawning Stock Biomass	1,645	1,752	1,823	1,831	1,454	956	504	235	135	147
F_{Full}	0.363	0.227	0.307	0.527	0.678	0.811	0.791	0.714	0.522	0.178
Recruitment (age 1)	3,511	3,208	6,326	1,646	1,209	274	125	105	775	905

Table 26: Comparison of reference points estimated in an earlier assessment and from the current assessment update. An $F_{40\%}$ proxy was used for the overfishing threshold and was based on long-term stochastic projections.

	2017	2019
F_{MSY} *proxy*	0.347	0.355
SSB_{MSY} (mt)	1,986	1,779 (993 - 2,725)
MSY (mt)	547	492 (277 - 749)
Median recruitment (age 1) (000s)	7,242	6,562
Overfishing	Yes	No
Overfished	Yes	Yes

Projections: Short term projections of biomass were derived by sampling from an empirical cumulative distribution function of 28 recruitment estimates from the ASAP model results. Following the previous and accepted benchmark formulation, recruitment was based on recent estimates of recruitments from the model time series (i.e. corresponding to year classes 1990 through 2017) to reflect the low recent pattern of recruitment in the stock. The annual fishery selectivity, maturity ogive, and mean weights at age used in projection are the most recent 5 year averages; retrospective adjustments were applied in the projections.

Table 27: Short term projections of total fishery catch and spawning stock biomass for Southern New England-Mid Atlantic yellowtail flounder based on a harvest scenario of fishing at F_{MSY} *proxy* between 2021 and 2022. Catch in 2019 was assumed to be 16 (mt).

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	16	95 (73 - 129)	0.227

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	31 (23 - 41)	111 (84 - 151)	0.355
2021	69 (33 - 127)	405 (112 - 905)	0.355
2022	173 (60 - 339)	878 (288 - 1,636)	0.355

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The persistence of retrospective patterns remains a source of unceratinty in this assessment. This has resulted in a decrease in adult biomass and recruitment and an increase in fishing mortality when more years of data are added. Although the magnitude of these retrospective patterns continues decrease for F and SSB relative to previous assessments (F by 33% and SSB by 36% relative to 2017 OA), rho adjusted projections were still conducted, which resulted in a reduction of starting abundance at age by approximately 61%.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}; see Table 8).

  The 7-year Mohn's ρ, relative to SSB, was 0.98 in the 2017 assessment and was 0.63 in 2018. The 7-year Mohn's ρ, relative to F, was -0.47 in the 2017 assessment and was -0.31 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimates of 2018 SSB (SSB_{ρ}=90) and 2018 F (F_{ρ}=0.259) were outside the approximate 90% confidence region around SSB (113 - 200) and F (0.12 - 0.25).

- Based on this stock assessment, are population projections well determined or uncertain?

  Population projections for Southern New England-Mid Atlantic yellowtail flounder are uncertain for reasons associated with the retrospective bias in this updated assessment. The

2018 estimates of SSB however are well within the bounds of the projected SSB in 2017. In contrast to SSB, total yield in the fishery is not within the bounds of the projected 2017 catch estimates. The stock is in a rebuilding plan with a rebuilding date of 2029. Estimated SSB in 2018 is below the $SSB_{Threshold}$.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  No major changes, other than the addition of recent years of data, were made to the Southern New England-Mid Atlantic yellowtail flounder assessment for this update. However, additional model explorations were carried out to examine the influence of the catchability estimates from the Cooperative Research chain sweep experiment in the ASAP model.

  In this Southern New England-Mid Atlantic yellowtail flounder assessment, experimental catchability-corrected swept area biomass was directly incorporated as the biomass data stream in a series of sensitivity runs (See the supplemental document for additional details.)

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The status of fishing for Southern New England-Mid Atlantic yellowtail flounder has changed since the last 2017 operational assessment from overfishing occurring to overfishing NOT occurring. The biomass stock status however remains unchanged and is still overfished in this update. The 2018 total catch for Southern New England-Mid Atlantic yellowtail flounder was estimated to be the lowest on record at 19mt and approximately 29% of the ACL. The continued decline in total catch of Southern New England-Mid Atlantic yellowtail flounder since the last operational assessment in 2017 and the moderate incoming year class in 2017 and 2018 (but still estimated below average since the 1990's) partly supports the change in the overfishing status. In the short term, SSB is expected to increase, assuming recruitment remains at average levels since the 1990s, but the projected increase is still below the biomass reference point.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Fishing mortality has been declining in recent years and is now below the overfishing reference point. In 2017, the relatively strong incoming year class has resulted in a moderate increase in Spawning Stock Biomass (SSB) in 2018, but remains well below SSBMSY. In the

short term, SSB is projected to increase due to another estimated incoming year class in 2018.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  Recruitment of Southern New England-Mid Atlantic yellowtail flounder continues to be weak compared to the pre-1990's. Should this pattern of poor recruitment continue into the future, the ability of the stock to recover could be compromised. Therefore, future studies should build on current knowledge to further investigate some of the underlying ecological mechanisms of poor recruitment in the stock as it may relate to the physical environment. Recent studies on evaluating environmental effects on Southern New England yellowtail stock productivity suggest that oceanographic features, such as the cold pool and Gulf Stream are likely important predictors of recruitment (Miller et al., 2016; Xu et al., 2017), however the mechanisms driving these predictions are not well known. Other areas of future work should continue to address the retrospective bias, including further work on the sensitivity analyses (i.e. determination of appropriate input data weighting by evaluating the CV and effective sample sizes in the model).

- Are there other important issues?

  None.

7.1 Reviewer Comments: Southern New England-Mid Atlantic yellowtail flounder

7.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

Recruitment continues to be at record lows and estimates of the current stock are 20% of what they were the mid-1990s when it was considered to be collapsed. Trying to conduct a survey or an analytical assessment for a stock in this depleted state is challenging.

A cooperative survey study was carried out to estimate catch efficiency of the NEFSC trawl survey gear, with focus on flatfish. Direct comparison of the base ASAP run with the biomass estimated from the surveys using the results of the catchability study indicated there may be a scaling issue in the ASAP run (the biomass from ASAP was below the biomass from the catchability studies). An ASAP model was run using just the NEFSC surveys with different time series for the Albatross and Bigelow years in order to use the cooperative survey study results in the model. The model fit to the catch data and survey indices from the Albatross time series were generally consistent with the base model. The model fit to the observed Bigelow indices showed strong residual patterning for both the aggregate index and the age composition data. Improvement occurred in the retrospective diagnostics when survey catchability was freely estimated, however catchability estimates were well above 1.00. The examination of model fit with fixed catchability resulted in a worse retrospective pattern compared to the base model.

7.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

Based on this assessments update, it is recommended that Southern New England-Atlantic Yellowtail flounder is overfished and overfishing is not occurring. This is a change in fishing status from the previous assessment which found the stock overfished and undergoing overfishing.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

The continued declining trend in the survey biomass to record low levels, despite reductions in catch to historical low levels, indicates a poor state of the resource. Recruitment continues to be weak.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were made from the ASAP model using numbers at age with retrospective adjustments. Historical recruitment appeared to spike periodically roughly every 10 years in the time series driving stock size, but now appears to be at record lows. The SAW 54 Working Group explored mechanisms such as the cold pool as to what might be causing this. There might be a shrinkage in desirable habitat leading to reduced recruitment in recent years. Projections from the assessment only sample from recent low recruitment because it seems likely that those earlier recruitments, including the spikes, may not be representative any longer.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

The persistence of a pattern in retrospective inconsistency is a source of uncertainty in this assessment. However, this update resulted in an improvement in retrospective diagnostics relative to the 2017 assessment.

The direct application of the catch efficiency results in the ASAP model formulation resulted in poorer model diagnostics and is not recommended for consideration in this update.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Alade, L. 2017. In Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-17; 259 p.

Xu, H., T.J. Miller, S. Hameed, L.A. Alade, J. Nye. (2017). Evaluating the Utility of the Gulf Stream Index for Predicting Recruitment of Southern New England- Mid Atlantic yellowtail flounder. Fisheries Oceanography. DOI: 10.1111/fog.12236

Miller, T.J., J. Hare, L. Alade. 2016. A State-Space approach to Incorporating Environmental Effects on Recruitment in Age-Structured Assessment Model with an Application to Southern New England Yellowtail Flounder. Canadian Journal of Fisheries and Aquatic Sciences. Vol. 73(8): pp1261-1270

Alade, L. 2015. In Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-24; 251 p.

Northeast Fisheries Science Center. 2012. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. US Dept Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 12-18.; 600 p. CRD12-18

Alade, L, C. Legault, S. Cadrin. 2008. In. 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. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15;

884 p + xvii. CRD08-15 Richardson, D, J. Hare, W. Overholtz, D. Johnson. 2009. Development of long-term larval indices for Atlantic herring (*Clupea harengus* on the northeast US continental shelf. ICES Journal of Marine Science, Volume 67, Issue 4, 1 May 2010, Pages 617627.

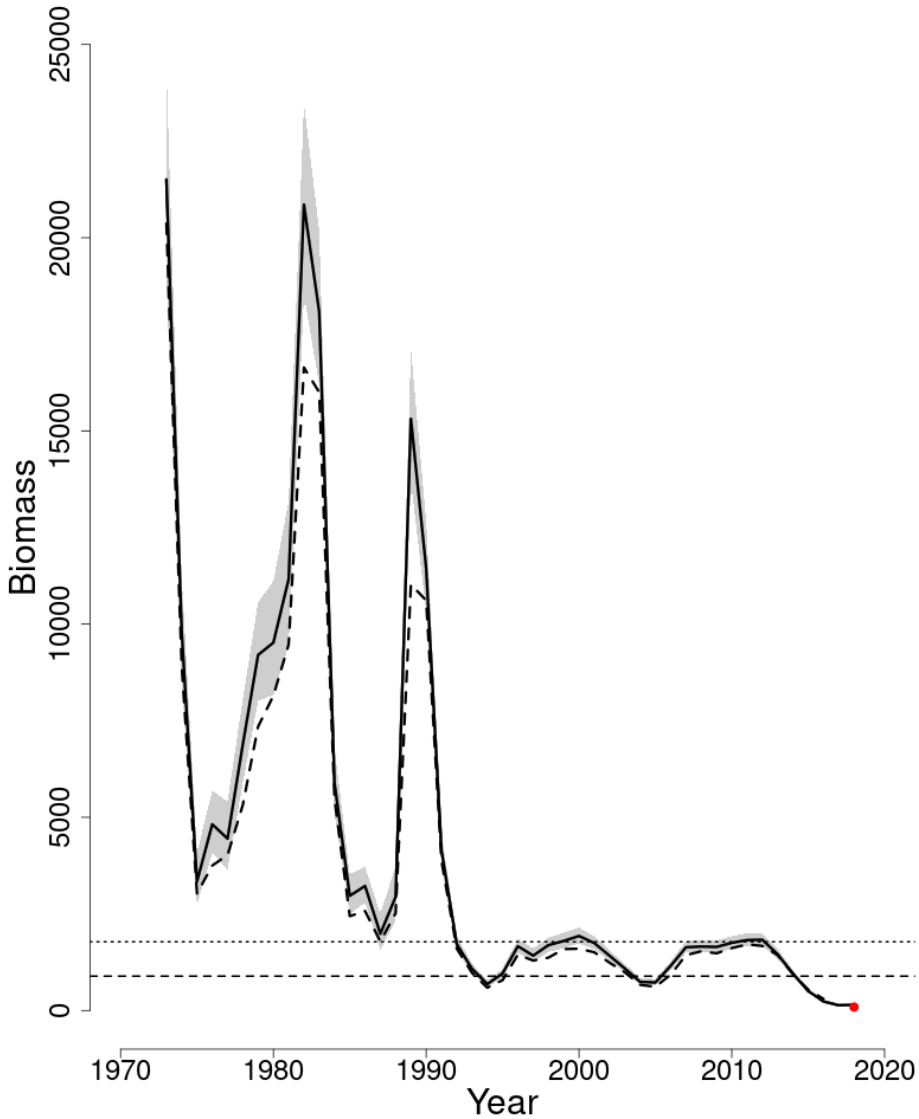

Figure 38: Trends in spawning stock biomass of Southern New England-Mid Atlantic yellowtail flounder between 1973 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% lognormal confidence intervals are shown.

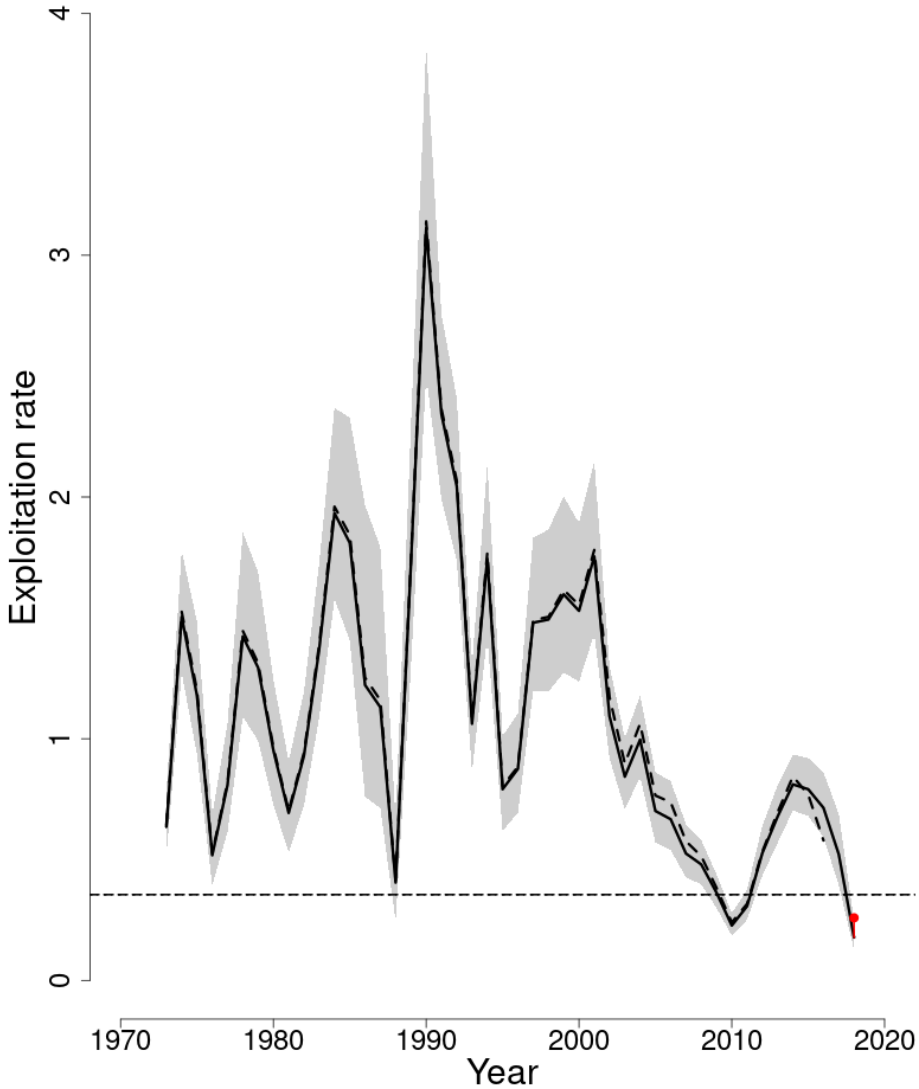

Figure 39: Trends in the fully selected fishing mortality (F_{Full}) of Southern New England-Mid Atlantic yellowtail flounder between 1973 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.355; horizontal dashed line). F_{Full} was adjusted for a retrospective pattern and the adjustment is shown in red based on the 2019 assessment. The approximate 90% lognormal confidence intervals are shown.

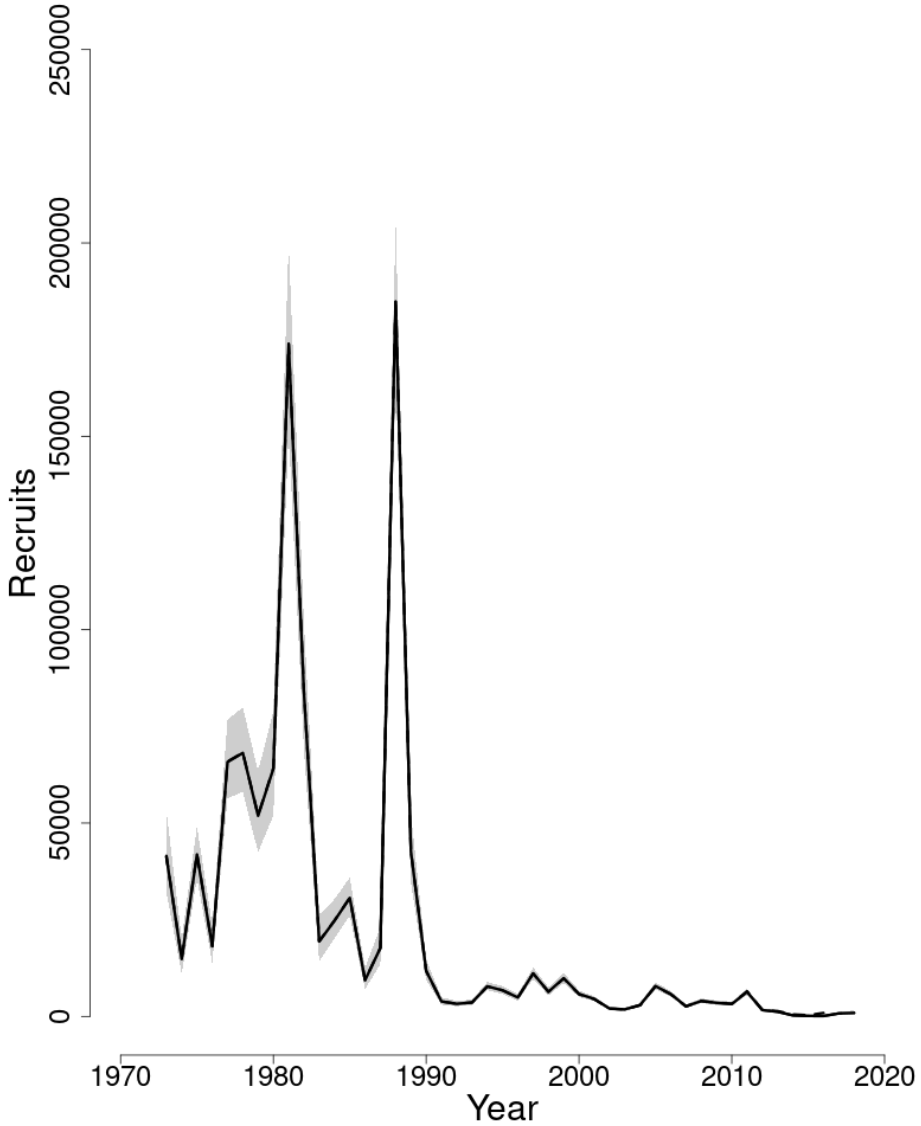

Figure 40: Trends in Recruitment (age 1) (000s) of Southern New England-Mid Atlantic yellowtail flounder between 1973 and 2018 from the current (solid line) and previous (dashed line) assessment. The approximate 90% lognormal confidence intervals are shown.

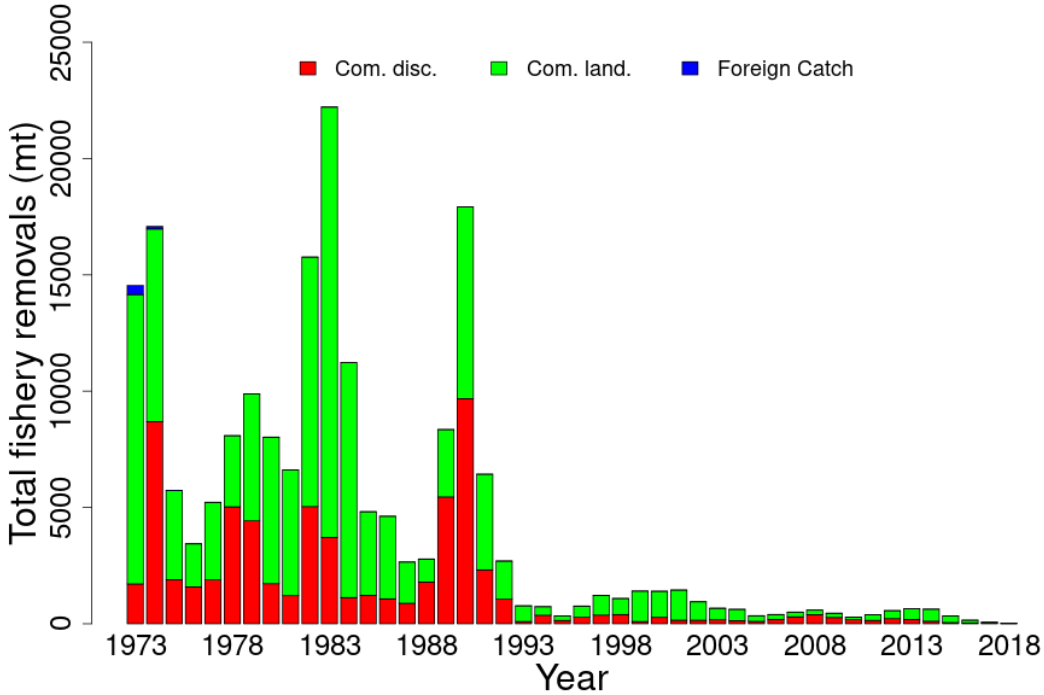

Figure 41: Total catch of Southern New England-Mid Atlantic yellowtail flounder between 1973 and 2018 by fleet (US domestic and foreign catch) and disposition (landings and discards).

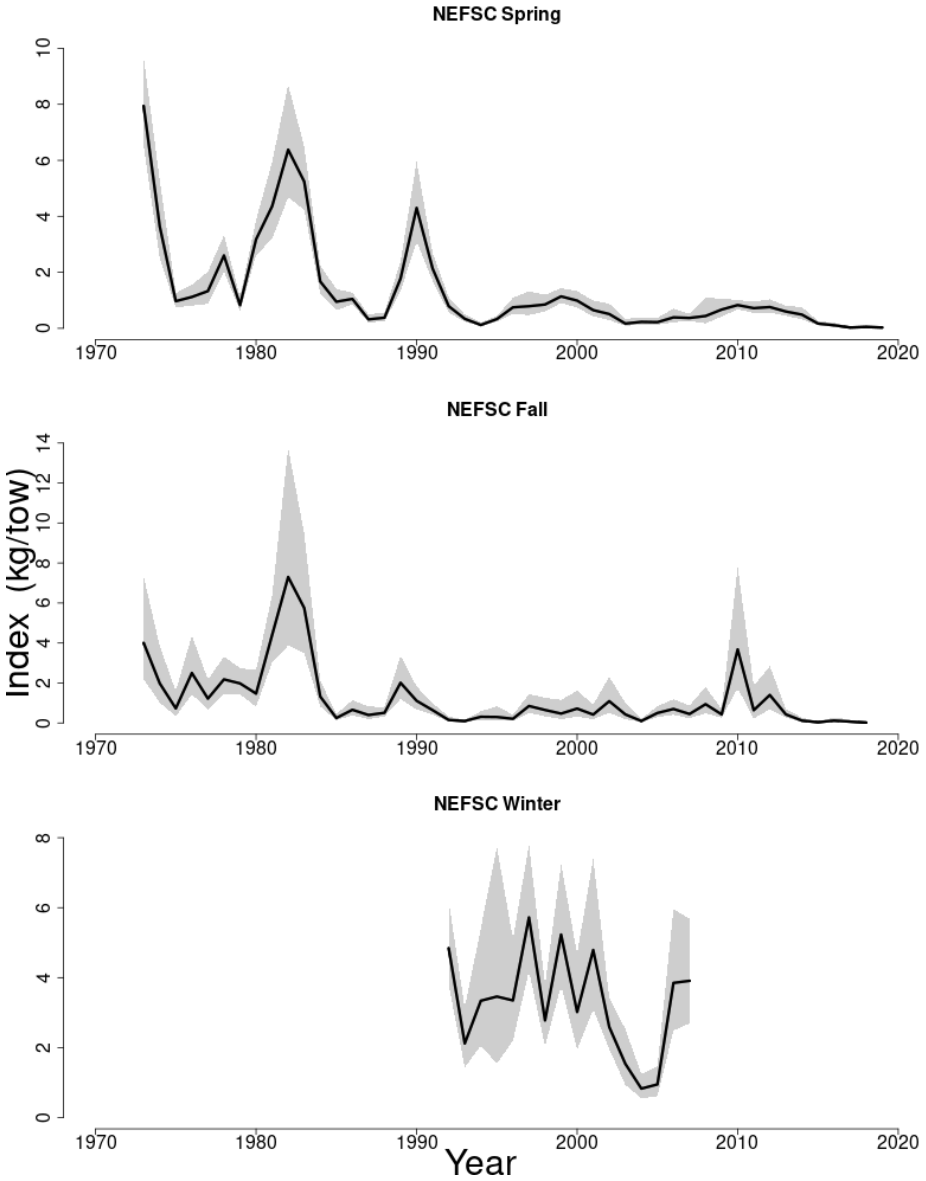

Figure 42: Indices of biomass for the Southern New England-Mid Atlantic yellowtail flounder between 1973 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring, fall and winter bottom trawl surveys. The approximate 90% lognormal confidence intervals are shown. Note: Larval index based on Richardson et al (2009) was also used in this assessment and is available in the supplemental documentation.

8 Georges Bank Winter Flounder

Lisa Hendrickson

This assessment of the Georges Bank Winter Flounder (Pseudopleuronectes americanus) stock is an operational update of the existing 2017 operational VPA assessment which included data for 1982-2016 (NEFSC 2017). Based on the previous assessment the stock was not overfished and overfishing was not ocurring. This assessment updates commercial fishery catch data, research survey biomass indices, and the analytical VPA assessment model and reference points through 2018. Additionally, stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, the Georges Bank Winter Flounder (*Pseudopleuronectes americanus*) stock is overfished and overfishing is not occurring (Figures 43-44). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 2,175 (mt) which is 24% of the biomass target for an overfished stock (SSB_{MSY} = 8,910 with a threshold of 50% of SSBMSY; Figure 43). The 2018 fully selected fishing mortality (F) was estimated to be 0.223 which is 43% of the overfishing threshold (F_{MSY} = 0.519; Figure 44). However, the 2018 point estimate of SSB and F, when adjusted for retrospective error (55% for SSB and -35% for F), is outside the 90% confidence interval of the unadjusted 2018 point estimate. Therefore, the 2018 F and SSB values used in the stock status determination were the retrospective-adjusted values of 0.223 and 2,175 mt, respectively.

Table 28: Catch input data and VPA model results for Georges Bank Winter Flounder. All weights are in (mt), recruitment is in (000s) and F_{Full} is the fishing mortality on fully selected ages (ages 4-6). Catch and model results are only for the most recent years (2009-2018) of the current updated VPA assessment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
				Data						
US landings	1,658	1,252	1,801	1,911	1,675	1,114	866	462	366	417
CA landings	12	45	52	83	12	12	13	4	6	9
US discards	79	110	127	126	46	46	19	5	14	42
CA scall dr discards	240	116	88	79	28	47	42	21	16	22
Catch for Assessment	1,989	1,523	2,069	2,199	1,761	1,219	940	492	402	490
				Model Results						
Spawning Stock Biomass	4,263	4,979	5,243	4,970	4,352	4,363	5,078	4,362	3,952	3,372
F_{Full}	0.462	0.329	0.496	0.486	0.48	0.374	0.17	0.13	0.099	0.145
Recruits (age 1)	12,091	6,276	5,942	4,455	3,205	4,275	1,806	2,041	2,969	532

Table 29: Comparison of reference points estimated in the 2017 assessment and the current assessment update and stock status during 2016 and 2018, respectively. An estimate of F_{MSY} was used for the overfishing threshold and was based on long-term stochastic projections of the parameter and variance estimates from a Beverton-Holt stock-recruit model. SSBMSY was used as the biomass target and was also based on long-term stochastic projections which included the 2014-2018 means for selectivity-, maturity- and mean weights-at-age.

	2017	2019
F_{MSY}	0.522	0.519
SSB_{MSY} (mt)	7,600	8,910 (4,196 - 21,143)
MSY (mt)	3,500	4,260 (2,049 - 9,632)
Median recruits (age 1) (000s)	9,164	8,608
Overfishing	No	No
Overfished	No	Yes

Projections: Short-term projections of biomass were derived by sampling from a cumulative distribution function of recruitment estimates (1982-2017 YC) from the final run of the ADAPT VPA model. The annual fishery selectivity, maturity ogive (a 3-year moving window), and mean weights-at-age used in the projection are the most recent five-year averages (2014-2018). An SSB retrospective adjustment factor of 0.643 was applied in the projections.

Table 30: Short-term projections of catch (mt) and spawning stock biomass (mt) for Georges Bank Winter Flounder based on a harvest scenario of fishing at F_{MSY} between 2020 and 2022. Catch in 2019 was assumed to be 334 (mt)

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	334	2,113 (1,597 - 2,806)	0.176

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	790	1,614 (1,180 - 2,243)	0.519
2021	868	1,592 (1,132 - 2,746)	0.519
2022	1,422	2,895 (1,326 - 4,045)	0.519

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The largest source of uncertainty is the estimate of natural mortality, which is based on longevity (max. age = 20). Natural mortality is not well studied in Georges Bank Winter Flounder and is assumed to be constant over time. Natural mortality affects the scale of the biomass and fishing mortality estimates. Other sources of uncertainty include the underestimation of catches. Discards from the Canadian bottom trawl fleet were not provided

by the CA DFO and the precision of the Canadian scallop dredge discard estimates, with only 1-2 trips per month, are uncertain. The lack of age data for the Canadian spring survey catches requires the use of the US spring survey age-length keys despite selectivity differences. In addition, there are no length or age composition data for the Canadian landings or discards of GB winter flounder. The steepness parameter used to estimate FMSY was inestimable and consequently was to be fixed (0.78).

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the 90% confidence intervals for SSB and F_{Full}; see Table 8).

  The 7-year Mohn's ρ, relative to SSB, was 0.540 in the 2017 assessment and was 0.555 in 2018. The 7-year Mohn's ρ, relative to F, was -0.308 in the 2017 assessment and was -0.347 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimates of 2018 SSB (SSB_{ρ}=2,175) and 2018 F (F_{ρ}=0.223) were outside the 90% confidence limits for SSB (2,725 - 4,346) and F (0.111 - 0.194). A retrospective adjustment was made for both the determination of stock status and for projections of catch in 2020. The retrospective adjustment changed the 2018 SSB from 3,372 to 2,175 and the 2018 F_{Full} from 0.145 to 0.223.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Georges Bank Winter Flounder were reasonably well determined and confidence bounds for projected biomass estimates from the current assessment were narrower than the confidence bounds of the biomass estimates from the previous assessment. This stock was required to be rebuilt by 2017, but this did not occur. The stock is in a revised rebuilding plan, based on fishing at 70% of FMSY, with rebuilding by 2029.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  Changes made to the Georges Bank Winter Flounder assessment included updating the most recent five-year averages (2014-2018) of fishery selectivity-, proportion mature-, stock weights-, catch weights-, and spawning stock weights-at-age. In addition, U.S. otter trawl discards were updated from 1964-2016, to correct a codend mesh size binning error in the SAS base code. This error related to binning of observer trips into small-mesh (< 5.5 in.) and large-mesh (>= 5.5 in.) fleet categories. Updating of the 1989-2016 discards were estimated using Northeast Fishery Observer Program (NEFOP) data and this required updating of the hindcast discards because the latter discards were computed from the NEFOP discard estimates. The updated otter trawl discard estimates included in the VPA model for 1982-2016 were 13% lower on average than the discard estimates included in the 2017 VPA model run. However, changes in the updated discard amounts were not unidirectional, and during this time period, total discards (US and CA, all gear types) comprised a small percentage (12% on average) of the total catch.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b;

Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

The catch efficiency studies were not focused on this stock and were not applicable to the 2019 assessment of Georges Bank Winter Flounder. As a result, the winter flounder length composition from the studies does not reflect the length composition of the Georges Bank stock (i.e., the studies included few fish > 38 cm total length).

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status of Georges Bank Winter Flounder has changed from not overfished and overfishing is not occurring to overfished and overfishing is not occurring. Although fishing mortality rates were at the lowest levels of the time series during 2015-2018, SSB remained near the SSBMSY threshold (4,455 mt) during 2004-2015 and then declined to the lowest level on record in 2018 (3,372 mt). As in the previous assessment, it was necessary to adjust the 2018 F and SSB point estimates for retrospective error. Mohn's rho values for both F and SSB were similar to the rho values from the previous assessment.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Fishing mortality was at or slightly below FMSY during 2011-2013, then declined rapidly and reached the lowest level of the time series in 2017. The 2018 fishing mortality rate (0.145) was only slightly higher. Following a decline in the catch mean weights-at-age for older fish (ages 4-7+), during 2007-2014, mean weights for these ages increased during 2015-2018. The mean length and weight of fish caught in the NEFSC fall and spring bottom trawl surveys increased during 2008-2014 and 2009-2017, respectively, but have decreased since then. Spawning stock biomass estimates were near the SSBthreshold during 2004-2015, but then decreased and reached the lowest level of the time series in 2018 (3,372 mt). Recruitment declined rapidly during the last decade, from about 13 million fish in 2008 to a time series low of 532,000 fish in 2018. Recruitment increased in 2019 and was similar to the 2017 value (about 3 million fish), but the 2019 estimate is uncertain because it is based solely on the geometric mean of age 1 stock numbers during 2011-2017.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Georges Bank Winter Flounder assessment could be improved with discard estimates from the Canadian bottom trawl fleet and age data from the Canadian spring bottom trawl surveys.

- Are there other important issues?

  None.

8.1 Reviewer Comments: Georges Bank Winter Flounder

8.1.1 Review Panel Summary

The completed Plan A operational assessment is appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference. However, the Panel is concerned about the reference point definitions and recruitment assumptions in projections for Georges Bank winter flounder (FLWGB). Specifically, using a fixed steepness value, as requested by the SARC 52 Review Panel, in the stock-recruitment relationship for defining the MSY reference points is probably not appropriate. Many problems exist with the assumptions here. The assumed F_{MSY} is much less conservative for maintaining spawning potential than the $F_{40\%MSP}$ used to define overfishing for other groundfish stocks. The steepness value cannot be independently estimated for the Georges Bank winter flounder stock. The 2011 benchmark assessment assumed a steepness value that generally fit the information available and was not significantly different from southern New England/Mid Atlantic winter flounder stock. However, the stock-recruitment relationship for Georges Bank winter flounder has deteriorated since then. The stock-recruitment fit has a strong pattern of residuals for the most recent nine years given the fixed steepness value, resulting in predictions much greater than VPA estimates of recruitment in the most recent nine years. The growth rate and maximum size of Georges Bank winter flounder is also much greater than that of the Southern New England/Mid-Atlantic winter flounder stock that constrained the steepness value, which is probably not appropriate. Similar to the GARM 2008 BRPs, an $F_{40\%MSP}$ F_{MSY} proxy might be a more stable and reliable estimator and this should be explored in the next management track assessment as an alternative to the approach used here.

The residual pattern in the stock-recruitment relationship indicates that recent recruitment has been weaker than expected. Projections that assume the either long-term recruitment or a stock-recruitment relationship suggest relatively rapid rebuilding, which may not be realistic for the current stock conditions. The Panel recommends that alternative projections should be considered that assume future recruitment will be similar to recent recruitment.

The Georges Bank winter flounder stock was assessed using a VPA and including a retrospective adjustment for stock status determination and catch projections.

8.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. The Panel approves the use of the corrections to the otter trawl discards to remove a previous error.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used, but with several caveats that have implications for short-term projections and rebuilding targets.

F_{MSY} as presented in the assessment report uses a Beverton-Holt stock-recruitment model with data for 1983-2018 (1982-2017 YC) and steepness fixed at 0.78 and using the most recent 5-yr means of stock weights, catch weights, selectivity and proportion mature-at-age

The B-H model with this assumed steepness value, however, does not satisfactorily characterize recruitment for this stock, thus influencing the quality of the F_{MSY} estimate due to the high dependence of F_{MSY} on steepness when a B-H model is used. This should be examined more closely for future assessments.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is overfished and overfishing is not occurring. A rebuilding plan was developed after the last assessment because it was approaching an overfished condition (e.g., the 2017 stock projection was at the overfished threshold). The updated assessment indicates this projection was realized.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

All three surveys indicate low current biomass with no signs of incoming strong recruitment.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the VPA with retrospective adjustments. However, recruitment has been at record lows over the last decade and projections that include the distribution of long-term recruitment or the fit B-H stock recruitment relationship with assumed steepness in the projections show population growth under any F_{MSY} scenario and thus will be overly optimistic if weak recruitment continues. The Panel notes that the current F_{MSY} is much greater than $F_{40\%MSP}$, the standard F_{MSY} proxy for groundfish, and the target 70% of F_{MSY} in the rebuilding plan is also greater than $F_{40\%MSP}$, which suggests that 70% of F_{MSY} is not be an appropriate rebuilding target.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

The Panel notes poor tracking of cohorts in any of the data streams, making a VPA less suitable as a stock assessment model and suggests that changing to a statistical catch-at-age or state-space model at the next available opportunity would be appropriate.

The Panel suggests explorations regarding the source of the retrospective pattern and recent poor recruitment for this stock.

The Panel supports the decision to not use the results of the chain sweep survey catchability studies for this stock, because the size range of winter flounder in the study does not sufficiently overlap with the size range caught in the NEFSC surveys. Information from other efficiency studies completed by the Northeast Trawl Advisory Panel and more directed experiments on Georges Bank for winter flounder could be conducted to allow appropriate calibration factors to be estimated for this stock.

As indicated in the discussion above, the Panel could not accept the model projections as valid and questioned the usefulness of the proxy F on the projections. However, the base model itself was useful for synthesizing the different pieces of information available and so should be kept for the purposes of setting the context for management recommendations for the upcoming ABC setting exercises. The stock continues to be in poor condition. It will be important for all of the analytical procedures used here and developed through previous working group and peer review committees to be revised and the present and future status of the stock reanalyzed. Therefore, the Panel recommends this stock go through a Level 3 review at next years management track assessment to allow changing the basis of projections and the reference points.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Sameoto, J., B. Hubley, L. Van Eeckhaute and A. Reeves. 2013. A Review of the standardization of effort for the calculation of discards of Atlantic cod, haddock and yellowtail flounder from the 2005 to 2011 Canadian scallop fishery on Georges Bank. Transboundary Resources Assessment Committee (TRAC) Reference Document 2013/04, 22 pp.

Northeast Fisheries Science Center. 2017. Operational assessment of 20 Northeast groundfish stocks, updated through 2016. U.S. Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 17-17; 259 p.

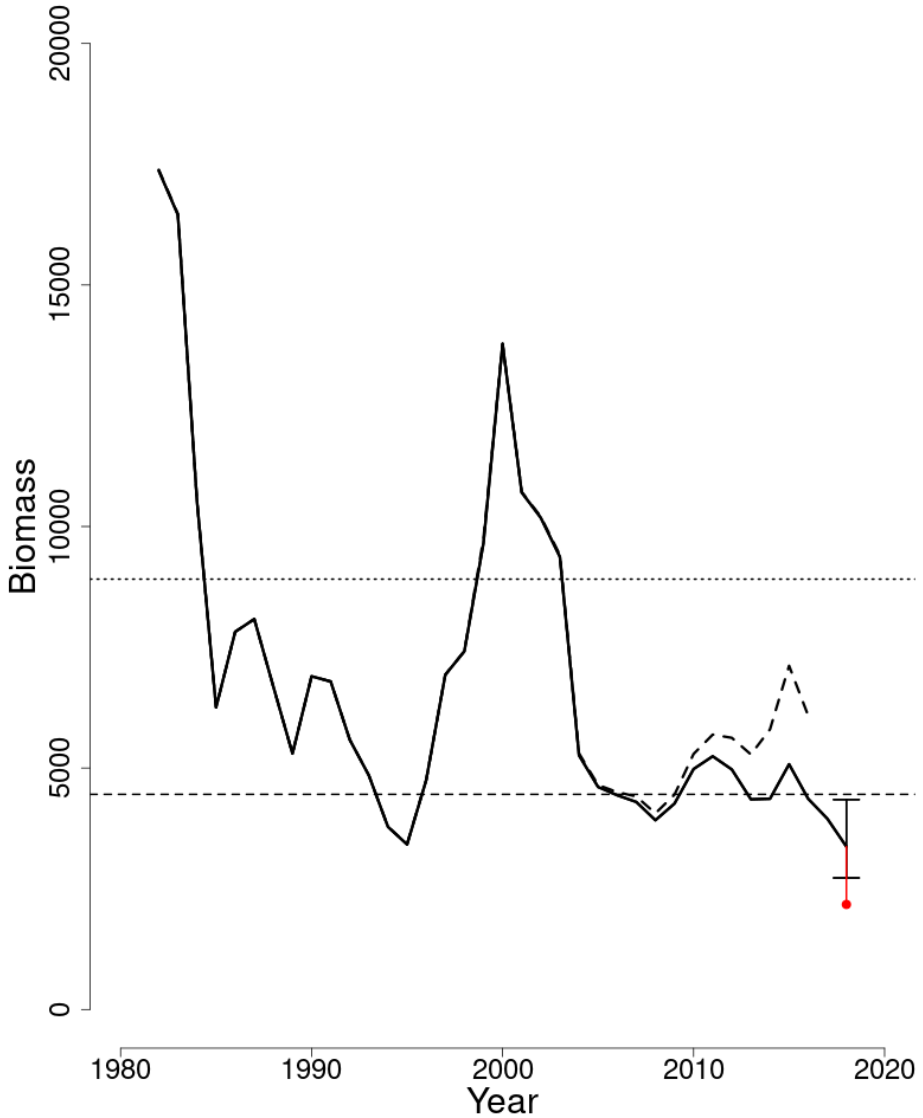

Figure 43: Trends in spawning stock biomass (mt) of Georges Bank Winter Flounder between 1982 and 2018 from the current (solid line) and previous (dashed line) assessments and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY}; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY}; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The 90% normal confidence interval is shown for 2018.

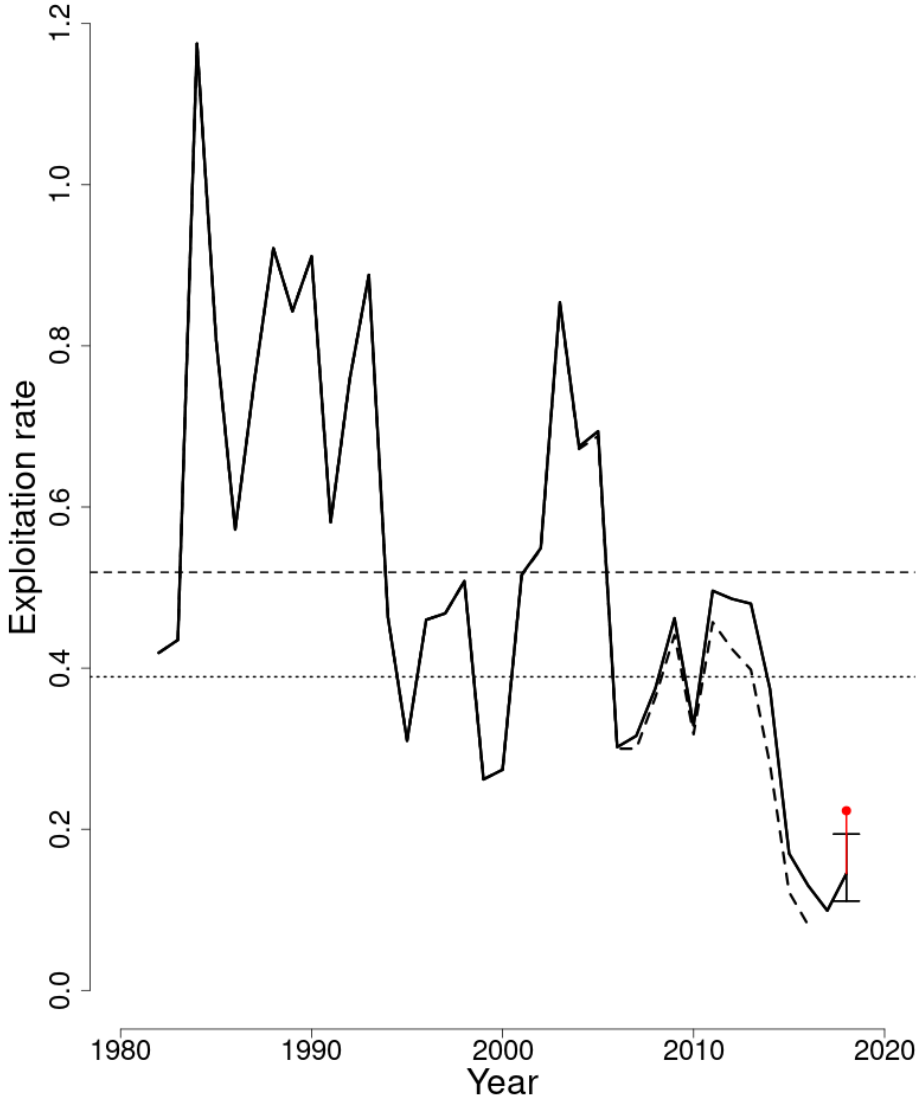

Figure 44: Trends in fully selected fishing mortality (F_{Full}) of Georges Bank Winter Flounder between 1982 and 2018 from the current (solid line) and previous (dashed line) assessments and the corresponding $F_{Threshold}$ (F_{MSY}=0.519; horizontal dashed line) as well as (F_{Target}= 75% of FMSY; horizontal dotted line). F_{Full} was adjusted for a retrospective pattern and the adjustment is shown in red. The 90% normal confidence interval is shown for 2018.

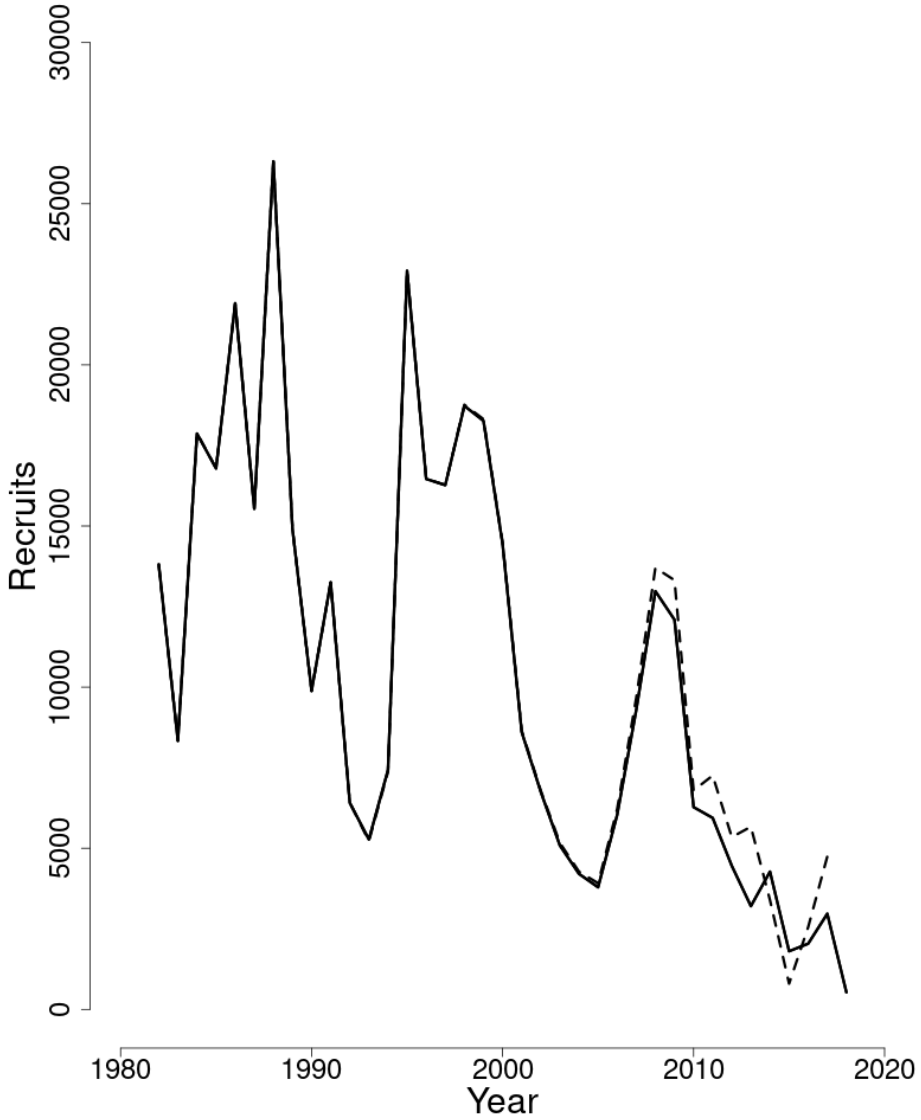

Figure 45: Trends in Recruits (age 1) (000s) of Georges Bank Winter Flounder between 1982 and 2018 from the current (solid line) and previous (dashed line) assessments.

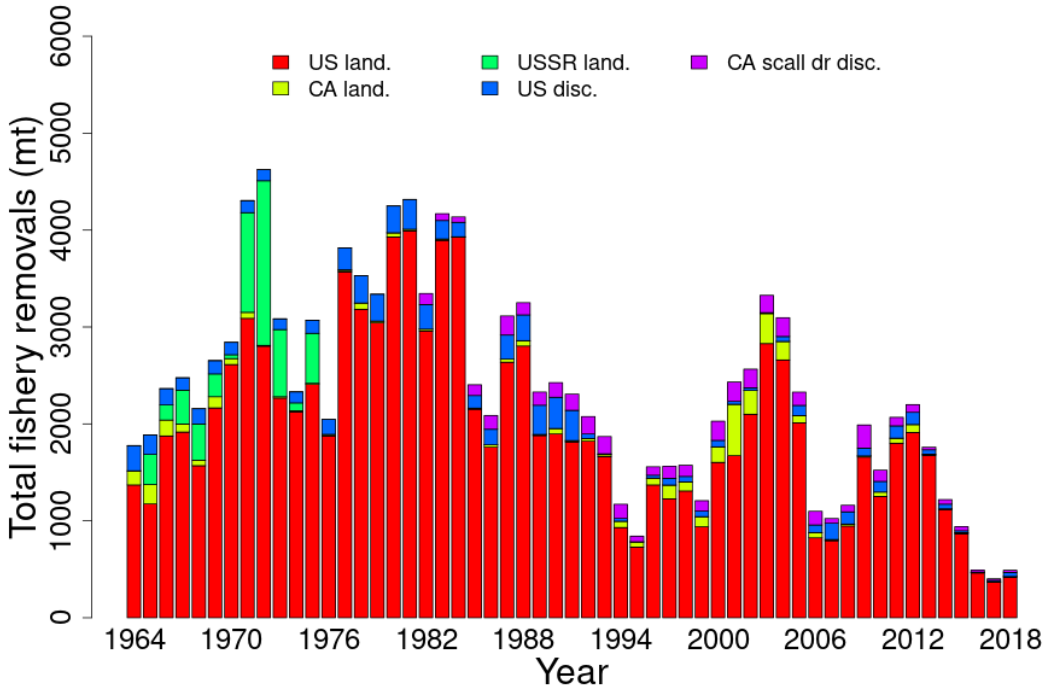

Figure 46: Total catches (mt) of Georges Bank Winter Flounder between 1982 and 2019 by country and disposition (landings and discards).

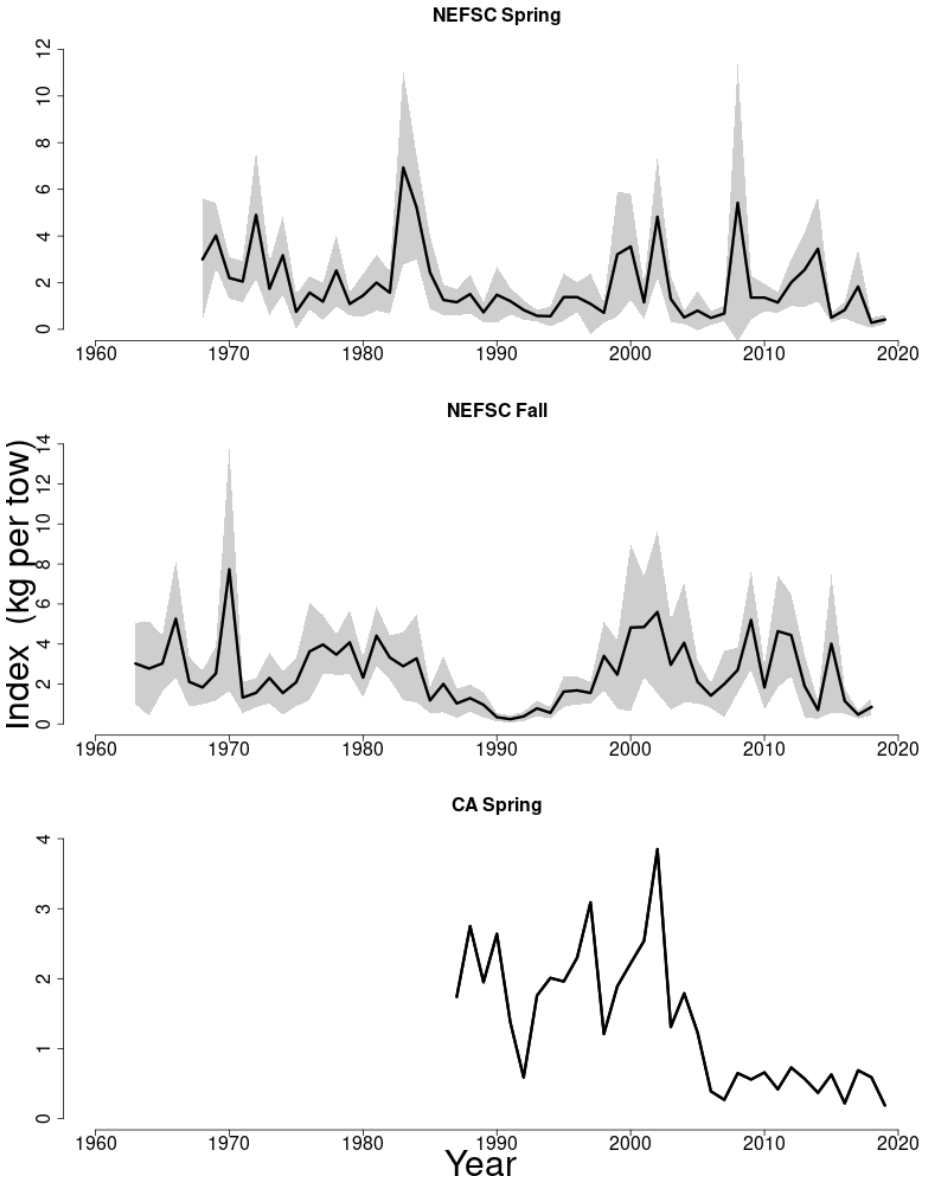

Figure 47: Indices of biomass for the Georges Bank Winter Flounder for the Northeast Fisheries Science Center (NEFSC) spring (1968-2019) and fall (1963-2018) bottom trawl surveys and the Canadian DFO spring survey (1987-2019). The 90% normal confidence interval is shown.

9 Gulf of Maine-Georges Bank American plaice

Larry Alade

This assessment of the Gulf of Maine-Georges Bank American plaice (Hippoglossoides platessoides) stock is an operational update of the existing 2012 benchmark assessment (O'Brien et al. 2012). Based on the previous assessment the stock was not overfished, and overfishing was not ocurring. This 2019 assessment updates commercial fishery catch data, research survey indices of abundance, the analytical VPA assessment model, and reference points through 2018. Additionally, stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, the Gulf of Maine-Georges Bank American plaice (*Hippoglossoides platessoides*) stock is not overfished and overfishing is not occurring (Figures 48-49). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 17,748 mt which is 116% of the biomass target for this stock (SSB_{MSY} *proxy* = 15,293; Figure 48). The 2018 fully selected fishing mortality was estimated to be 0.089 which is 34% of the overfishing threshold proxy (F_{MSY} *proxy* = 0.258; Figure 49).

Table 31: Catch and model results for Gulf of Maine-Georges Bank American plaice. All weights are in (mt), recruitment is in (000s), and F_{Full} is the fishing mortality on fully selected ages (ages 6-9). Model results are unadjusted values from the current updated VPA assessment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
				Data						
GM Commercial landings	866	901	771	762	764	738	828	718	871	911
GM Commercial discards	115	239	96	161	88	36	42	60	72	71
GB Commercial landings	501	492	595	699	528	498	400	287	259	171
GB Commercial discards	274	152	102	123	64	53	44	40	23	39
SNE landings	13	11	3	1	5	3	2	3	1	0
CA landings	0	0	1	0	0	0	0	0	0	0
Catch for Assessment	1,770	1,795	1,569	1,747	1,449	1,328	1,316	1,108	1,226	1,192
				Model Results						
Spawning Stock Biomass	10,258	10,539	10,884	10,893	11,304	13,164	15,202	20,124	24,167	22,490
F_{Full}	0.256	0.185	0.159	0.179	0.138	0.096	0.091	0.064	0.062	0.071
Recruits (age 1)	13,607	13,225	18,368	17,881	24,748	52,719	8,818	17,651	4,785	37,810

Table 32: Comparison of reference points estimated in the previous assessment and from the current assessment update. An $F_{40\%}$ proxy was used for the overfishing threshold and SSB_{MSY} *proxy* was based on long-term stochastic projections.

	2017	2019
F_{MSY} *proxy*	0.216	0.258
SSB_{MSY} (mt)	13,503	15,293 (11,706 - 20,432)
MSY (mt)	2,942	3,301 (2,531 - 4,386)
Median recruits (age 1) (000s)	21,969	22,414
Overfishing	No	No
Overfished	No	No

Projections: Short term projections of biomass were derived by sampling from an empirical cumulative distribution function of 38 recruitment estimates from VPA model results. The annual fishery selectivity, maturity ogive, and mean weights at age used in projections are the most recent 5 year averages; retrospective adjustments were applied in the projections.

Table 33: Short term projections of total fishery catch and spawning stock biomass for Gulf of Maine-Georges Bank American plaice based on a harvest scenario of fishing at F_{MSY} *proxy* between 2020 and 2022. Catch in 2019 was assumed to be 1,131 (mt).

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	1,131	18,954 (16,193 - 22,104)	0.066

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	4,078	17,768 (15,107 - 20,640)	0.258
2021	3,543	15,873 (13,107 - 21,233)	0.258
2022	3,364	15,414 (11,704 - 28,619)	0.258

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Sources of uncertainty in this assessment are the estimates of historical landings at age, prior to 1984, and the magnitude of historical discards, prior to 1989. Both of these affect the scale of the biomass and fishing mortality estimates, and influence reference point estimations. Retrospective patterns also remain a source of uncertainty in the assessment. This has persisted for a number of years causing a decrease in estimates of adult biomass and recruitment and increased estimates of fishing mortality when more years of data are added. However, the magnitude of retrospective biases in this assessment were notably reduced by approximately 35% for both fishing mortality and adult biomass when compared to the previous 2017 operational assessment. Despite the improvement in retrospective bias, the

rho adjusted projections were still conducted which reduced starting numbers at age by an average of 36% (Note that the rho adjustment for the projections are based on numbers at age with Mohns rho ranging from 16-81 %).

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full} see Table 8).

  The 7-year Mohn's ρ, relative to SSB, was 0.35 in the 2017 assessment and was 0.27 in 2018. The 7-year Mohn's ρ, relative to F, was -0.33 in the 2017 assessment and was -0.20 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimates of 2018 SSB (SSB_{ρ}=17,748) and 2018 F (F_{ρ}=0.089) were outside the approximate 90% confidence regions around SSB (19,592 - 26,220) and F (0.063 - 0.084). A retrospective adjustment was made for both the determination of stock status and for projections of catch in 2020. The retrospective adjustment changed the 2018 SSB from 22,490 to 17,748 and the 2018 F_{Full} from 0.071 to 0.089.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for Gulf of Maine-Georges Bank American plaice are reasonably well determined.The stock is in a rebuilding plan, but based on the 2019 assessment, the stock is now considered rebuilt.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  The only major change made to the Gulf of Maine-Georges Bank American plaice assessment was the exclusion of the MassDMF inshore state survey resulting in a considerable improvement in model diagnostics including a reduction in the CV's for the terminal year plus one abundances, reduction in the magnitude of the residuals as well in the retrospective bias. The exclusion of the MassDMF survey resulted in an upward scaling of the VPA total biomass estimates by 12-89% between 2010 and 2019.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  In this Gulf of Maine-Georges Bank American plaice assessment the model's derived catchability estimate was direcrtly compared with the experimental estimate for use as a diagnostic. Averages of the NEFSC sprng and fall survey values were calculated to account for inter-survey variation and also to provide an estimate that could be considered for the start of the calendar year. The 2018 VPA predicted January-1 biomass (25,475 mt) was only 5% lower than the average survey biomass (26,8440 mt) and well within the confidence

bounds of the chain sweep study biomass estimates. This suggests that there is some consistency between the VPA model and the chain sweep study results.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  As in recent assessments for Gulf of Maine-Georges Bank American plaice the stock status remains as not overfished and overfishing not occurring.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The current fishing mortality rate is relatively low, and so recent above average recruitment has resulted in an increase in SSB. SSB is projected to decrease in the short term.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  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.

- Are there other important issues?

  A difference in growth between GM and GB fish has been documented, however, historical catch data for GB may not be sufficient to conduct a separate assessment. Also, the growth difference may not persist in the most recent years. This could be explored further in a research track assessment.

9.1 Reviewer Comments: Gulf of Maine-Georges Bank American plaice

9.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference. Swept area biomass estimates derived from sweep study research were similar in magnitude to the VPA model biomass estimates, supporting the use of the VPA model despite a retrospective pattern for American plaice (PLA).

9.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment.

The Massachusetts DMF survey was excluded from this assessment due to concerns that the declining trend may reflect a movement of the stock offshore instead of decline in the population itself. Many fish species have undergone a gradual shift to deeper waters (e.g., Nye et al. 2009 Mar. Ecol. Prog. Ser. 393:111-129), thus larger plaice may be shifting out of the coastal area off of Massachusetts. Exclusion of the MA DMF survey resulted in higher biomass estimates that are more consistent with those from the area-swept survey estimates. The Panel supports the exclusion of the MA DMF survey in the VPA model.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use with the retrospective adjustments.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is not overfished and overfishing is not occurring. The stock biomass exceeds the estimated SSB_{MSY}, so the stock should be considered rebuilt.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

Surveys indicate relatively high biomass recently will full age distributions and some recent high recruitments.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the VPA with retrospective adjustments.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

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 MA DMF survey data might be used given the concern with movement of the stock offshore.

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.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

O'Brien, L. and J. Dayton (2012). E. Gulf of Maine - Georges Bank American plaice Assessment for 2012 in Northeast Fisheries Science Center, 2012, Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p. http://www.nefsc.noaa.gov/publications/crd/crd1206/.

O'brien, L. 2015. In Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci CentRef Doc. 15-24; 251 p.

Terceiro, M. 2017. In Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci CentRef Doc. 15-17; 259 p.

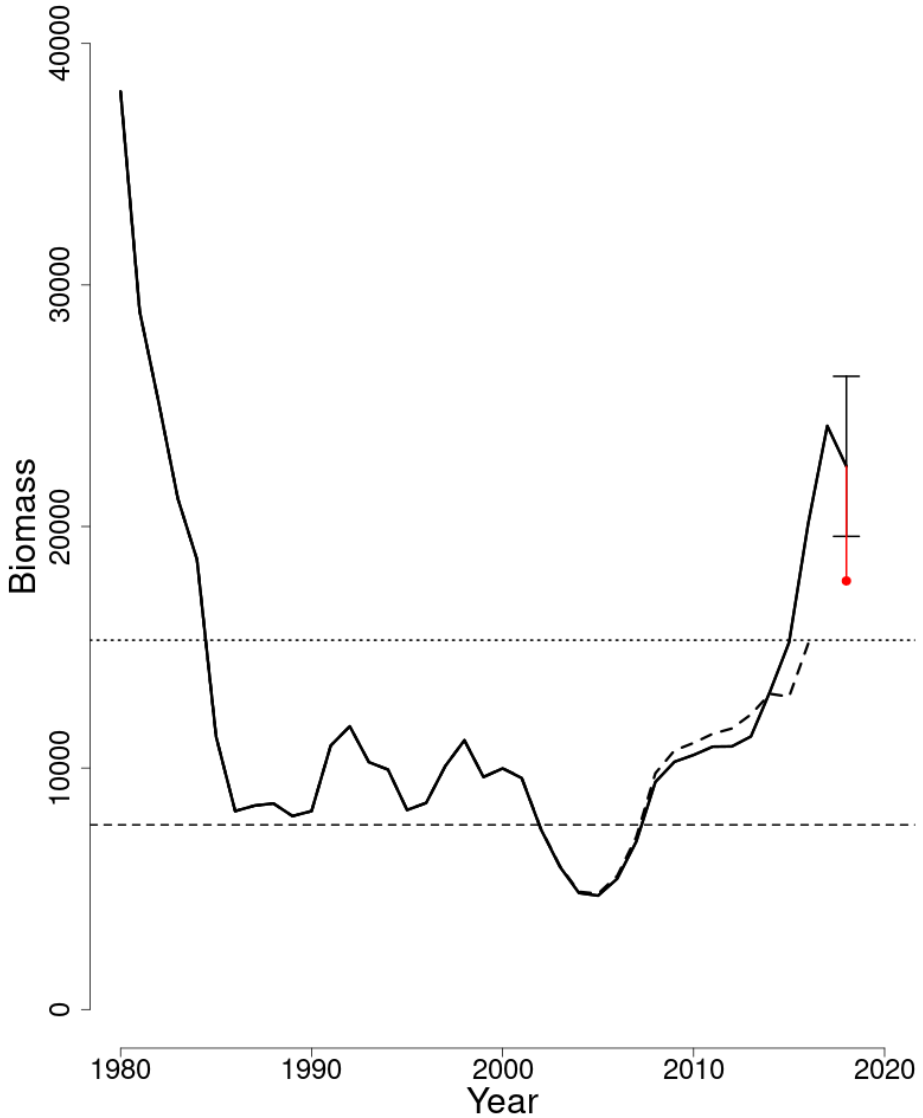

Figure 48: Trends in spawning stock biomass of Gulf of Maine-Georges Bank American plaice between 1980 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% normal confidence intervals are shown.

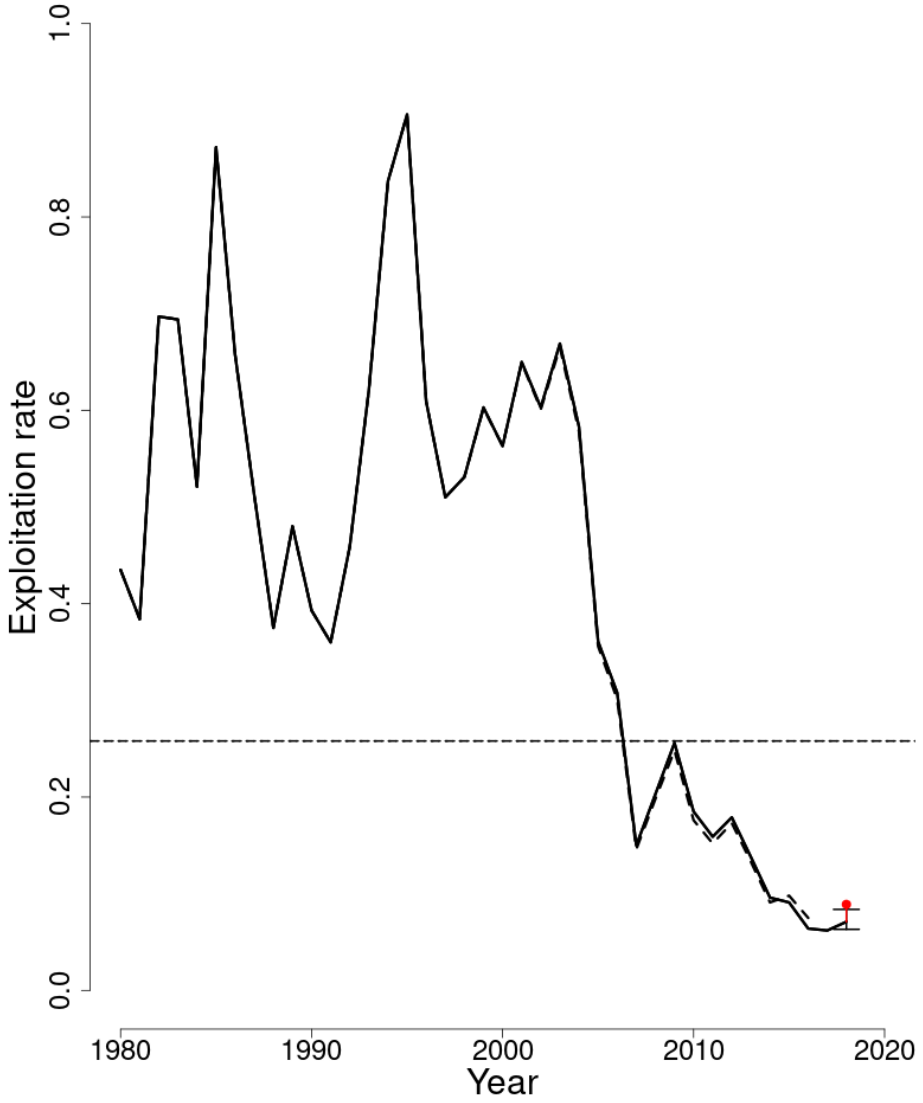

Figure 49: Trends in the fully selected fishing mortality (F_{Full}) of Gulf of Maine-Georges Bank American plaice between 1980 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.258; horizontal dashed line). F_{Full} was adjusted for a retrospective pattern and the adjustment is shown in red, based on the 2019 assessment. The approximate 90% normal confidence intervals are shown.

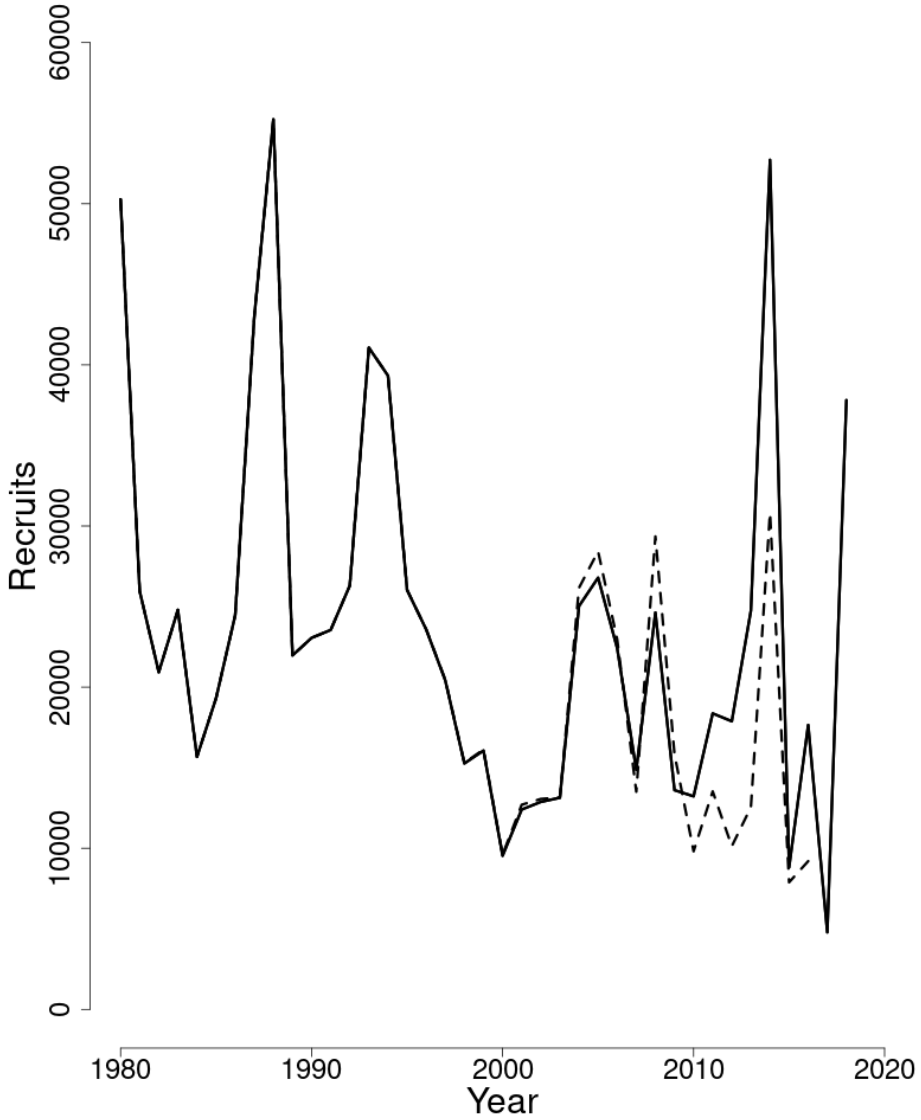

Figure 50: Trends in Recruits (age 1) (000s) of Gulf of Maine-Georges Bank American plaice between 1980 and 2018 from the current (solid line) and previous (dashed line) assessment.

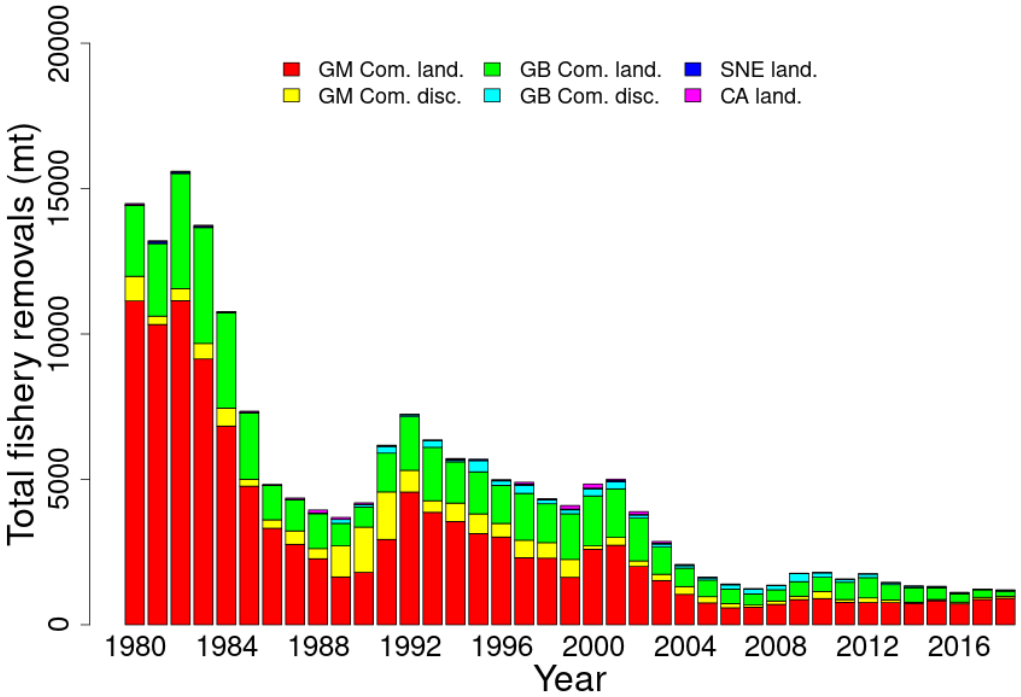

Figure 51: Total catch of Gulf of Maine-Georges Bank American plaice between 1980 and 2018 by fleet (Gulf of Maine, Georges Bank, Southern New England, and Canadian) and disposition (landings and discards).

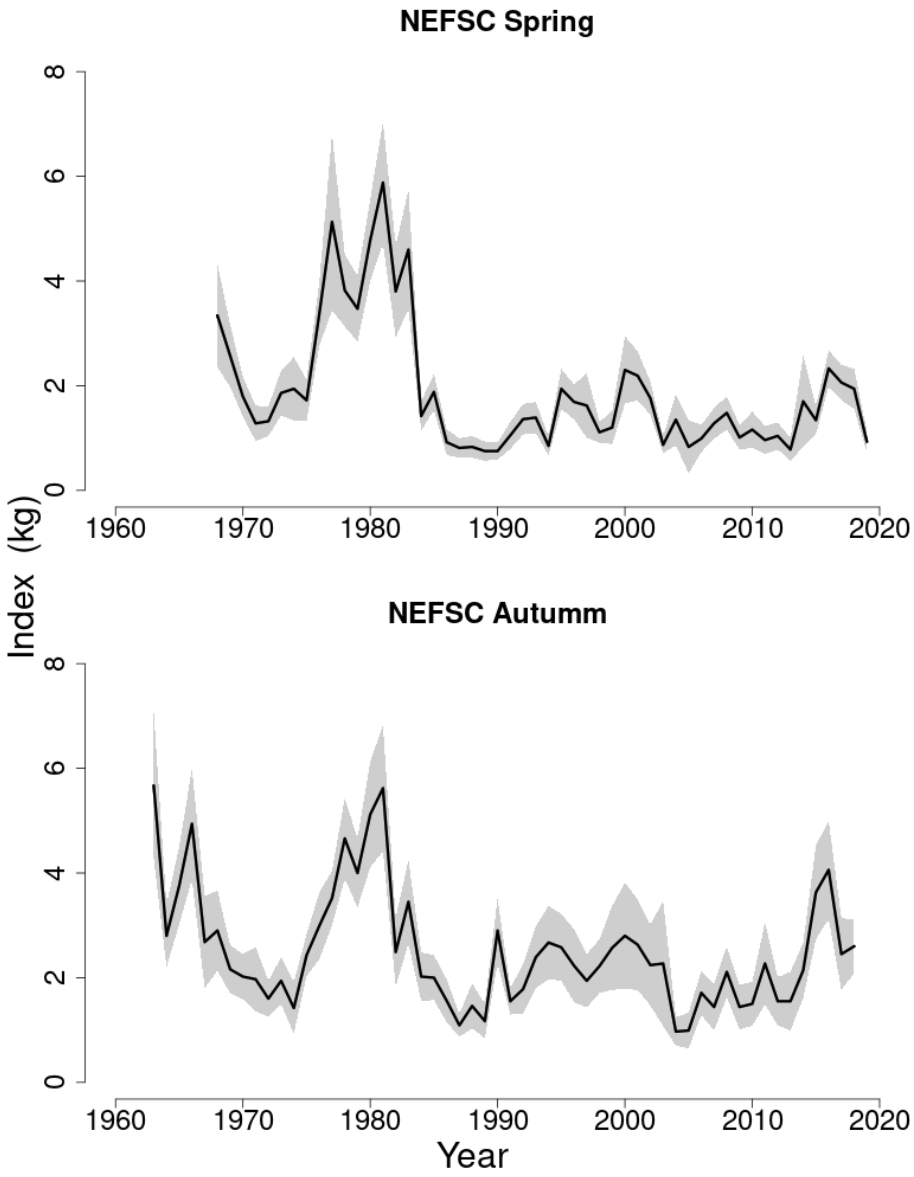

Figure 52: Indices of biomass for the Gulf of Maine-Georges Bank American plaice between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and autumn research bottom trawl surveys. The approximate 90% normal confidence intervals are shown.

10 Witch flounder

Susan Wigley

This assessment of the witch flounder (Glyptocephalus cynoglossus) stock is an operational assessment of the existing 2017 assessment (NEFSC 2017a). Based on the 2017 assessment the stock status was overfished and overfishing unknown, and stock condition was poor. This assessment updates commercial fishery catch data through 2018 (Table 34, Figure 55), and updates research survey biomass indices and the empirical approach assessment through 2018 (Figure 56). No stock projections can be computed using the empirical approach.

State of Stock: Based on this updated assessment, witch flounder (*Glyptocephalus cynoglossus*) recommended stock status cannot be determined analytically due to a lack of biological reference points associated with the empirical approach; stock condition remains poor. Retrospective adjustments were not made to the model results. The exploitable biomass in 2018 (defined as the arithmetic average of the 2018 NEFSC spring and 2017 NEFSC fall surveys population biomass estimates and converted to exploitable biomass using 0.9 based on examination of survey and fishery selectivity patterns) was estimated to be 35,585 (mt) (Figure 53). The 2018 exploitation rate (2018 catch divided by 2018 exploitable biomass) was estimated to be 0.02 (Figure 54).

Table 34: Catch and model results table for witch flounder. All weights are in (mt). The exploitable biomass in year y is the arithmetic average of the year y NEFSC spring and year y-1 NEFSC fall surveys then converted to exploitable biomass using 0.9. The exploitation rate is the year y catch divided by the year y exploitable biomass. Model results are from the current updated empirical approach assessment.

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
					Data						
Commercial Landings	1,009	954	759	870	1,038	686	570	492	397	446	606
Commercial Discards	127	204	153	201	232	124	106	94	115	106	115
Catch for Assessment	1,136	1,158	913	1,072	1,270	811	676	586	512	552	722
					Model Results						
Exploitable Biomass	39,131	22,689	18,403	17,986	20,390	13,634	16,690	19,670	18,331	24,820	35,585
Exploitation Rate	0.029	0.051	0.05	0.06	0.062	0.059	0.04	0.03	0.028	0.022	0.02

Table 35: Comparison of reference points estimated in an earlier assessment and from the current assessment update.

	2017	2019
F_{MSY} *proxy*	NA	NA
SSB_{MSY} (mt)	NA	NA
MSY (mt)	NA	NA
Overfishing	Unknown	Unknown
Overfished	Yes	Yes

Projections: Short term projections cannot be computed using the empirical approach. The estimated 2019 exploitable biomass is 30,371 mt. Using the January 2017 NEFMC PDT/SSC approach for catch advice, application of the mean exploitation rate of 4.9% (based on nine years, 2007-2015) to the 3 year (2017- 2019) moving average of exploitable biomass (30,259 mt) results in an estimated catch for 2020 of 1,483 mt.

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Uncertainty in the catch has increased due to recent criminal convictions in a case involving catch misreporting.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}; see Table 8).

  The model used to estimate status of this stock does not allow estimation of a retrospective pattern.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for witch flounder are not computed. Catch advice is derived from applying a mean exploitation rate of 0.049 (based on nine years, 2007-2015) to the 3 year average (2017-2019) of the exploitable biomass. The stock is in a revised rebuilding plan, rebuilding by 2043.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Recent landings and discards were updated and the time series of survey indices were updated; however, this has no impact on the stock status.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  In the 2019 assessment of witch flounder, the catch efficiency analyses were directly incorporated into the assessment model. Estimates of population biomass used revised catchability coefficients that varied by year; the revised catchability coefficients had a minor impact on catch advice in 2020. The 2018 NEFSC fall survey stratum 30 was not sampled; survey indices were not adjusted because this stratum represents less than 1% of total expanded catch weight and has negligible impact on survey indices and swept area biomass.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  No change in stock status has occurred for witch flounder since the previous assessment. Biological references points remain unknown.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The witch flounder stock condition remains poor. Fishery landings and survey catch by age indicate truncation of age structure and a reduction in the number of older fish in the population. NEFSC relative indices of abundance and biomass remain below their time series average.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The witch flounder assessment could be improved with accurate catch statistics; catch statistics have been undermined by misreporting, as partially documented in the recent criminal case. Additional research recommendations are given in NEFSC 2017b.

- Are there other important issues?

  The empirical approach does not incorporate age structure information. Consideration of incoming recruitment is critical for catch advice that supports stock rebuilding. This assessment uses revised catchability coefficients (q vary by year) in the estimates of population biomass. The 2016 and 2017 assessments applied a constant catchability coefficient (0.291). Minimum estimates of scientific research removals of witch flounder ranged between 0.1 and 15.9 mt, with an average of 1 mt between 1963 and 2018. The NEFSC bottom trawl surveys, Massachusetts Division of Marine Fisheries inshore surveys, Atlantic States Marine Fisheries Commission summer shrimp surveys, and various Cooperative Research surveys (e.g., such as Industry-based surveys for cod and for yellowtail flounder) and gear studies have contributed to scientific research removals. The August 2016 Gear Efficiency Study removed 14.0 mt of witch flounder.

10.1 Reviewer Comments: Witch flounder

The witch flounder stock assessment was not reviewed by the 2019 Review Panel because it was determined to be a level 1 assessment at the AOP meeting in June of 2019 (Appendix 16.2), according to the stock assessment process adopted for this and future management track assessments (Appendix 16.3).

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2017a. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. CRD17-17

Northeast Fisheries Science Center. 2017b. 62nd Northeast Regional Stock Assessment Workshop Assessment Report, Northeast Fisheries Science Center, Woods Hole, Massachusetts, January 2017. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 17-03; 822 p. CRD17-03

Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated through 2014. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 15-24; 251 p. CRD15-24

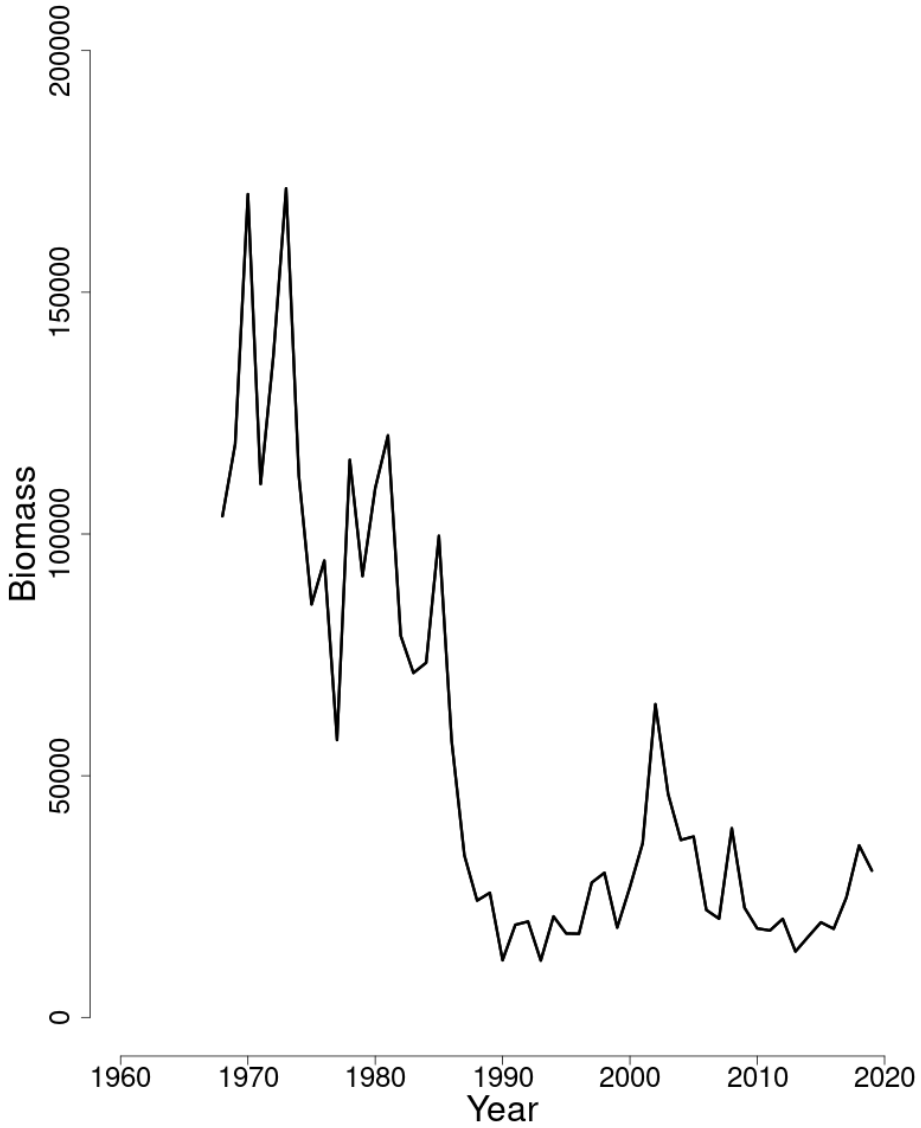

Figure 53: Trends in exploitable biomass (mt) of witch flounder between 1968 and 2019 from the current assessment.

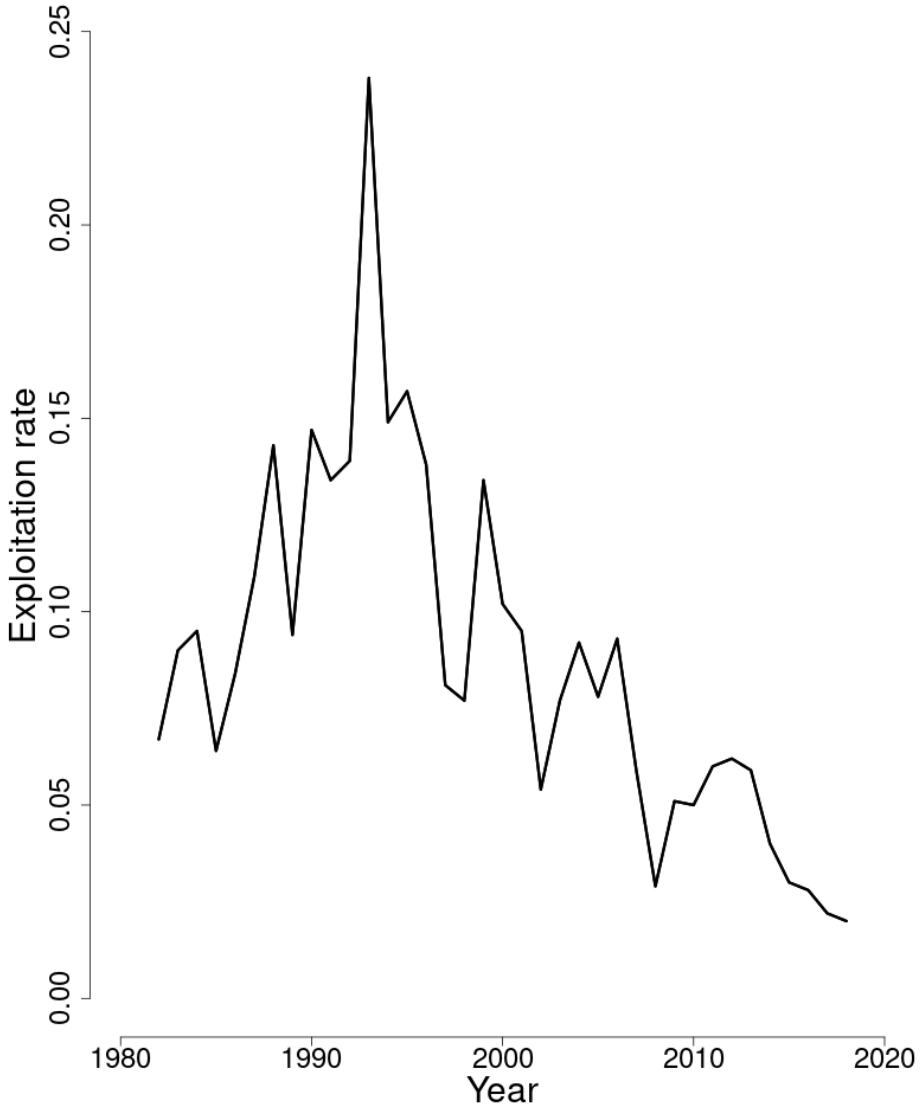

Figure 54: Trends in the exploitation rate (catch/ exploitable biomass) of witch flounder between 1982 and 2018 from the current assessment.

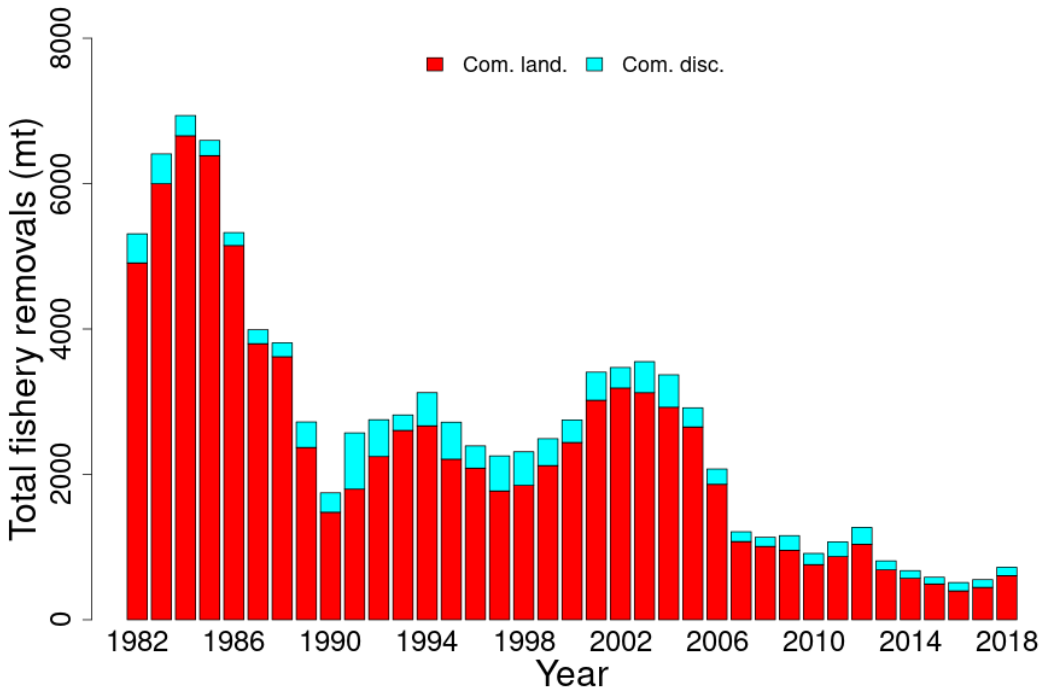

Figure 55: Total catch of witch flounder between 1982 and 2018 by fleet (commercial) and disposition (landings or discards).

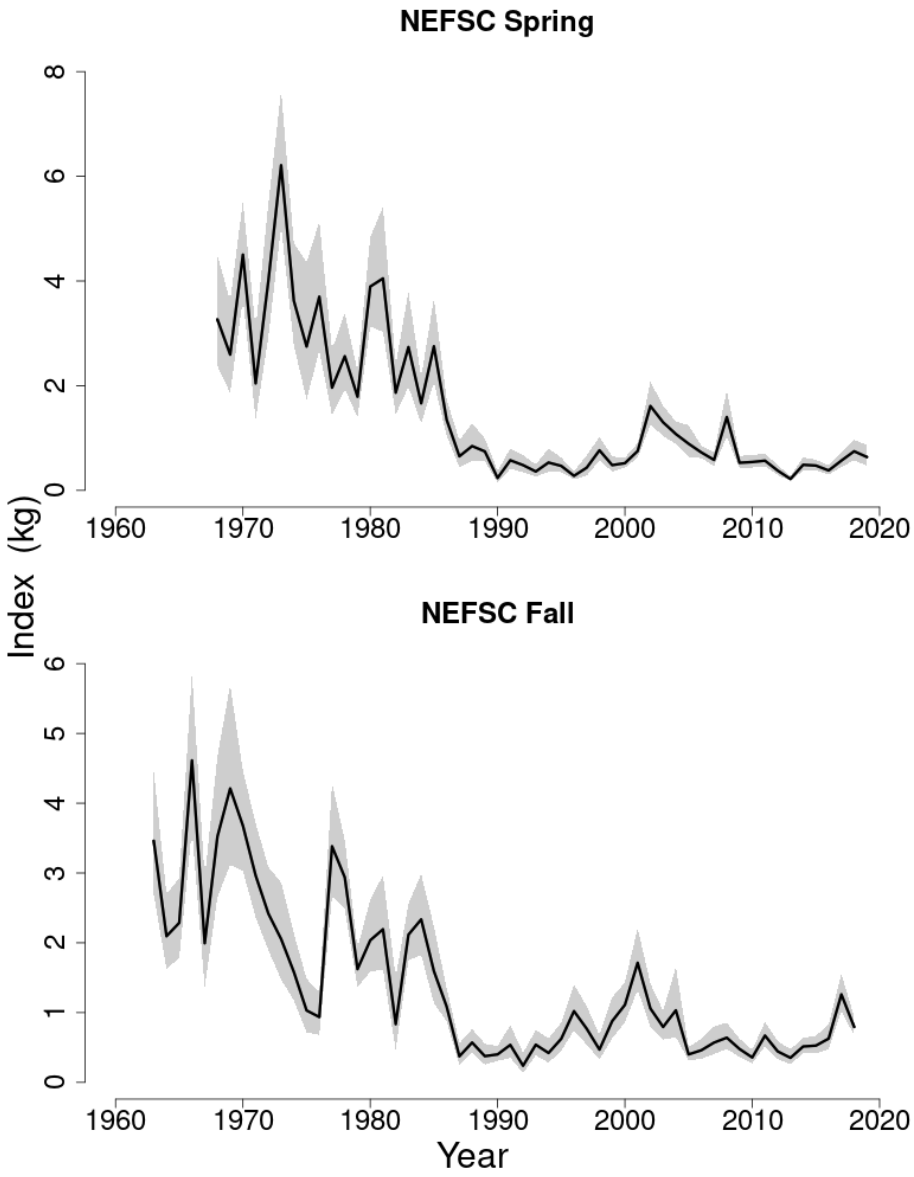

Figure 56: Indices of biomass for the witch flounder between 1963 (Fall) and 2019 (Spring) for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys. The approximate 90% lognormal confidence intervals are shown.

11 White hake

Katherine Sosebee

This assessment of the white hake (Urophycis tenuis) stock is an operational update of the 2017 operational assessment (NEFSC 2017) and the last benchmark assessment (NEFSC 2013). Based on the previous assessment the stock was not overfished and overfishing was not ocurring. This assessment updates commercial fishery catch data, research survey indices of biomass, and the ASAP assessment model and reference points through 2018. Stock projections have been updated through 2022.

State of Stock: Based on this updated assessment, the white hake (*Urophycis tenuis*) stock is overfished and overfishing is not occurring (Figures 57-58). Retrospective adjustments were made to the model results. Spawning stock biomass (SSB) in 2018 was estimated to be 15,891 (mt) which is 50% of the biomass target (SSB_{MSY} *proxy* = 31,828; Figure 57). The 2018 fully selected fishing mortality was estimated to be 0.129 which is 77% of the overfishing threshold proxy (F_{MSY} *proxy* = 0.1677; Figure 58).

Table 36: Catch and ASAP results table for white hake. All weights are in (mt) recruitment is in (000s) and F_{Full} is the fishing mortality on fully selected ages (ages 6 - 9+). Model results are from the current ASAP assessment.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
					Data					
Commercial discards	82	89	49	50	38	33	24	33	36	29
Commercial landings	1,712	1,820	2,899	2,771	2,235	1,887	1,632	1,325	1,976	1,969
Canadian landings	79	104	86	83	43	35	25	39	32	45
Other landings	0	0	0	0	0	0	0	0	0	0
Catch for Assessment	1,873	2,012	3,034	2,903	2,316	1,955	1,680	1,396	2,043	2,044
					Model Results					
Spawning Stock Biomass	12,471	15,998	19,158	19,778	20,222	19,739	18,986	22,494	24,386	20,757
F_{Full}	0.174	0.139	0.181	0.169	0.129	0.114	0.097	0.066	0.091	0.107
Recruits (age 1)	3,483	3,034	2,867	2,820	3,238	2,960	3,334	2,080	3,168	4,038

Table 37: Comparison of reference points estimated in the 2017 assessment and from the current assessment update. An $F_{40\%}$ proxy was used for the overfishing threshold and SSB_{MSY} was based on long-term stochastic projections which sampled from a cumulative distribution function of recruitment estimates from ASAP from 1963-2016. The annual fishery selectivity, maturity ogive, and mean weights at age used in the projection are the most recent 5 year averages.

	2017	2019
F_{MSY} *proxy*	0.1839	0.1677
SSB_{MSY} (mt)	30,948	31,828 (25,398 - 40,317)
MSY (mt)	4,867	4,601 (3,665 - 5,828)
Median recruits (age 1) (000s)	4,616	4,471
Overfishing	No	No
Overfished	No	Yes

Projections: Short term projections of catch and SSB were derived by sampling from a cumulative distribution function of recruitment estimates from ASAP from 1995-2016. The annual fishery selectivity, maturity ogive, and mean weights at age used in the projection are the most recent 5 year averages. The numbers-at-age used to start the projections were adjusted for retrospective bias using age-specific rho estimates.

Table 38: Short term projections of total fishery catch and spawning stock biomass for white hake based on a harvest scenario of fishing at F_{MSY} *proxy* between 2020 and 2022. Catch in 2019 was assumed to be 2,140 (mt) which is 55% of the 2019 OFL.

Year	Catch (mt)	SSB (mt)	F_{Full}
2019	2,140	19,412 (16,665 - 22,697)	0.106

Year	Catch (mt)	SSB (mt)	F_{Full}
2020	2,857	19,580 (16,730 - 22,856)	0.1677
2021	2,809	19,474 (16,764 - 22,315)	0.1677
2022	2,791	19,343 (16,885 - 21,914)	0.1677

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  1. Catch at age information is not well characterized due to possible mis-identification of species in the commercial and observer data, particularly in early years, low sampling of commercial landings in some years, and sparse discard length data.

  2. Since the commercial catch is aged primarily with survey age/length keys, there is considerable augmentation required, mainly for ages 5 and older. The numbers at age and mean weights at age in the catch for these ages may therefore not be well specified.

3. White hake may move seasonally into and out of the defined stock area.

4. There are no commercial catch at age data prior to 1989 and the catchability of older ages in the surveys is very low. This results in a large uncertainty in starting numbers at age.

5. Since 2003, dealers have apparently been culling extra-large fish out of the large category. However, there was no market category for landings until June 2014. The length compositions are distinct from fish characterized as large and have been identified since 2011. This may bias the age composition of the landings, particularly in 2014 when 2000 of the 5000 large samples were these extra-large fish.

6. A pooled age/length key is used for 1963-1981 and fall 2003 (second half of commercial key).

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}

  The 7-year Mohn's ρ, relative to SSB, was 0.22 in the 2017 assessment and was 0.31 in 2018. The 7-year Mohn's ρ, relative to F, was -0.15 in the 2017 assessment and was -0.22 in 2018. There was a major retrospective pattern for this assessment because the ρ adjusted estimate of 2018 SSB (SSB_{ρ}=15891) was outside the approximate 90% confidence regions around SSB (17,792 - 24,216). A retrospective adjustment was made for both the determination of stock status and for projections of catch in 2020. The retrospective adjustment changed the 2018 SSB from 20,757 to 15,891 and the 2018 F_{Full} from 0.107 to 0.129.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for white hake are not well determined and projected biomass from the last assessment was near the edge the confidence bounds of the biomass estimated in the current assessment. The rebuilding deadline for this stock was 2014 and the stock is not yet rebuilt and is now likely overfished.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  The experimental catchability data were not applicable to the white hake stock assessment because the catchability experiments were focused on flatfish species.

  The discard time series was re-estimated to incorporate changes made to the underlying data. This had almost no impact on the assessment.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  Stock status of white hake has changed from not overfished to overfished and the stock has not rebuilt even with a very low fishing mortality. The numbers for the 2008-2010 year classes, which were included in the age 5-7 starting numbers in the projections, were over-estimated which led to over-estimating of the 2016 SSB. In addition, the 2014 year class(age 1 in 2015) was over-estimated by 167% in the 2017 assessment which contributed to the optimistic projections.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  The white hake stock shows no truncation of age structure. Estimates of commercial landings and discards have decreased over time.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  Age structures collected by the observer program are available and should be aged to augment the survey keys. They are also available from the ASMFC shrimp survey and would allow another survey to be added to the model. Otoliths are currently being collected from the market category for heads and these should also be aged.

- Are there other important issues?

  None.

11.1 Reviewer Comments: White hake

11.1.1 Review Panel Summary

The completed Plan A operational assessment is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

The White Hake (HKW) stock was assessed using the ASAP model. The retrospective pattern appears to be worsening, and retrospective adjustment was deemed necessary for this years assessment, producing a change in stock status to overfished.

11.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. There is some concern about being able to separate out red hake from white hake at young ages and this may lead to variability in the estimates of recruitment over time. There is also a concern for the largest fish being sampled not in proportion to the landings due to the combining of large and extra-large market categories. The lack of cohort signals may be further evidence of both of these issues.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is now overfished, but overfishing is not occurring.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

Surveys have indicated increasing biomass recently, but they remain below the time series means. There are no signals of large incoming recruitment. Age and size structure in the surveys are not informative due to strong doming in selectivity.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were made using ASAP using Rho-adjusted numbers-at-age with age-specific rho values.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

Consider bringing in the Gulf of Maine longline survey information into the assessment.

Consider exploring changing the index cv as some are falling outside the confidence intervals of the observations. Consider adding another selectivity block to the assessment (e.g., 2010 management changes may justify an assumed change in selectivity). Complete ageing of the survey (fall 2003). Complete ageing of the observer samples from 2001 on. Complete aging of the other surveys (shrimp, MENH) and attempt to use as recruitment indices.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

NEFSC. 2013. 56th Northeast Regional Stock Assessment Workshop (56th SAW) Assessment Report.US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 13-10; 868 p. http://www.nefsc.noaa.gov/publications/crd/crd1310/

NEFSC. 2017. Operational Assessment of 19 Northeast Groundfish Stocks Updated Through 2016.US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. http://www.nefsc.noaa.gov/publications/crd/crd1717/

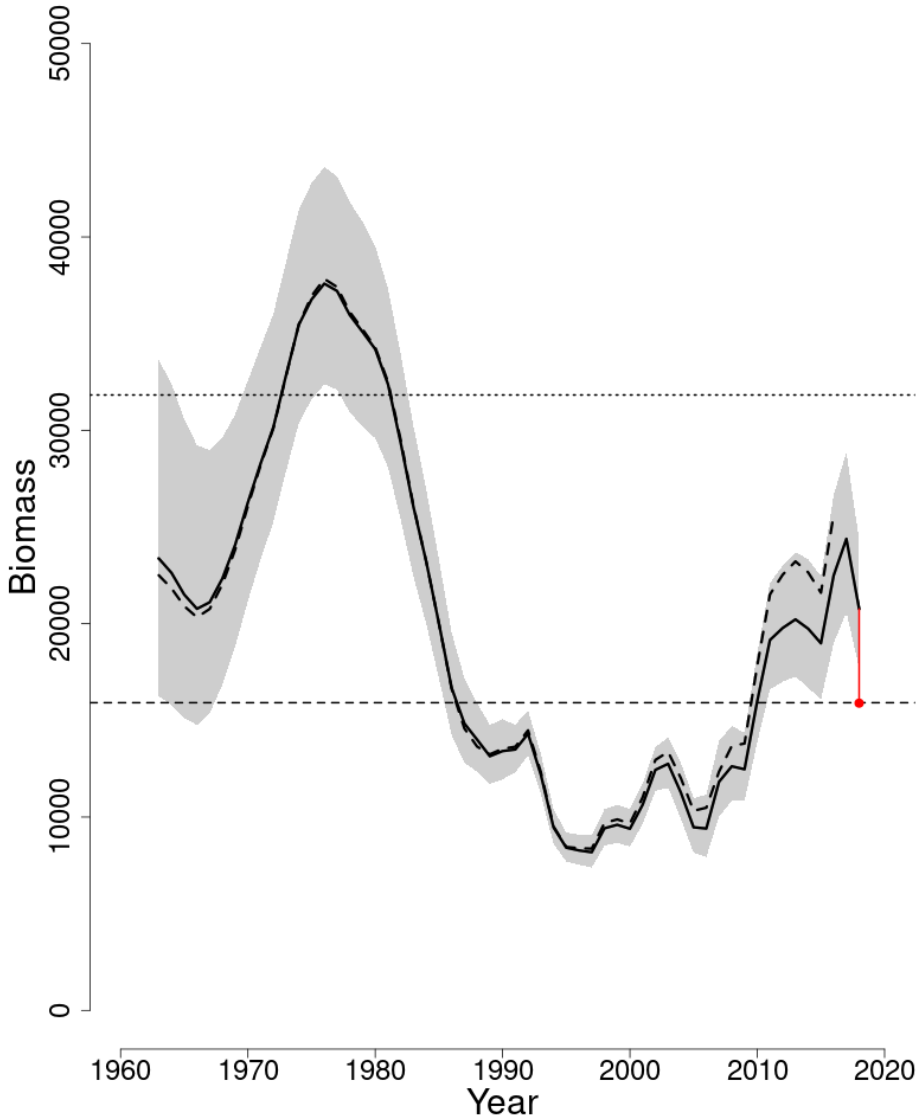

Figure 57: Trends in spawning stock biomass of white hake between 1963 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ ($\frac{1}{2}$ SSB_{MSY} *proxy*; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} *proxy*; horizontal dotted line) based on the 2019 assessment. Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% lognormal confidence intervals are shown.

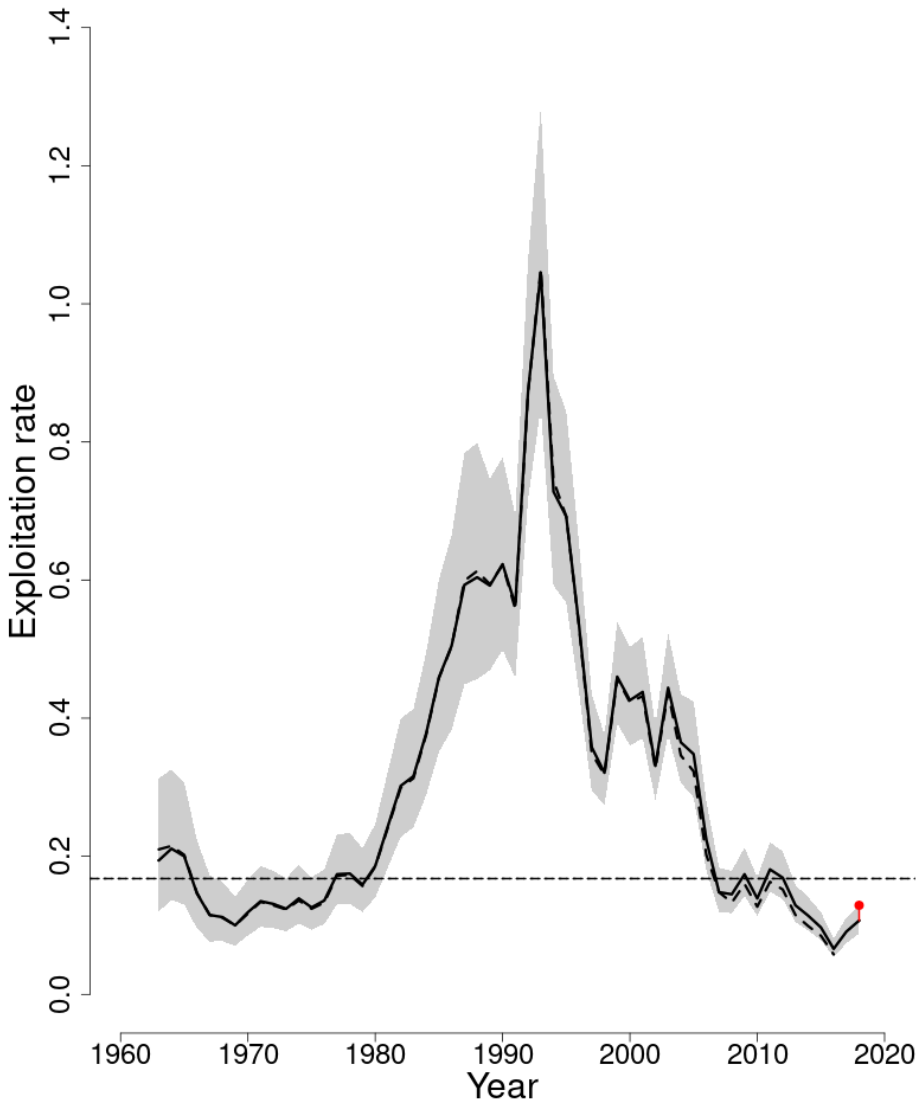

Figure 58: Trends in the fully selected fishing mortality (F_{Full}) of white hake between 1963 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} *proxy*=0.1677; horizontal dashed line). based on the 2019 assessment.The F_{Full} was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% lognormal confidence intervals are shown.

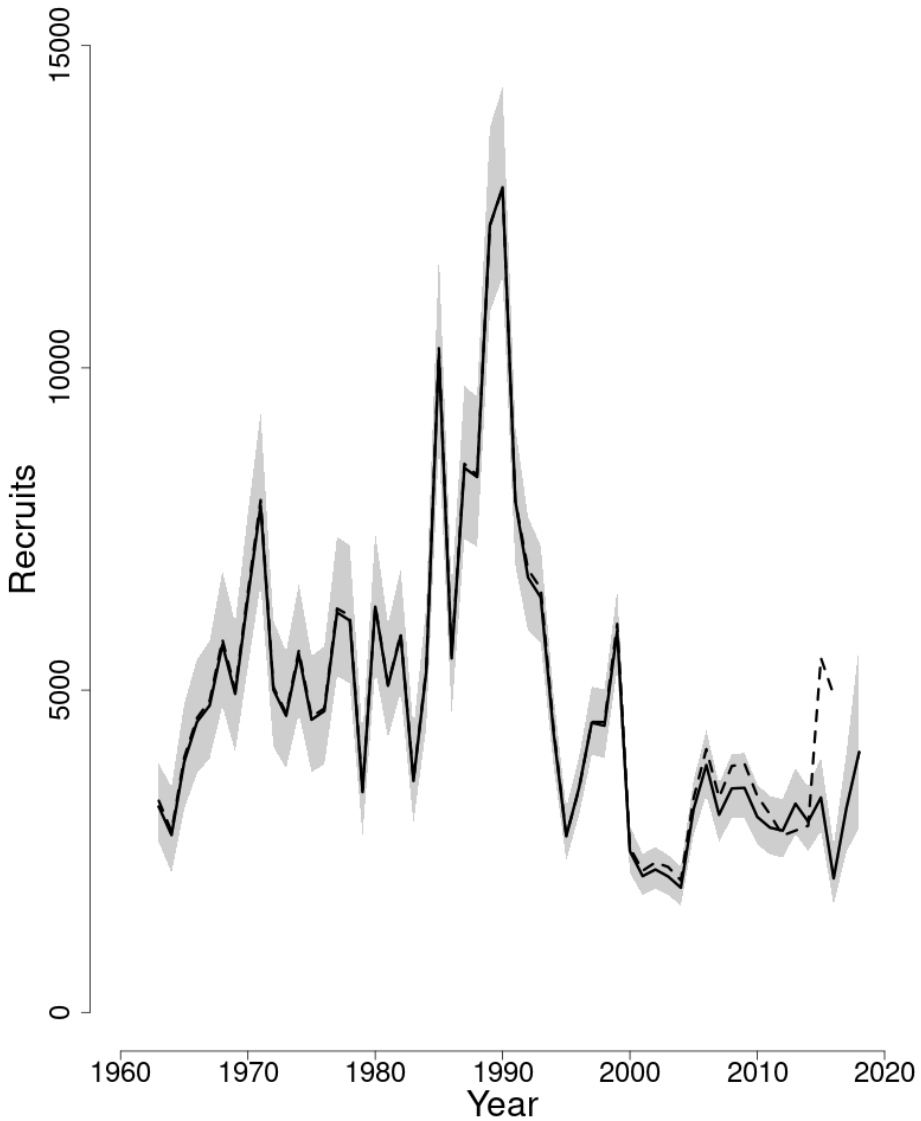

Figure 59: Trends in Recruits (age 1) (000s) of white hake between 1963 and 2018 from the current (solid line) and previous (dashed line) assessment. The approximate 90% lognormal confidence intervals are shown.

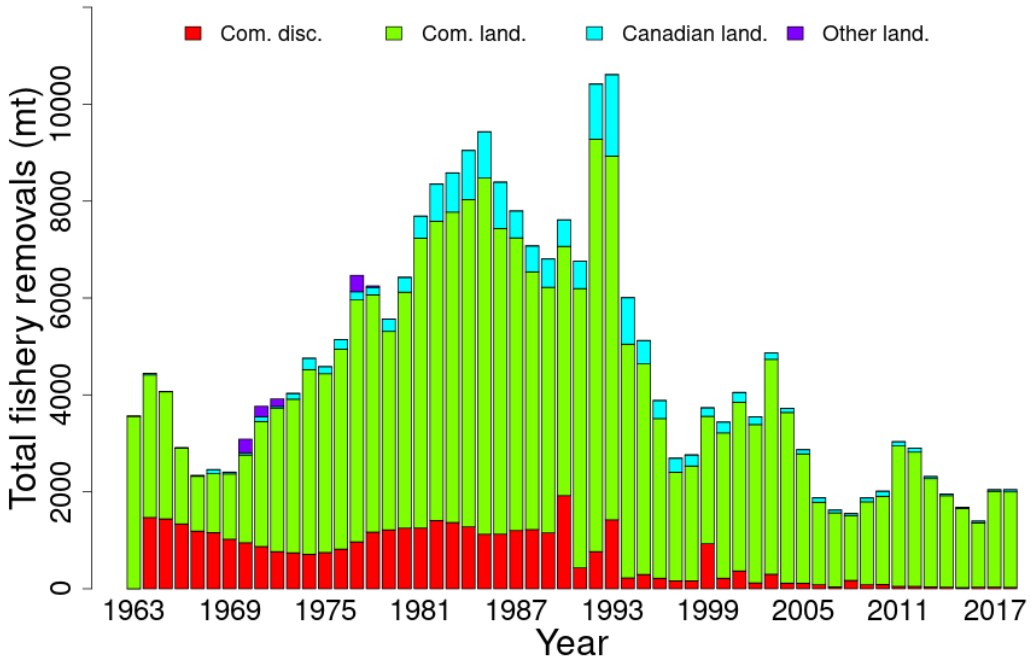

Figure 60: Total catch of white hake between 1963 and 2018 by fleet (commercial, recreational, or Canadian) and disposition (landings and discards).

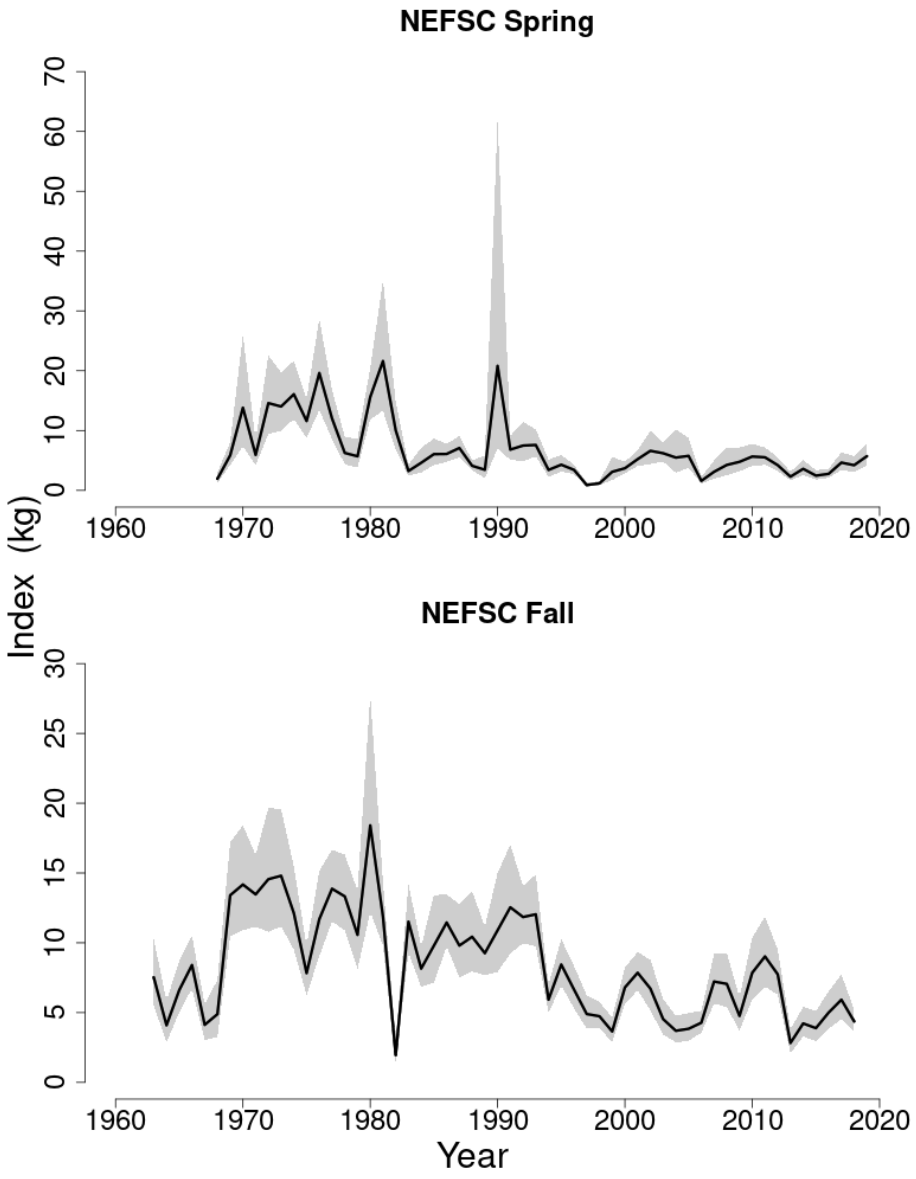

Figure 61: Indices of biomass for the white hake between 1963 and 2019 for the Northeast Fisheries Science Center (NEFSC) spring and fall bottom trawl surveys. The approximate 90% lognormal confidence intervals are shown.

12 Pollock

Brian Linton

This assessment of the pollock (Pollachius virens) stock is an update of the existing 2017 operational assessment (NEFSC 2017). This assessment updates commercial and recreational fishery catch data, research survey indices of abundance, the ASAP analytical models, and biological reference points through 2018. Additionally, stock projections have been updated through 2022. In what follows, there are two population assessment models brought forward from the 2017 operational assessment: the base model (dome-shaped survey selectivity), which is used to provide management advice; and the flat sel sensitivity model (flat-topped survey selectivity), which is included for the sole purpose of demonstrating the sensitivity of assessment results to survey selectivity assumptions. The most recent benchmark assessment of the pollock stock was in 2010 as part of the 50th Stock Assessment Review Committee (SARC 50; NEFSC 2010), which includes a full description of the model formulations.

State of Stock: The pollock (*Pollachius virens*) stock is not overfished and overfishing is not occurring (Figures 62-63). Retrospective adjustments were made to the model results. Retrospective adjusted spawning stock biomass (SSB) in 2018 was estimated to be 212,416 (mt) under the base model and 71,322 (mt) under the flat sel sensitivity model which is 170 and 101% (respectively) of the biomass target, an SSB_{MSY} proxy of SSB at $F_{40\%}$ (124,639 and 70,721 (mt); Figure 62). Retrospective adjusted 2018 age 5 to 7 average fishing mortality (F) was estimated to be 0.038 under the base model and 0.094 under the flat sel sensitivity model, which is 14 and 36% (respectively) of the overfishing threshold, an F_{MSY} proxy of $F_{40\%}$ (0.272 and 0.26; Figure 63).

Table 39: Catch and status table for pollock. All weights are in (mt), recruitment is in (000s), and F_{AVG} is the age 5 to 7 average F. Unadjusted SSB and F estimates are reported. Model results are from the current base model and flat sel sensitivity model.

	2011	2012	2013	2014	2015	2016	2017	2018
			Data					
Commercial landings	7,211	6,742	5,058	4,545	3,043	2,582	3,249	3,078
Commercial discards	176	121	169	135	155	97	49	70
Recreational landings	3,447	1,355	4,078	1,511	752	1,030	1,239	687
Recreational discards	2,958	2,151	4,123	2,441	2,190	1,522	2,059	944
Catch for Assessment	13,792	10,370	13,428	8,632	6,139	5,231	6,597	4,779
			Model Results (base)					
Spawning Stock Biomass	234383	208817	196520	184110	208798	221237	250282	276305
F_{AVG}	0.136	0.108	0.157	0.108	0.068	0.048	0.044	0.027
Recruits $age1$	29695	51121	50567	75056	49903	36034	32358	24169
			Model Results (flat sel sensitivity)					
Spawning Stock Biomass	88172	76164	70252	62825	73521	84802	100368	112633
F_{AVG}	0.279	0.231	0.366	0.261	0.163	0.11	0.098	0.058
Recruits $age1$	16057	27367	27264	40406	27095	19710	17940	13950

Table 40: Comparison of biological reference points for pollock estimated in the 2017 assessment and from the current base model and flat sel sensitivity model. An F_{MSY} proxy of $F_{40\%}$ was used for the overfishing threshold, and was based on yield per recruit analysis. F_{MSY} is reported as the age 5 to 7 average F. Recruits represent the median of the predicted recruits. Intervals shown are 5th and 95th percentiles.

	2017 base	2017 flat sel sensitivity	base	flat sel sensitivity
F_{MSY}	0.260	0.249	0.272	0.260
SSB_{MSY} (mt)	105,510	60,738	124,639 (98,701 - 158,416)	70,721 (55,964 - 89,609)
MSY (mt)	19,427	11,692	19,856 (14,471 - 27,709)	12,007 (8,876 - 16,407)
Median recruits (age 1) (000s)	22,183	13,067	25,312	14,503
Overfishing	No	No	No	No
Overfished	No	No	No	No

Projections: Short term projections of median total fishery yield and spawning stock biomass for pollock were conducted based on a harvest scenario of fishing at an F_{MSY} proxy of $F_{40\%}$ between 2020 and 2022. Catch in 2019 has been estimated at 5,140 (mt). Recruitments were sampled from a cumulative distribution function derived from ASAP estimated age 1 recruitment between 1970 and 2016. Recruitments in 2017 and 2018 were not included due to uncertainty in those estimates. The annual fishery selectivity, natural mortality, maturity ogive, and mean weights used in projections are the most recent 5 year averages. Retrospective adjusted age 5 to 7 average F in 2018 fell outside the 90% confidence intervals of the unadjusted 2018 value under the base model (Figure 63). Retrospective adjusted SSB and age 5 to 7 average F in 2018 fell outside the 90% confidence intervals of the unadjusted 2018 values under the flat sel sensitivity model (Figures 62-63). Therefore, age-specific abundance rho values were applied to the initial numbers at age in the projections for the base model and the flat sel sensitivity model.

Table 41: Retrospective adjusted short term projections of median total fishery yield and spawning stock biomass for pollock from the current base model and flat sel sensitivity model based on a harvest scenario of fishing at an F_{MSY} proxy of $F_{40\%}$ between 2020 and 2022. Catch in 2019 has been estimated at 5,140 (mt). F_{AVG} is the age 5 to 7 average F.

Year	Catch (mt)	SSB (mt)	F_{AVG}	Catch (mt)	SSB (mt)	F_{AVG}
		base			*flat sel sensitivity*	
2019	5,140	190,927	0.036	5,140	65,237	0.092

Year	Catch (mt)	SSB (mt)	F_{AVG}	Catch (mt)	SSB (mt)	F_{AVG}
		base			*flat sel sensitivity*	
2020	35,358	200,992	0.272	14,522	69,808	0.260
2021	26,765	176,117	0.272	11,924	63,273	0.260
2022	19,889	160,156	0.272	9,388	59,921	0.260

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  One of the greatest sources of uncertainty in the pollock assessment is selectivity, as the base model with dome-shaped survey and fishery selectivities implies the existence of a large cryptic biomass that neither current surveys nor the fishery can confirm. Assuming that survey selectivity is flat-topped leads to lower estimates of SSB and higher estimates of F (Figures 62-63). Stock status is insensitive to the shape of the survey selectivity patterns at older ages. Another source of uncertainty is the major retrospective pattern (see Question 2). In addition, the strength of the 2013 year class is a source of uncertainty in short term stock projections. For both models, the 2013 year class is estimated to be smaller in size than in the previous assessment, but it is still estimated to be the largest year class in the assessment time series, 1970-2018. The 2013 year class has begun to enter the commercial fishery, and uncertainty in the year class' strength should decrease as it moves through the fishery in subsequent years.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{AVG} lies outside of the approximate joint confidence region for SSB and F_{AVG}; see Table 8).

  The 7-year Mohn's ρ, relative to SSB, was 0.231 under the base model and 0.407 under the flat sel sensitivity model in the 2017 assessment and was 0.301 and 0.579, respectively, in 2018. The 7-year Mohn's ρ, relative to F, was -0.278 under the base model and -0.35 under the flat sel sensitivity model in the 2017 assessment and was -0.282 and -0.389, respectively, in 2018. There was a major retrospective pattern for the base model because the ρ adjusted estimate of 2018 F (F_ρ=0.038) was outside the approximate 90% confidence region around F (0.019-0.035). There was a major retrospective pattern for the flat sel sensitivity model because the ρ adjusted estimates of 2018 SSB (SSB_ρ=71,322 (mt)) and 2018 F (F_ρ=0.094) were outside the approximate 90% confidence region around SSB (83,067-142,199 (mt)) and F (0.042-0.073). A retrospective adjustment was made for both the determination of stock status and for projections of catch in 2020. The base model retrospective adjustment changed the 2018 SSB from 276,305 (mt) to 212,416 (mt) and the 2018 F_{AVG} from 0.027 to 0.038. The flat sel sensitivity model retrospective adjustment changed the 2018 SSB from 112,633 (mt) to 71,322 (mt) and the 2018 F_{AVG} from 0.058 to 0.094.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  Population projections for pollock appear to be reasonably well determined for both the base model and the flat sel sensitivity model. The stock is not in a rebuilding plan.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  Two changes were made to the pollock assessment as part of this update. First, the new calibrated recreational catch estimates were used in the assessment. The new recreational catch estimates are greater than the old estimates, particularly at the beginning and end of

the time series, 1981-2018. In both models, the new recreational catch estimates may contribute to the increased scaling of SSB compared to SSB estimates from the previous assessment, which used the old recreational catch estimates. Second, evaluation of the commercial age composition residuals led to the inclusion of a new commercial selectivity time block, beginning in 2010. In both models, the new time block improved the residual patterns, and led to an increased scaling of SSB compared to runs without the new time block. This rescaling of SSB likely is due to the difficulty that both models have in scaling the stock size (see Question 8). In addition to these two changes, the impact of survey stratum 1300 not being sampled in the 2018 fall bottom trawl survey was explored. No adjustments were made to the 2018 fall survey index value, because stratum 1300 makes up an average of only 1% of the expanded survey catch in numbers over the entire time series, 1970-2018, and only 3% of the expanded survey catch in numbers in recent years, 2009-2018.

Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

The experimental catchability data were not applicable to the pollock stock assessment, because the catchability experiments were focused on flatfish species.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  Stock status based on the base and flat sel sensitivity models has not changed since the previous assessment.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Total removals of pollock have declined since 2013. The spring survey index increased from 2013 to 2018, before decreasing in 2019. The fall survey index has decreased since 2014. Fishery and survey data suggest the exisitence of a relatively strong 2013 year class, which has just begun to enter the commercial fishery. Survey data suggests that older fish have begun to reappear in the stock since the 1990s.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The pollock assessment could be improved with additional studies on gear selectivity. These studies could cover topics such as physical selectivity (e.g., multi-mesh gillnet), behavior (e.g., swimming endurance, escape behavior), geographic and vertical distribution by size and age, tag-recovery at size and age, and evaluating information on length-specific selectivity at older ages.

- Are there other important issues?

  As in the previous assessment, both of the pollock assessment models had difficulty converging on a solution in some of the retrospective peels and jitter analysis runs. One

possible explanation for this issue is that the models may be overparameterized, with the base and flat sel sensitivity models estimating 223 and 221 parameters, respectively. The high number of parameters is due to the fact that the commercial and recreational fisheries are modeled as separate fleets. The effects of combining the two fleets into a single fleet should be explored during the next benchmark assessment. In addition, both of the models have a tendency to rescale the population size when years of data are dropped or added to the assessment, while the relative trends in stock size over time remain the same. This difficulty in scaling the stock may be tied to the convergence issue.

12.1 Reviewer Comments: Pollock

12.1.1 Review Panel Summary

The completed Plan A operational assessment of pollock is technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

The assessment was conducted in ASAP for years 1970-2018 and ages: 1-9+. Commercial and recreational fisheries were modeled as separate fleets. The base model assumes dome-shaped selectivity for both the fishery and survey. An application of ASAP with dome-shaped fishery selectivity and flat-topped survey selectivity was explored as a sensitivity analysis to the dome shaped selectivity assumption for the surveys, as recommended in the benchmark assessment.

There are several diagnostic problems with this stock assessment. Domed selectivity across both the survey indices and the fleets in the model results in cryptic biomass (biomass that is assumed to be in the system, but does not appear in either the fishery or the survey). The base case model with domed selectivity fits the data better. The sensitivity run with flat-topped selectivity shows what the estimates would be with the "cryptic" portion excluded. Both models have some difficulty with parameter estimation. The new assessment format should be helpful in allowing shorter term exploration of alternative data and modeling assumptions since changes can be made in a stepwise fashion without the need of a full benchmark, and there is more room to explore different options.

12.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The Plan A assessment is recommended for use with retrospective adjustments. A new selectivity period was assumed for 2010-2018, consistent with the transition to the Annual Catch Limit and sector management system. The Panel accepts that the new selectivity period was an appropriate model revision.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The biological reference points (BRPs) for this stock were updated and used.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

The stock is not overfished and overfishing is not occurring.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

The two survey indices show conflicting trends in recent years, with the spring index increasing and the fall index decreasing. The spring 2019 index value is not included in the ASAP model, but does show a decrease, which might reduce conflicts in the next assessment.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

Projections were provided based on the ASAP model with retrospective adjustments.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

A major source of uncertainty is the assumed shape of survey selectivity curve. The base model predicts a large cryptic biomass that cannot be confirmed by fisheries or surveys. However, stock status is insensitive to shape of selectivity curve at older ages. Resolving the question of how to best characterize survey selectivity may be evaluated in the next research track assessment, but additional information may be needed to find a reasonable solution (e.g. where are the older fishing going if they exist).

Another source of uncertainty is a strong retrospective pattern. The retrospective inconsistency has increased slightly from the 2017 assessment for all estimates and for all models except for recruitment in the flat-top survey selectivity model, which showed a slight decrease in retrospective pattern.

In addition to model scaling problems associated with the survey selectivity assumptions, model convergence was an issue suggesting that the model may be over parameterized or contains parameters that are highly colinear. Model over parameterization and colinearity should be explored in the next assessment as a means to avoid convergence issues.

The benchmark assessment begins with a starting point in 1970, but the series of recreational catches begins in 1981. The implicit assumption of no recreational fishery for 1970-1980 was justified in the benchmark assessment, because recreational catch was usually less than 10% of the total catch. However, revised MRIP estimates are greater. The Panel suggests either changing the starting year of the model or hindcasting recreational catches so that estimates of recreational catch are not assumed to be zero in the starting year.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-17; 844 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD10-17

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. CRD17-17

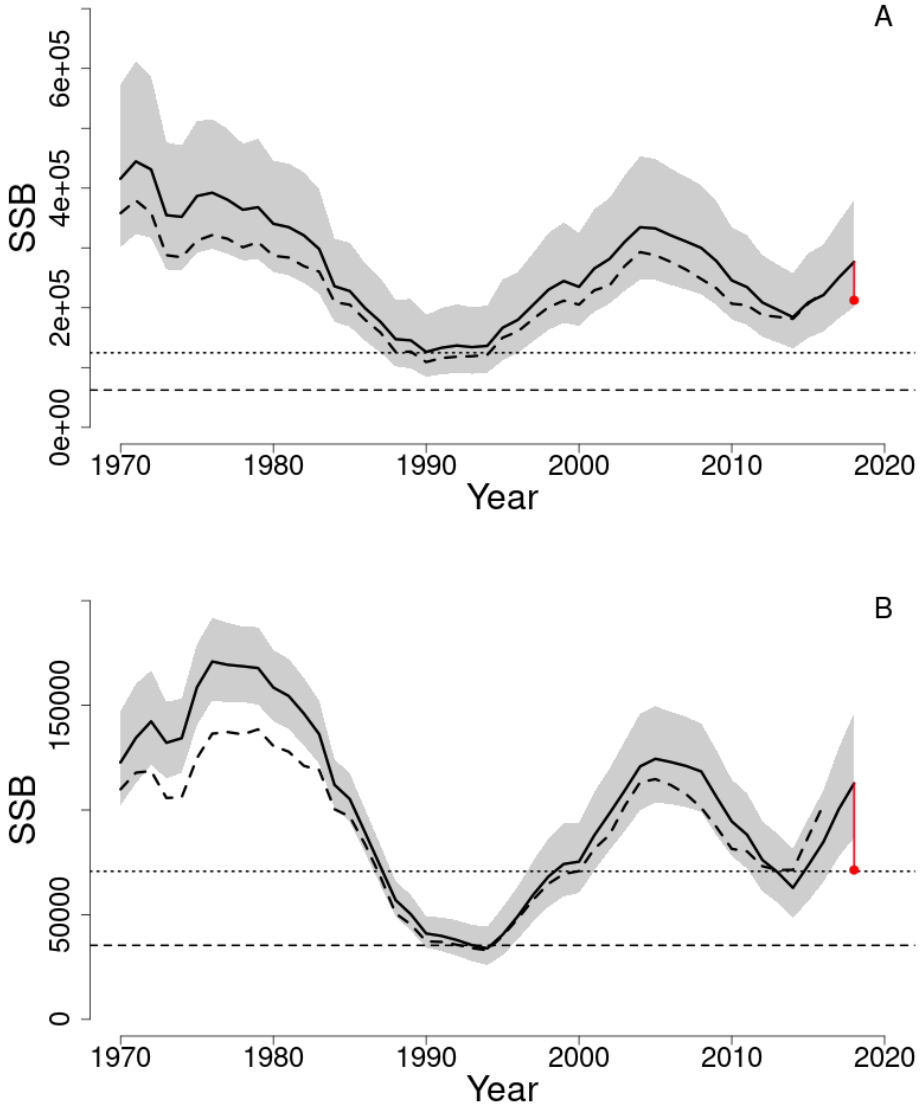

Figure 62: Estimated trends in the spawning stock biomass of pollock between 1970 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $SSB_{Threshold}$ (0.5 * SSB_{MSY} proxy; horizontal dashed line) as well as SSB_{Target} (SSB_{MSY} proxy; horizontal dotted line) based on the 2019 assessment models base (A) and flat sel sensitivity (B). Biomass was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% lognormal confidence intervals are shown.

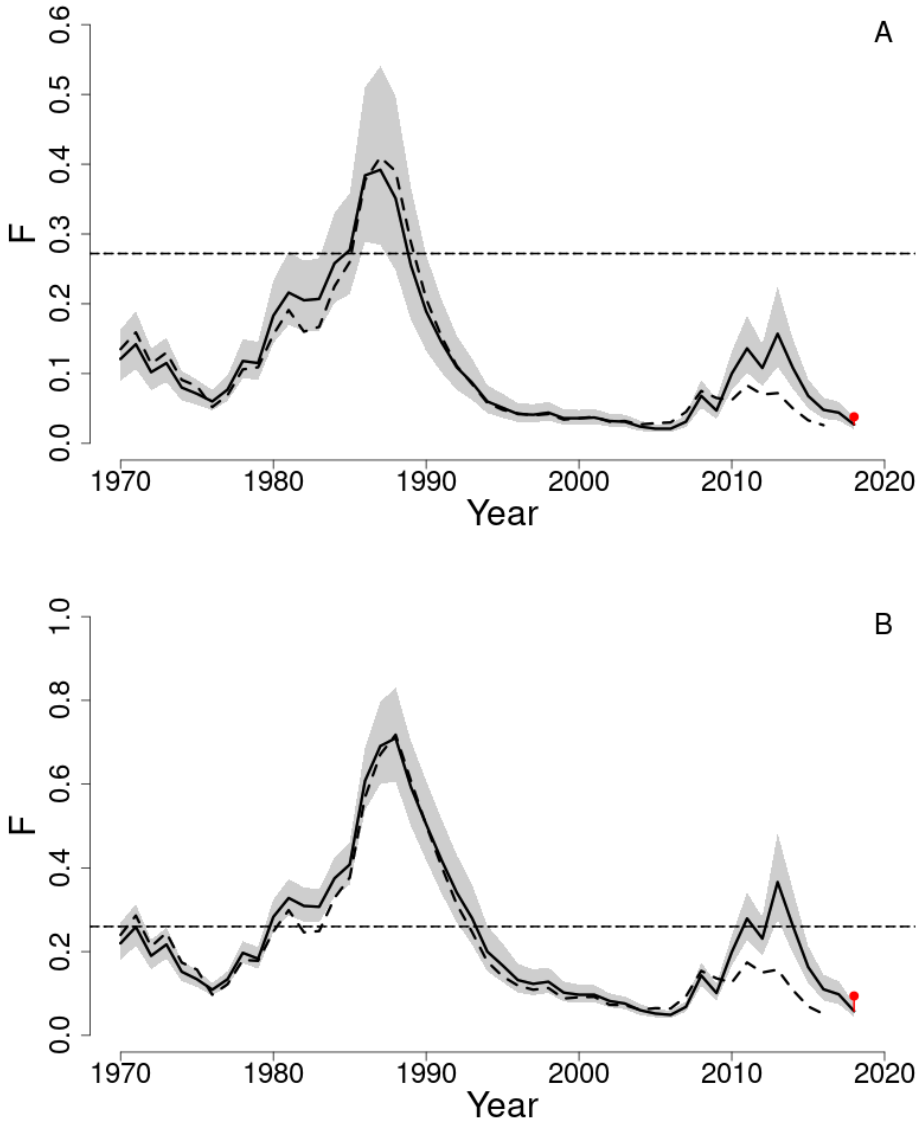

Figure 63: Estimated trends in age 5 to 7 average F (F_{AVG}) of pollock between 1970 and 2018 from the current (solid line) and previous (dashed line) assessment and the corresponding $F_{Threshold}$ (F_{MSY} proxy; dashed line) based on the 2019 assessment models base (A) and flat sel sensitivity (B). F_{AVG} was adjusted for a retrospective pattern and the adjustment is shown in red. The approximate 90% lognormal confidence intervals are shown.

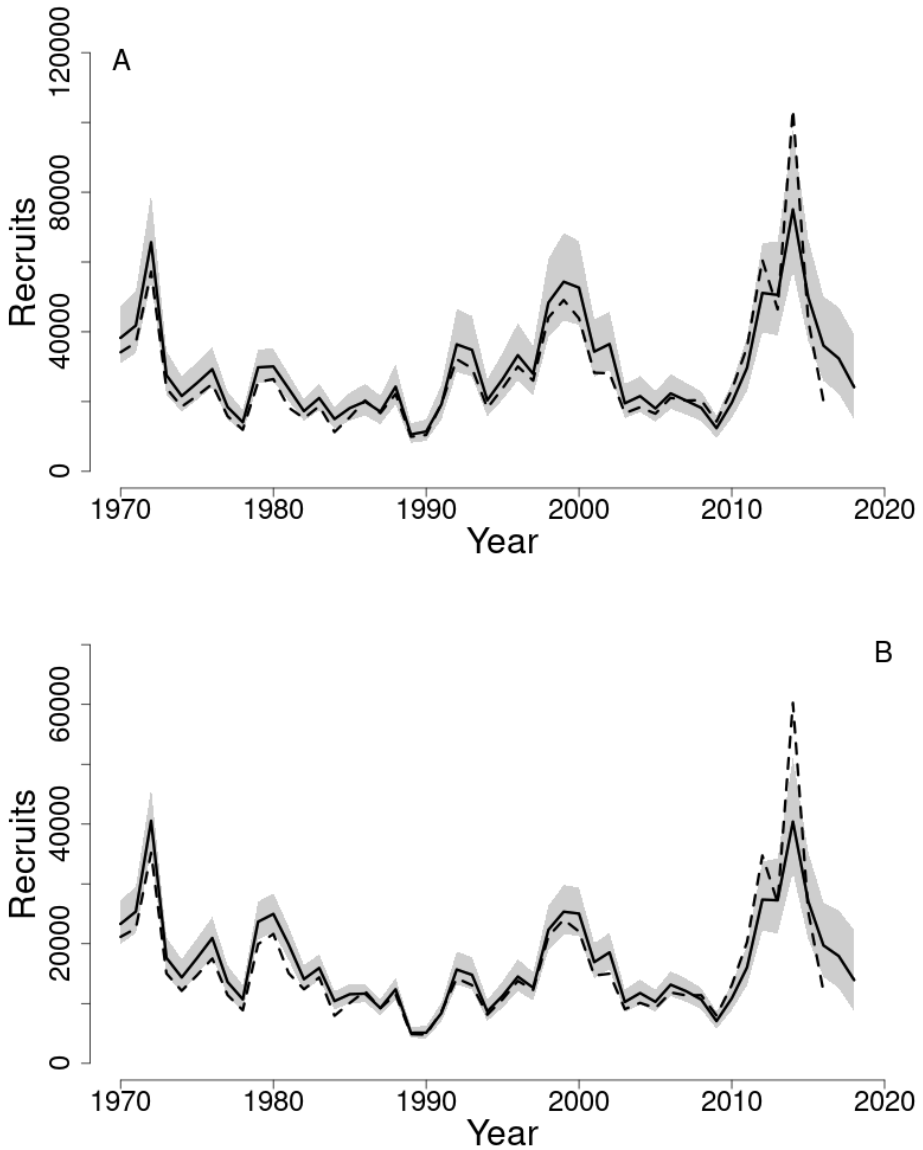

Figure 64: Estimated trends in age 1 recruitment (000s) of pollock between 1970 and 2018 from the current (solid line) and previous (dashed line) assessment for the assessment models base (A) and flat sel sensitivity (B). The approximate 90% lognormal confidence intervals are shown.

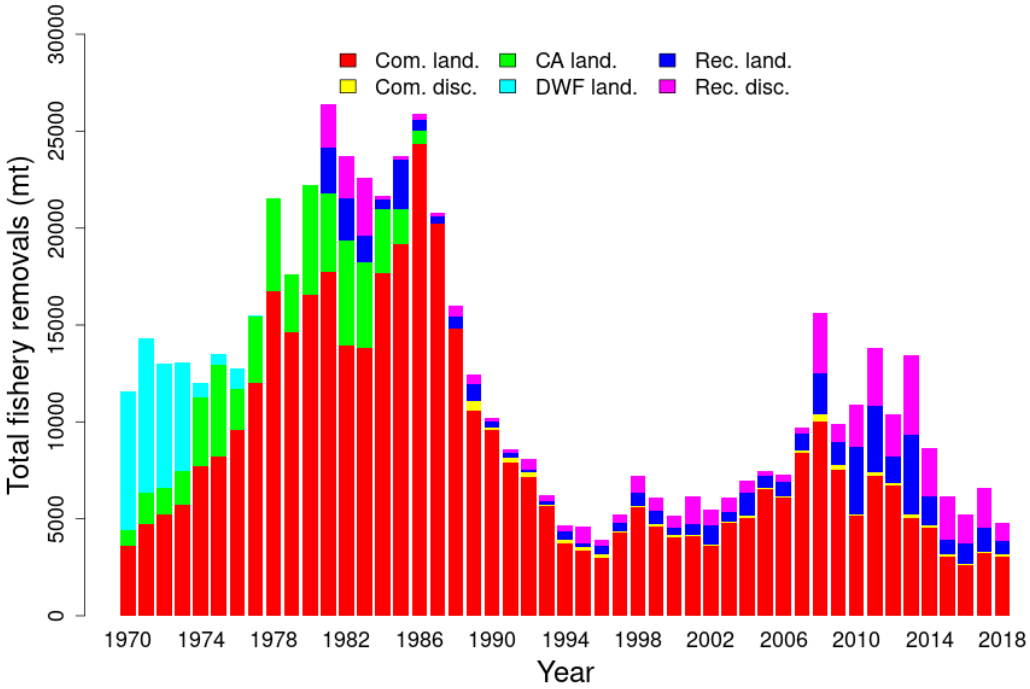

Figure 65: Total catch of pollock between 1970 and 2018 by fleet (commercial, Canadian, distant water fleet, and recreational) and disposition (landings and discards).

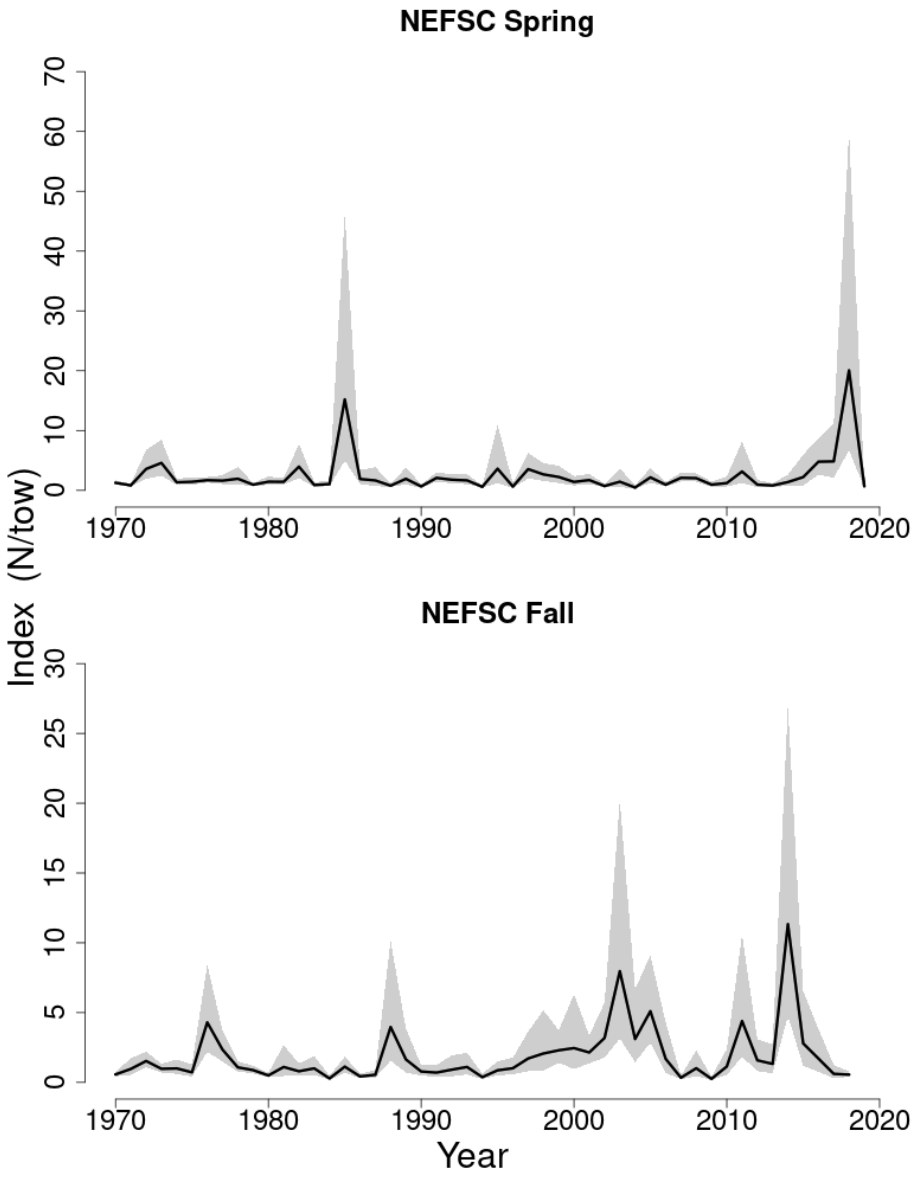

Figure 66: Indices of abundance for pollock from the Northeast Fisheries Science Center (NEFSC) spring (1970 to 2019) and fall (1970 to 2018) bottom trawl surveys. The approximate 90% lognormal confidence intervals are shown.

13 Atlantic halibut

Daniel Hennen

This assessment of the Atlantic halibut (Hippoglossus hippoglossus) stock is an update of the existing 2017 plan B assessment (Rago, 2018). This assessment updates commercial fishery catch data, commerical and survey indices of abundance, and the First Second Derivative (FSD) model through 2018. Reference points are unknown and have not been updated.

State of Stock: Based on this updated assessment, Atlantic halibut (*Hippoglossus hippoglossus*) stock status cannot be determined analytically due to a lack of biological reference points associated with the FSD method. Biomass (SSB) in 2018 was unknown. The 2018 fully selected fishing mortality was unknown.

Table 42: Catch and status table for Atlantic halibut. All weights are in (mt).

	2011	2012	2013	2014	2015	2016	2017	2018
			Data					
Commercial discards	27	41	42	26	23	31	27	46
Commercial landings	26	35	35	45	62	67	63	54
CA landings	29	32	38	33	30	34	34	56
Catch for Assessment	82	108	115	104	115	132	124	156
			Model Results					
Catch Multiplier	1.28	1.38	1.23	1.02	1.17	1.02	1.01	0.94
Catch Advice	75	104	148	142	106	135	135	126

Table 43: There are no current reference points for Atlantic halibut which is on a 'plan B' assessment that does not allow for the estimation of reference points. Therefore the status of the stock relative to overfishing and overfished status is unknown. Note: based on NOAA policy, the Agency previously decided the stock status was overfished and overfishing not occurring.

	2017	2019
F_{MSY} *proxy*	NA	
SSB_{MSY} (mt)	NA	
MSY (mt)	NA	
Overfishing	Unknown	Unknown
Overfished	Unknown	Unknown

Projections: Short term projections are not possible using the FSD approach. The FSD approach is based on applying a multiplier to the catch from the previous year and cannot be projected beyond the catch time series. The catch multiplier for 2019 resulting from the FSD model is 0.94

and the estimated catch for 2018 is 156.4 mt, which results in catch advice of 147.1 mt for 2019. The FSD model is explained in (Rago, 2018) and is graphically depicted in a document called 'FSDmodelResults.pdf', both are available at the data portal (SASINF).

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  The assessment model (FSD) used for Atlantic halibut is a 'plan B' assessment method. It uses recent trends in 3 abundance indices as well as recent changes in those trends to adjust the previous year's catch. For example, If the abundance indices are increasing, the catch will be adjusted up. If that increasing trend in abundance is increasing in magnitude over time, the adjustment to catch will be commensurately higher. The FSD method was rigorously tested in simulation (Rago, 2018) and should perform well for Atlantic halibut in the US. Sources of uncertainty in the FSD method include process error related to potential changes in stock productivity over time, the choice of relative weights for the control parameters used in the model, and the lag in information inherent in using change in trend as one of the control parameters, which requires dropping one data point from the regression fit to generate a comparison. Other sources of uncertainty include the observation error in the abundance indices. The FSD method also relies on the assumption that abundance can be described with linear dynamics, but that assumption should be relatively unimportant if the stock abundance is well below it's theoretical carrying capacity.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major?

  The FSD model does not support retrospective analysis.

- Based on this stock assessment, are population projections well determined or uncertain?

  The FSD model provides catch advice in the year following the terminal year of the input data. It is not intended to to project further ahead than one year. It is possible however to assume that catch in the year following the terminal year will equal the catch advice from the FSD model and that the population abundance indices will continue to follow the same trend and that the change in trend will be identical to the previous five years of data. These assumptions allow for a projection any number of years into the future. The relative quality of these projections degrades as the indices of abundance depart from the behavior of the most recent data available to the model.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  No changes were made beyond the inclusion of updated data.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the

F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

In the 2019 Atlantic halibut assessment, the catch efficiency studies and data were not used because not enough Atlantic halibut were caught to provide a comparison between the gear types and produce an estimate of catchability.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  Stock status cannot be determined and remains unchanged. Rago in his 2018 report argued that because the catch multiplier estimated in the FSD model had been greater than one for several years, that overfishing was unlikely. Because the catch multiplier is now less than one, overfishing may be the more likely determination in 2019. There is however, no way to credibly determine stock status without reference points.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The Atlantic halibut assessment could be improved with more precise fishery independent indices of abundance, additional age and length composition data, and a better understanding of stock structure. These would allow for alternative assessment methods, and potential development of a more sophisticated stock assessment model.

- Are there other important issues?

  The FSD method does not allow for the estimation of tradtional reference point quantities and thus the stock status cannot be determined. It is possible to infer that the stock is low relative to it's virgin biomass, which, based on historical catch records, was likely much higher than current abundances. It is unclear however, that biomass reference points based on historical abundance are useful for current management. There are indications that abundance has increased significantly over the last decade (Rago, 2018), which would support a hypothesis that the stock was not experiencing overfishing during that period. It should be noted however, that the FSD model has recently recommended reducing catch, which might be an indication that the stock no longer increasing.

13.1 Reviewer Comments: Atlantic halibut

The Atlantic halibut stock assessment was not reviewed by the 2019 Review Panel because it was determined to be a level 1 assessment at the AOP meeting in June of 2019 (Appendix 16.2), according to the stock assessment process adopted for this and future management track assessments (Appendix 16.3).

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://nefsc.noaa.gov/publications/

Col, L.A., Legault, C.M. 2009. The 2008 Assessment of Atlantic halibut in the Gulf of Maine Georges Bank region. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 09-08; 39 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Rago, P.J. 2018. Halibut Assessment Report for 2017 for New England Fishery Management Council, January 24, 2018.Unpublished, online at SASINF

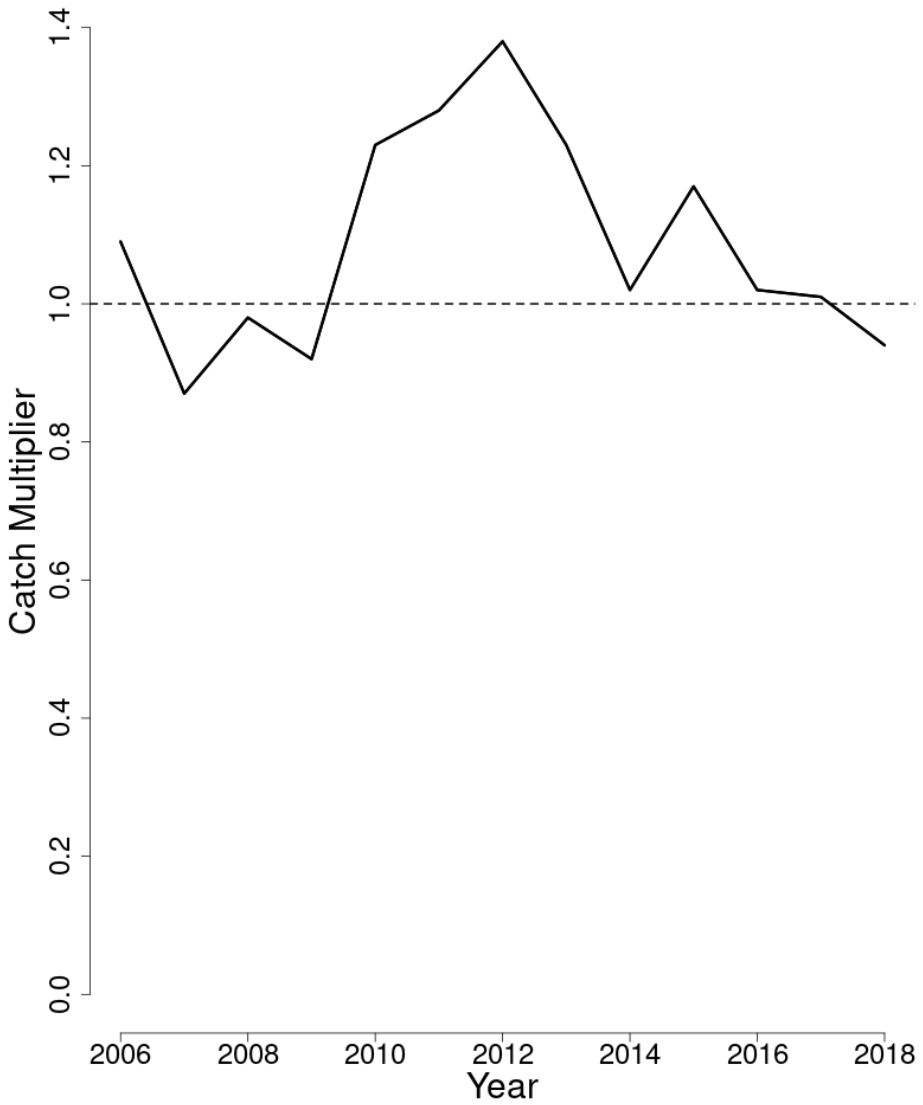

Figure 67: The catch multiplier resulting from the FSD model for Atlantic halibut between 2006 and 2018 from the current (solid line) assessment. A dashed line at 1 is added for reference.

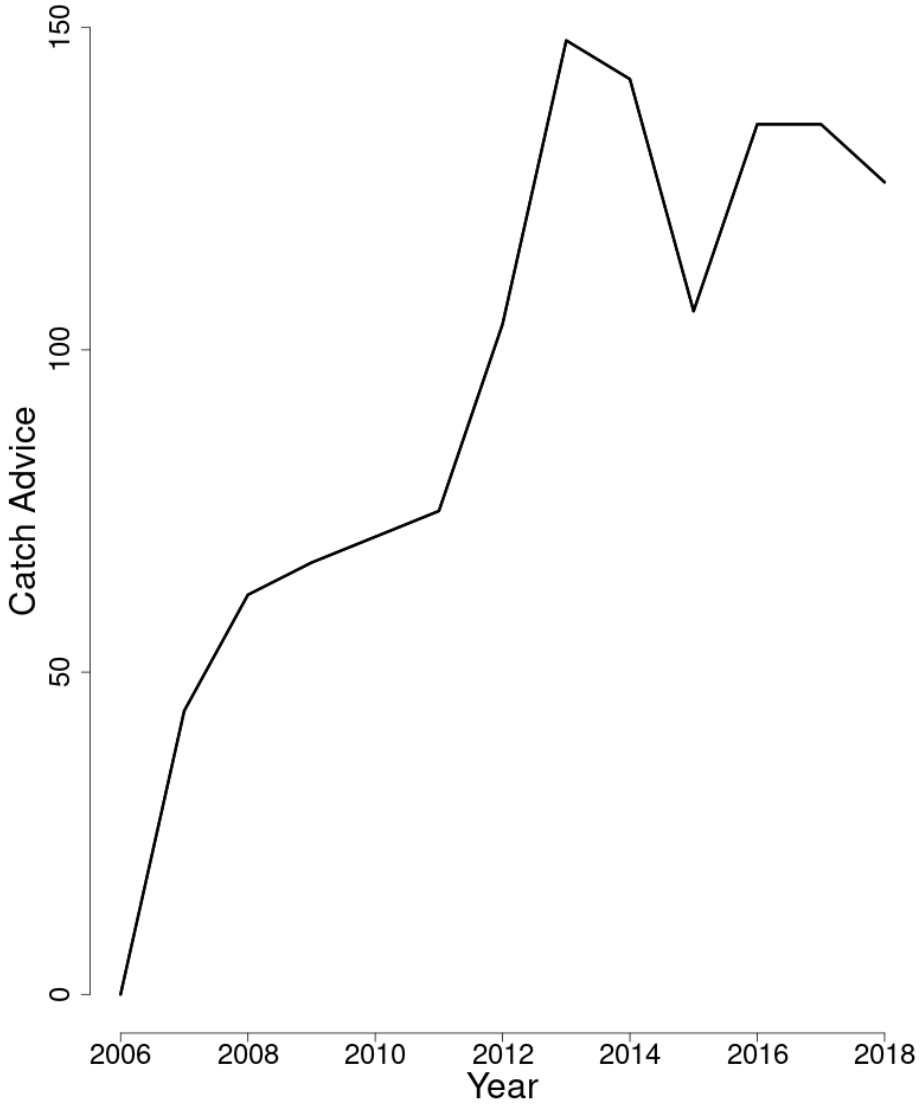

Figure 68: The catch advice resulting from multiplying catch and the catch multiplier from the FSD model for Atlantic halibut between 2006 and 2018 from the current assessment.

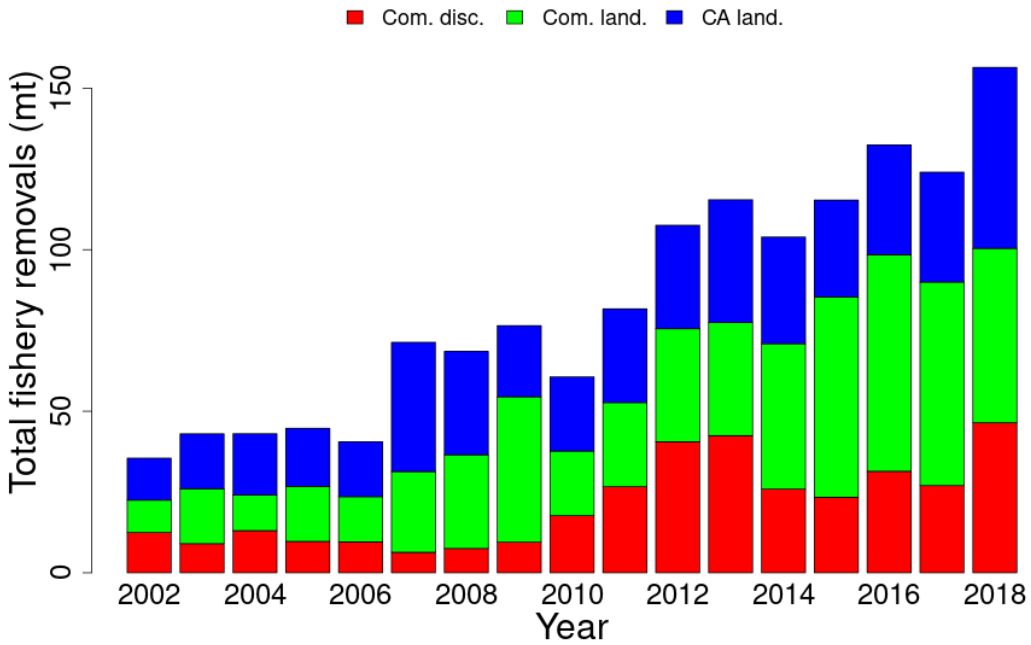

Figure 69: Total catch of Atlantic halibut between 2006 and 2018 by disposition (landings and discards).

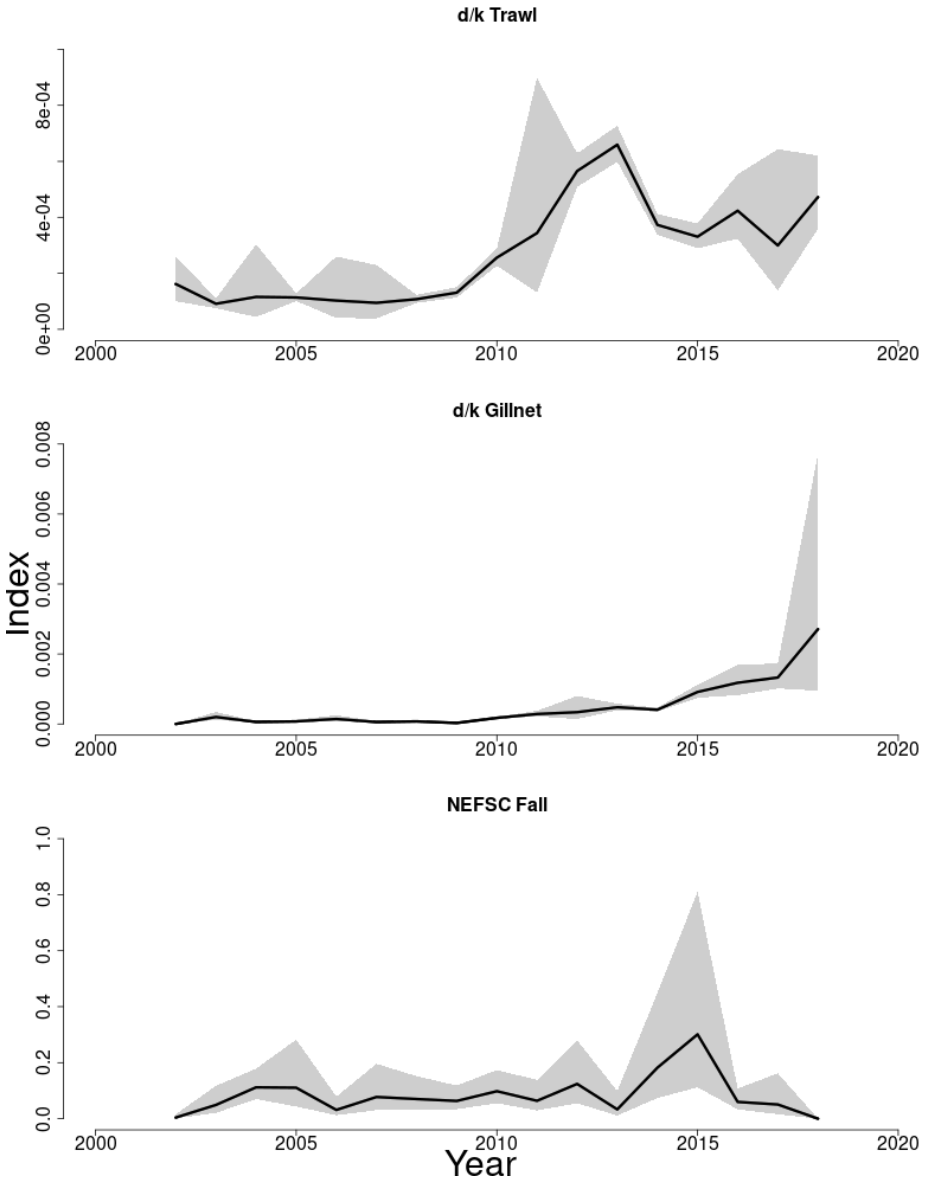

Figure 70: Indices of biomass for the Atlantic halibut between 2002 and 2018 for the Northeast Fisheries Science Center (NEFSC) fall bottom trawl survey and 2 discard ratio estimators. Discard mortality is assumed to be 0.76 for trawl gear and 0.3 for gillnet gear. The 90% lognormal confidence intervals are shown.

14 Gulf of Maine - Georges Bank windowpane flounder

Toni Chute

This assessment of the Gulf of Maine - Georges Bank windowpane flounder (Scophthalmus aquosus) stock is an update of the 2017 assessment which was based on survey and fishery data through 2016 (NEFSC 2017). Based on the 2017 assessment the stock was overfished, but overfishing was not ocurring. This assessment updates commercial fishery catch data, survey biomass indices, AIM model results, and reference points through 2018.

State of Stock: Based on this updated assessment, the Gulf of Maine - Georges Bank windowpane flounder (*Scophthalmus aquosus*) stock is overfished and overfishing is occurring (Figures 71-72). Retrospective adjustments were not made to the model results. The mean NEFSC fall bottom trawl survey index from years 2016, 2017 and 2018 (a 3-year moving average is used as a biomass index) was 0.248 kg/tow which is lower than the $B_{Threshold}$ of 1.745 kg/tow. The 2018 relative fishing mortality was estimated to be 0.335 kt per kg/tow which is higher than the F_{MSY} *proxy* of 0.185 kt per kg/tow. This F_{MSY} *proxy* was not accepted by the peer review panel. They recommended using the F_{MSY} *proxy* of 0.340 from the previous assessment as a reference point.

Table 44: Catch and model results table for Gulf of Maine - Georges Bank windowpane flounder. All landings and discard weights are rounded to the nearest metric ton. Less than half a metric ton has been landed annually since 2013. Biomass index (a 3-year moving average of the NEFSC fall bottom trawl survey index) is in units of kg/tow and relative F is in units of kt per kg/tow (catch in kt per kg/tow of the survey index).

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
					Data						
Commercial Discards	332	417	241	181	197	356	220	194	90	96	83
Commercial Landings	45	26	0	0	1	0	0	0	0	0	0
Total Catch	377	443	241	181	199	356	220	195	90	96	83
					Model Results						
Biomass Index	0.448	0.442	0.467	0.433	0.343	0.518	0.535	0.536	0.36	0.287	0.248
Relative F	0.843	1.004	0.516	0.418	0.581	0.687	0.411	0.364	0.25	0.334	0.335

Table 45: Reference points estimated in the 2017 assessment and in the current assessment update. F_{MSY} *proxy* is in units of kt per kg/tow. For previous assessments, the MSY proxy of 0.669 kt per kg/tow was rounded up to 0.700. The MSY proxy was re-calculated for this update and not rounded. The peer review panel did not recommend using the 2019 F_{MSY} *proxy* as a reference point for status determination.

	2017	2019
F_{MSY} *proxy*	0.340	0.185 (0.0001 - 0.726)
B_{MSY} *proxy* (kg/tow)	2.060	3.489
MSY proxy (mt)	700	647
Overfishing	No	Yes
Overfished	Yes	Yes

Projections: Short term projections from the AIM model are not used. Applying the updated F_{MSY} *proxy* (0.185) to the terminal year biomass index (0.248) produces a catch of 46 mt. Applying the F_{MSY} *proxy* of 0.340 from the 2017 Operational Assessment, as recommended by the peer review panel, to the 2019 terminal year biomass index produces a catch of 84 mt.

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Even though estimated catch has decreased in recent years, the survey index has not shown any resulting increase as a result despite evidence of regular recruitment from survey length frequencies. Since there has been a 'no possession' rule in place since 2010, between 99 and 100 percent of Gulf of Maine - Georges Bank windowpane flounder catch has consisted of estimated discards. These annual estimates have low CVs, however (averaging 0.162 since the year 2000).

  The AIM model fit for Gulf of Maine - Georges Bank windowpane flounder is poor.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}).

  The AIM (An Index Model) model used to estimate status of this stock does not allow estimation of a retrospective pattern.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  The GARM benchmark indicated that projections should not be made based on discards, so no projections are run for windowpane flounder. Northern windowpane flounder was declared overfished in the 2008 assessment (terminal data year 2007), and was supposed to be rebuilt by 2017. However the 2017 Operational Update indicated that the stock was still overfished. A new rebuilding plan was developed with F-rebuild equal to 70% F_{MSY} with a target of rebuilding by 2029.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the effect these changes had on the assessment and stock status.

  Several changes were made in the AIM model input for this Gulf of Maine - Georges Bank windowpane flounder assessment update. First, the entire time series of discards, including hindcast years, were re-estimated using SBRM. This changed the MSY proxy (median catch 1995-2001, a time period where the replacement ratio was fairly high) from 0.669 kt per kg/tow to 0.647 kt per kg/tow, not enough of a change to affect reference points very much. Also, estimated windowpane discards from the general category scallop fleet were added to the catch stream starting in 2002 (when observer coverage of this fleet increased) to more realistically reflect windowpane removals. This additional fleet represents an annual increase in estimated discards of up to 20% (2018) but averaging 5.5% since 2010. Re-estimating discards and adding the general category scallop fleet improved the fit of the AIM model slightly.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b; Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

  In this Gulf of Maine - Georges Bank windowpane flounder assessment, experimental catchability estimates were used to calculate a survey swept area biomass for the alternative Plan B assessment. The primary AIM assessment provides only relative indices of abundance and fishing mortality, and so catchability estimates would not have affected those results.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  Initial results from this assessment indicated that the stock status of Gulf of Maine - Georges Bank windowpane flounder changed since the previous assessment from overfished with no overfishing to overfished with overfishing occurring. The peer review panel did not recommend accepting the F_{MSY} proxy produced for this assessment and recommended using the F_{MSY} proxy from the 2017 Operational Assessment for status determination. This changed the recommended status to overfished with no overfishing occurring, consistent with the 2017 assessment results. When the estimated catch decreases and the survey index also decreases, it is difficult to model. There is no obvious reason why Gulf of Maine - Georges Bank windowpane flounder have not responded to the reduction in catch by increasing in abundance.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Since the year 2000, Gulf of Maine - Georges Bank windowpane flounder has shown a decreasing trend in survey indices despite reductions in catch. In 2008 (with data through

2007) the stock was declared overfished and still remains below the biomass threshold despite recent catch estimates being the very lowest in the time series. According to the State of the Ecosystem Report for 2019, both male and female windowpane flounder are currently showing high condition indices. There are also new recruits regularly present in the fall bottom trawl survey catches.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The AIM model fit for Gulf of Maine - Georges Bank windowpane flounder might be improved by identifying any potential sources of mortality or addional removals from the population. There may be catches (such as from Canadian fishing on Georges Bank), discards, or incidental mortality unaccounted for in the model. There may also be value in looking carefully at the Gulf of Maine - Georges Bank windowpane flounder stock area strata set to see if there might be reason to change it.

- Are there other important issues?

  None.

14.1 Reviewer Comments: Gulf of Maine - Georges Bank windowpane flounder

14.1.1 Review Panel Summary

Updated fishery and survey data from the completed Plan A operational assessment and previously estimated reference points are technically appropriate for assessing stock status, providing scientific advice and successfully addresses the terms of reference.

The Panel accepted the use of updated fishery and survey data for calculating the current biomass index and relative fishing mortality rate (F), but did not accept the updated F_{MSY} estimate from the AIM model. There was concern about using the estimated F_{MSY} proxy in determining management actions for this update, because the updated relationship of biomass replacement to relative F is uninformative. There are also numerous negative residuals at the end of the assessment time series where observed survey values are less that model predictions. Furthermore, F_{MSY} estimates have been unstable over the last several assessments, with a large change in the estimate between the 2017 and 2019 update assessments. The stock is not responding as expected from with low catch quotas. More specifically, low catches are not leading to stock growth. However, the Panel felt that the declining trend in biomass despite low catches was captured by the empirical approach and felt the previously estimated reference points (i.e., the estimates from the 2017 update assessment) provide an objective way to derive catch advice for this stock.

14.1.2 Operational Stock Assessment Terms of Reference

1. Update fishery-dependent data (landings, discards, catch-at-age, etc.) and fishery-independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment.

The data were fully updated for use in this assessment. The Panel supports the change in estimation method for discards to be consistent with the Standardized Bycatch Reporting Methodology commonly used in the region and the addition of general category scallop fishery discard estimates in this assessment. The addition of general category scallop fishery discard estimates aligns the assessment with the current quota monitoring of this stock.

2 a.) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A").Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

The empirical components of the updated Plan A assessment are recommended for use, but updated estimates of reference points are not recommended for use. The 2017 estimates of reference points are recommended until the 2020 research track for index-based methods.

2 b.) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

Plan B was not considered for use.

3. Update the values of biological reference points (BRPs) for this stock.

The updated estimate of the F_{MSY} proxy was deemed unusable for determining management actions. The biomass replacement ratio to relative F regression is poorly informed in part due to an extended period over which stock sizes have remained low despite low relative F values. The stock is not increasing as expected under low catch and relative F, indicating that productivity of the stock has changed or that the data (catch and indices) are not reflective of the actual removals or population in recent years.

There are few observations of relative F in the updated F_{MSY} regression that are less than the estimate of F_{MSY}) so there is little information informing the estimate of F_{MSY}. The 2017 update assessment had a stronger relationship between replacement rate and relative F. We therefore recommend using the 2017 estimate of F_{MSY} and associated B_{MSY} estimate for consistency. The overfishing definition and entire assessment approach should be re-considered at the 2020 research track workshop on index-based assessments, particularly because the updated discard estimates include fleet components that were not included in the previous assessments.

4 a.) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

Using the 2017 estimates of F_{MSY} and SSB_{MSY} proxies the stock is overfished, but overfishing is not taking place. This is the same status as determined from the last assessment.

4 b.) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

The population is low and not increasing even under conditions of apparently low catch and fishing mortality, despite fish being in good condition and regular recruitment apparent in survey data.

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

No projections were provided.

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

There are essentially no landings, so the assessment is relying almost entirely on discard information to estimate fishery removals and determine stock status. This creates a large amount of uncertainty, because discards are poorly estimated for some fleets.

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Most recent assessment update:
Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. doi: 10.7289/V5/RD-NEFSC-17-17.

Most recent benchmark assessment:
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, MA, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p.

State of the Ecosystem Reports, 2019, Mid-Atlantic Region. Available at:
https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

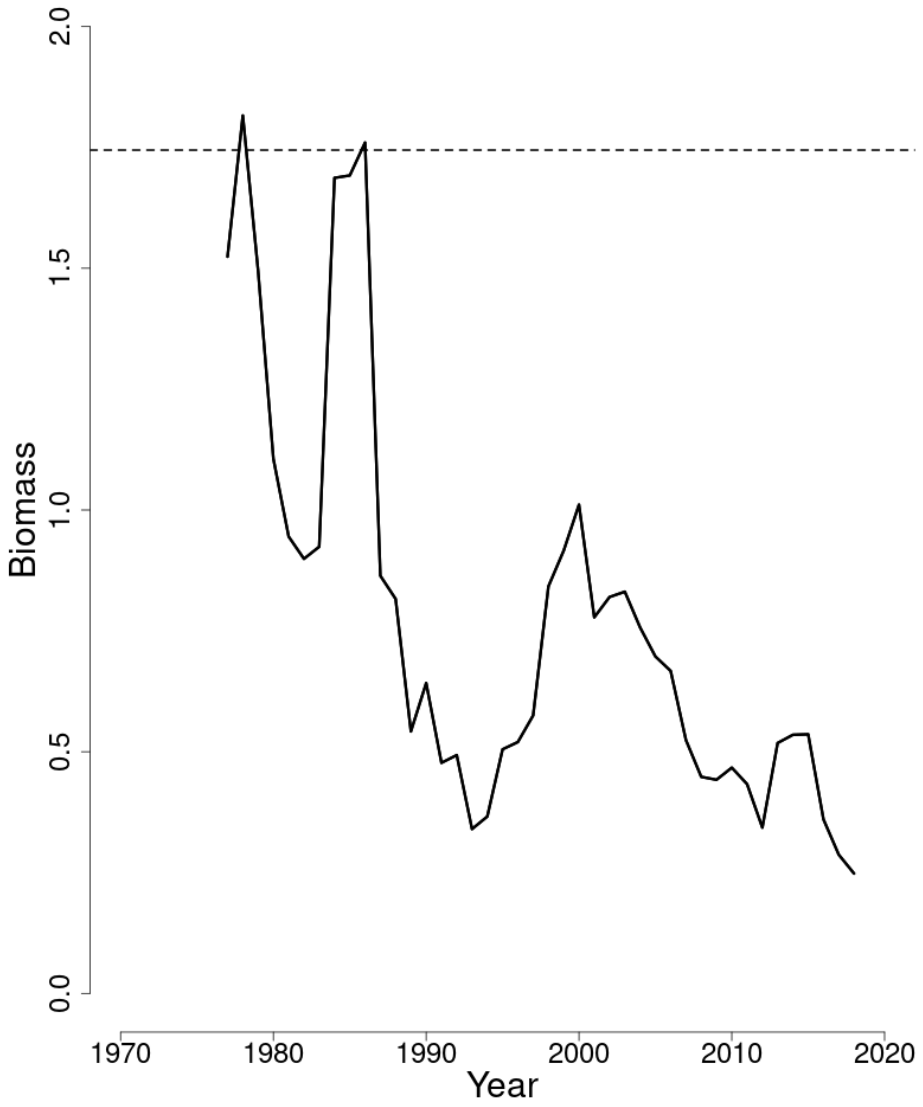

Figure 71: Trends in the biomass index (a 3-year moving average of the NEFSC fall bottom trawl survey index) of Gulf of Maine - Georges Bank windowpane flounder between 1975 and 2018 from the current assessment, and the corresponding $B_{Threshold} = \frac{1}{2}\ B_{MSY}$ *proxy* = 1.745 kg/tow (horizontal dashed line).

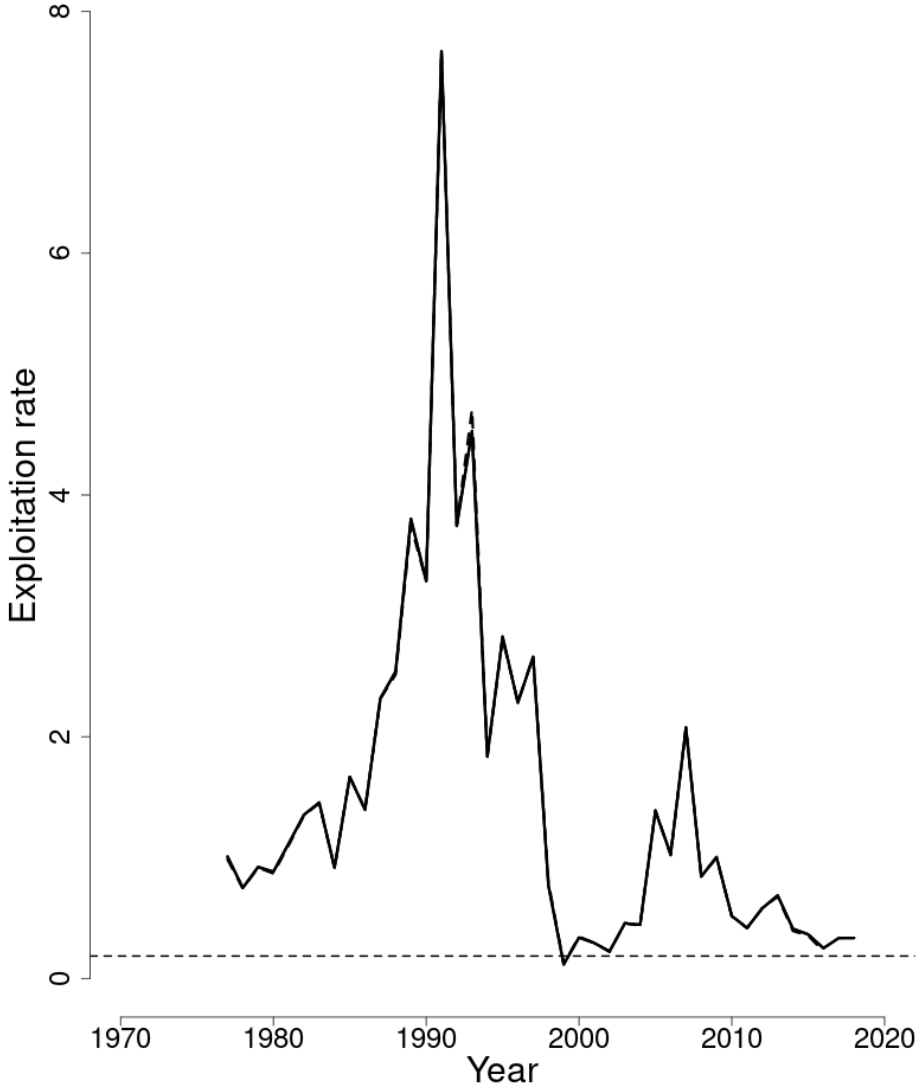

Figure 72: Trends in estimated relative fishing mortality of Gulf of Maine - Georges Bank windowpane flounder between 1975 and 2018 from the current assessment, and the corresponding F_{MSY} *proxy* = 0.185 (horizontal dashed line).

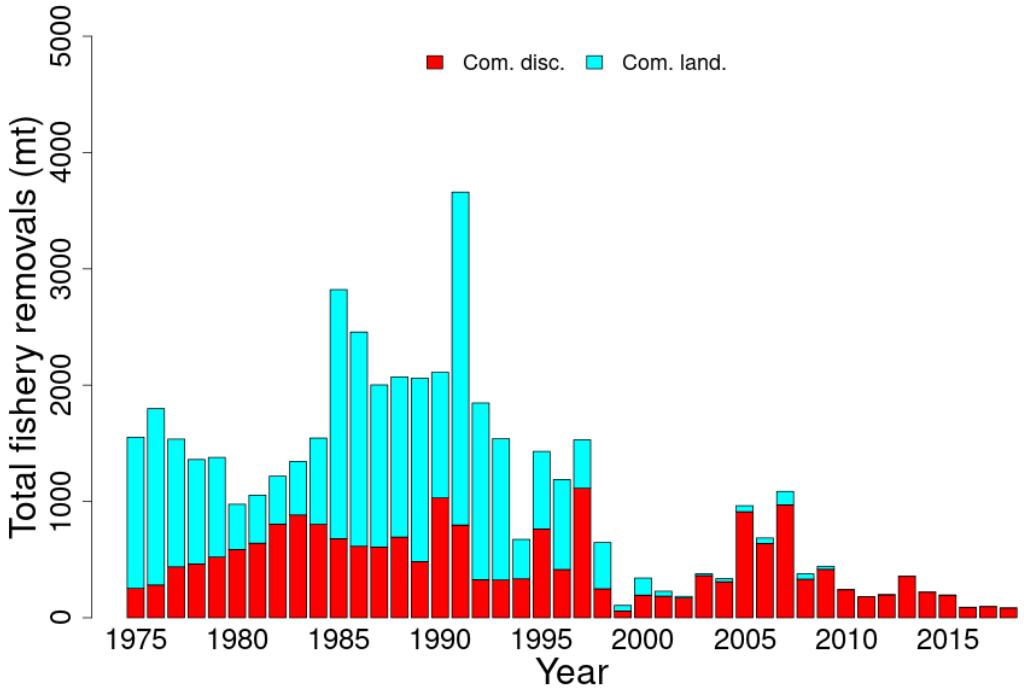

Figure 73: Total catch of Gulf of Maine - Georges Bank windowpane flounder between 1975 and 2018 by disposition (landings and discards).

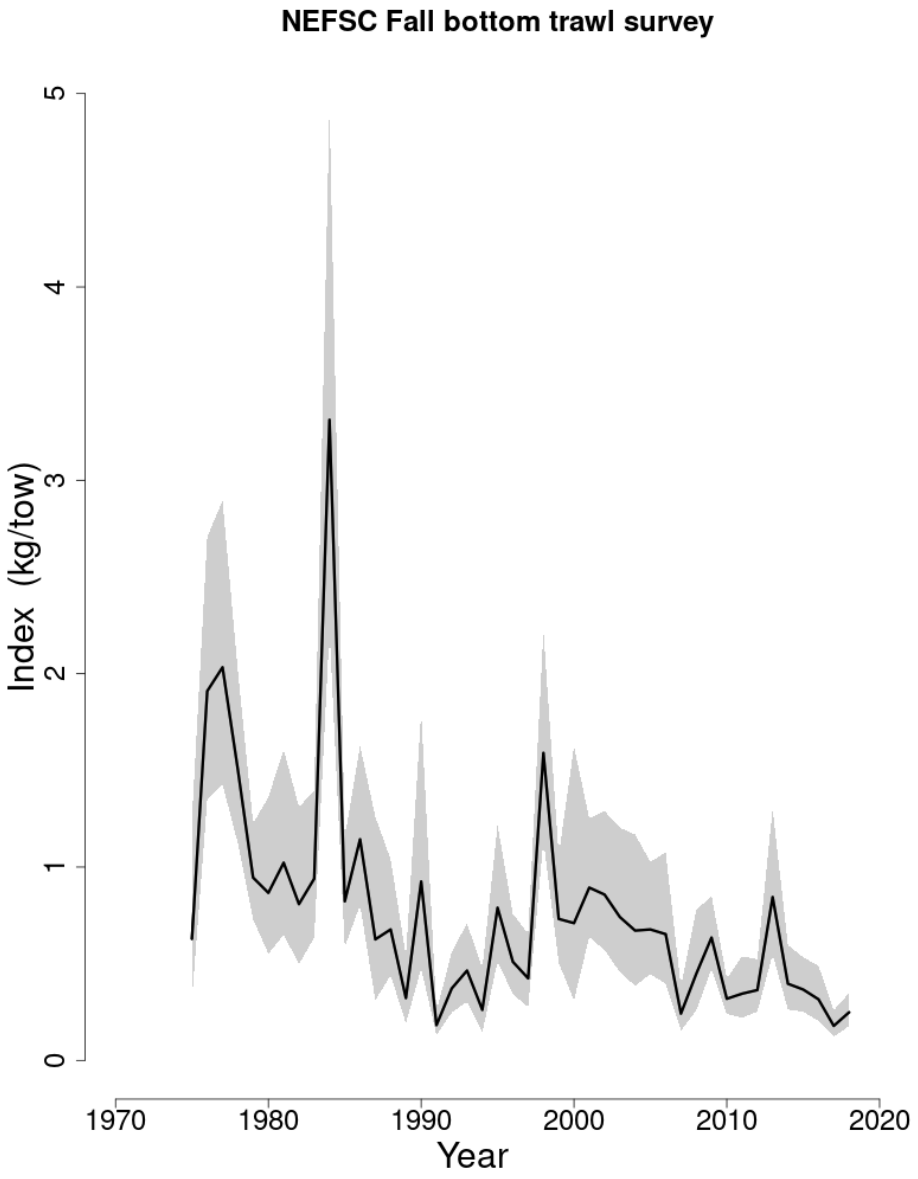

Figure 74: NEFSC fall bottom trawl survey indices in kg/tow for Gulf of Maine - Georges Bank windowpane flounder between 1975 and 2018. The approximate 90% lognormal confidence intervals are shown.

15 Southern New England - mid-Atlantic windowpane flounder

Toni Chute

This assessment of the southern New England - mid-Atlantic windowpane flounder (Scophthalmus aquosus) stock is an update of the 2017 assessment which was based on fishery and survey data through 2016 (NEFSC 2017). Based on the 2017 assessment the stock was not overfished, and overfishing was not ocurring. This assessment updates commercial fishery catch data, survey indices of abundance, AIM model results, and reference points through 2018.

State of Stock: Based on this updated assessment, the southern New England - mid-Atlantic windowpane flounder (*Scophthalmus aquosus*) stock is not overfished and overfishing is not occurring (Figures 75-76). Retrospective adjustments were not made to the model results. The mean NEFSC fall bottom trawl survey index from years 2016, 2017, and 2018 (a 3-year moving average is used as a biomass index) was 0.319 (kg/tow) which is higher than the $B_{Threshold}$ of 0.094 (kg/tow). The 2018 relative fishing mortality was estimated to be 1.632 (kt per kg/tow) which is lower than the F_{MSY} *proxy* of 1.780 (kt per kg/tow).

Table 46: Catch and model results table for southern New England - mid-Atlantic windowpane flounder. All landings and discard weights are rounded to the nearest metric ton. Biomass index (a 3-year moving average of the NEFSC bottom trawl survey index) is in units of kg/tow and relative F is in units of kt per kg/tow (catch in kt per kg/tow of the survey index).

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
					Data						
Commercial Discards	247	410	459	466	788	709	566	547	580	545	503
Commercial Landings	75	55	53	32	29	22	14	22	13	13	17
Total Catch	322	465	513	498	817	731	580	569	593	558	520
					Model Results						
Biomass Index	0.204	0.245	0.345	0.435	0.517	0.464	0.413	0.318	0.329	0.334	0.319
Relative F	1.582	1.9	1.485	1.144	1.581	1.574	1.405	1.789	1.801	1.671	1.632

Table 47: Reference points estimated in the 2012 assessment and in the current assessment update. F_{MSY} *proxy* is in units of kt per kg/tow. For previous assessments, the MSY proxy of 0.485 kt per kg/tow was rounded up to 0.500. The MSY proxy was re-calculated for this update and not rounded.

	2017	2019
F_{MSY} *proxy*	1.918	1.780 (1.046 - 2.191)
B_{MSY} *proxy* (kg/tow)	0.261	0.187
MSY proxy (mt)	500	333
Overfishing	No	No
Overfished	No	No

Projections: Short term projections from the AIM model are not used. Applying the updated F_{MSY} *proxy* (1.780) to the terminal year biomass index (0.319) produces a catch of 568 mt.

Special Comments:

- What are the most important sources of uncertainty in this stock assessment? Explain, and describe qualitatively how they affect the assessment results (such as estimates of biomass, F, recruitment, and population projections).

  Since there has been a 'no possession' rule in place since 2010, commercial windowpane landings have been extremely low. As a result, in recent years over 95% of the catch input to the model has been estimated discards. The CVs for these estimates have been small, however, with a mean of 0.21 since 2000, so it is unlikely discards are being severely overestimated or underestimated.

- Does this assessment model have a retrospective pattern? If so, is the pattern minor, or major? (A major retrospective pattern occurs when the adjusted SSB or F_{Full} lies outside of the approximate joint confidence region for SSB and F_{Full}).

  The AIM (An Index Model) model used to estimate status of this stock does not allow estimation of a retrospective pattern.

- Based on this stock assessment, are population projections well determined or uncertain? If this stock is in a rebuilding plan, how do the projections compare to the rebuilding schedule?

  The GARM benchmark indicated that projections should not be made based on discards, so no projections are run for windowpane flounder.

- Describe any changes that were made to the current stock assessment (e.g., catch efficiency studies), beyond incorporating additional years of data and the affect these changes had on the assessment and stock status.

  Several changes were made in the AIM model input for this southern New England - mid-Atlantic windowpane flounder assessment update. First, the entire time series of discards, including hindcast years, were re-estimated using SBRM. This changed the MSY proxy (median catch 1995-2001, a time period where the replacement ratio was fairly high) from 0.48 kt per kg/tow to 0.33 kt per kg/tow. The lower MSY proxy led to a lower Bmsy proxy. Also, the estimated windowpane discards from the general category scallop fleet were added to the catch stream starting in 2002 (when observer coverage of this fleet increased) to more realistically reflect windowpane removals. These additional discards represent an annual increase of about eight percent, although since 2010 it has only been about five percent of total windowpane discards.

  In addition, the NEFSC fall bottom trawl survey was unable to cover any of the southern New England - mid-Atlantic windowpane flounder strata in 2017, leaving a missing value in the survey index data. To impute a value for 2017, the mean survey biomass per tow values from 2016 and 2018 were averaged, by stratum, and a stratum-weighted index was calculated from these values.

  Methods to estimate relative catch efficiency, and its uncertainty, for rockhopper and chainsweep gears for the NEFSC bottom trawl survey and generate calibrated swept area numbers at length and biomass estimates are described in Miller 2013; Miller et al. 2017a,b;

Miller 2018. The data came from studies carried out in 2015, 2016, and 2017 aboard the F/V Karen Elizabeth twin trawl vessel and focused primarily on flatfish species. Models took into account body size and diel effects on relative efficiency. The best performing model was used to convert bottom trawl survey numbers at length into predicted catches using chainsweep gear, followed by estimation of calibrated stratified mean swept area numbers at length and calibrated biomass estimates.

In this southern New England - mid-Atlantic windowpane flounder assessment, experimental catchability estimates were used to calculate a survey swept area biomass for the alternative Plan B assessment. The primary AIM assessment provides only relative indices of abundance and fishing mortality, and so catchability estimates would not have affected those results.

- If the stock status has changed a lot since the previous assessment, explain why this occurred.

  The stock status of southern New England - mid-Atlantic windowpane flounder has not changed since the previous assessment.

- Provide qualitative statements describing the condition of the stock that relate to stock status.

  Since 2012, southern New England - mid-Atlantic windowpane flounder survey biomass indices have declined by half from 0.596 kg/tow to 0.266 kg/tow. However, the larger trend has been upward since the series low of 0.039 kg/tow in 1993. Catch and relative F have been stable. The replacement ratio model output has been bouncing around one since 1994, and the 2018 estimate of 0.81 is higher than 30% of the values since 1994. The stock was declared overfished in 2005, but had recovered by the 2008 assessment update, so there is a recent history of the stock falling below reference points for biomass, but also having the ability to recover. Overfishing was occurring in 2007 (the final year of data used for the 2008 assessment) but was not occurring in the three most recent assessment updates. According to the State of the Ecosystem Report for 2019, female windowpane flounder (there are no data for males) in the mid-Atlantic are in good body condition. Survey length frequencies indicate ongoing new recruits to the stock.

- Indicate what data or studies are currently lacking and which would be needed most to improve this stock assessment in the future.

  The AIM model fit is presently good with a randomization test indicating the correlation between ln(relative F) and ln(replacement ratio), a measure of the relationship between catch and survey index values, is significant ($p = 0.001$) so it is not clear what new information would help acheive better results from this model. There has been some ageing work for southern windowpane done at Virginia Institute of Marine Science which we might explore for use in an age-based model such as ASAP.

- Are there other important issues?

  None.

15.1 Reviewer Comments: Southern New England - mid-Atlantic windowpane flounder

The southern New England - mid-Atlantic windowpane flounder stock assessment was not reviewed by the 2019 Review Panel because it was determined to be a level 1 assessment at the AOP meeting in June of 2019 (Appendix 16.2), according to the stock assessment process adopted for this and future management track assessments (Appendix 16.3).

References:

Miller, T. J. 2013. A comparison of hierarchical models for relative catch efficiency based on paired-gear data for U.S. northwest Atlantic fish stocks. Canadian Journal of Fisheries and Aquatic Sciences 70(9): 1306-1316,

Miller, T. J., Martin, M. Politis, P., Legault, C. M., Blaylock, J. 2017a. Some statistical approaches to combine paired observations of chain sweep and rockhopper gear and catches from NEFSC and DFO trawl surveys in estimating Georges Bank yellowtail flounder biomass. TRAC Working Paper 2017/XX. 36. pp.,

Miller, T. J., Richardson, D. E., Politis, P. Blaylock, J. 2017b. NEFSC bottom trawl catch efficiency and biomass estimates for 2009-2017 for 8 flatfish stocks included in the 2017 Northeast Groundfish Operational Assessments. Working paper. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 11-15, 2017.,

Miller, T. J., Politis, P., Blaylock, J., Richardson, D., Manderson, J., Roebuck, C. 2018. Relative efficiency of a chain sweep and the rockhopper sweep used for the NEFSC bottom trawl survey and chainsweep-based swept area biomass estimates for 11 flatfish stocks. SAW 66 summer flounder Data/Model/Biological Reference Point (BRP) meeting. National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA. September 17-21, 2018.,

Most recent assessment update:
Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. doi: 10.7289/V5/RD-NEFSC-17-17.

Most recent benchmark assessment:
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, MA, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p.

State of the Ecosystem Reports, 2019, Mid-Atlantic Region. Available at:
https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

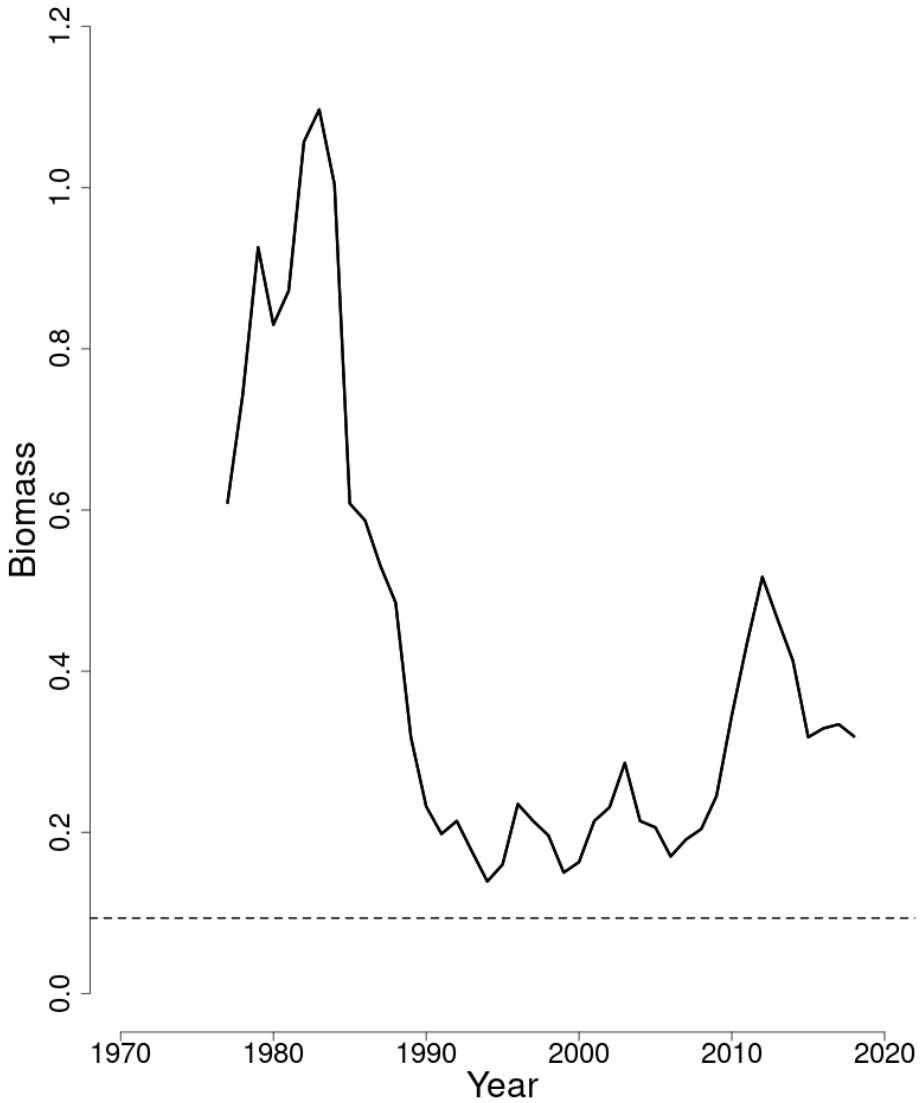

Figure 75: Trends in the biomass index (a 3-year moving average of the NEFSC fall bottom trawl survey index) of southern New England - mid-Atlantic windowpane flounder between 1975 and 2018 from the current assessment, and the corresponding $B_{Threshold} = \frac{1}{2}\ B_{MSY}$ *proxy* = 0.094 kg/tow (horizontal dashed line).

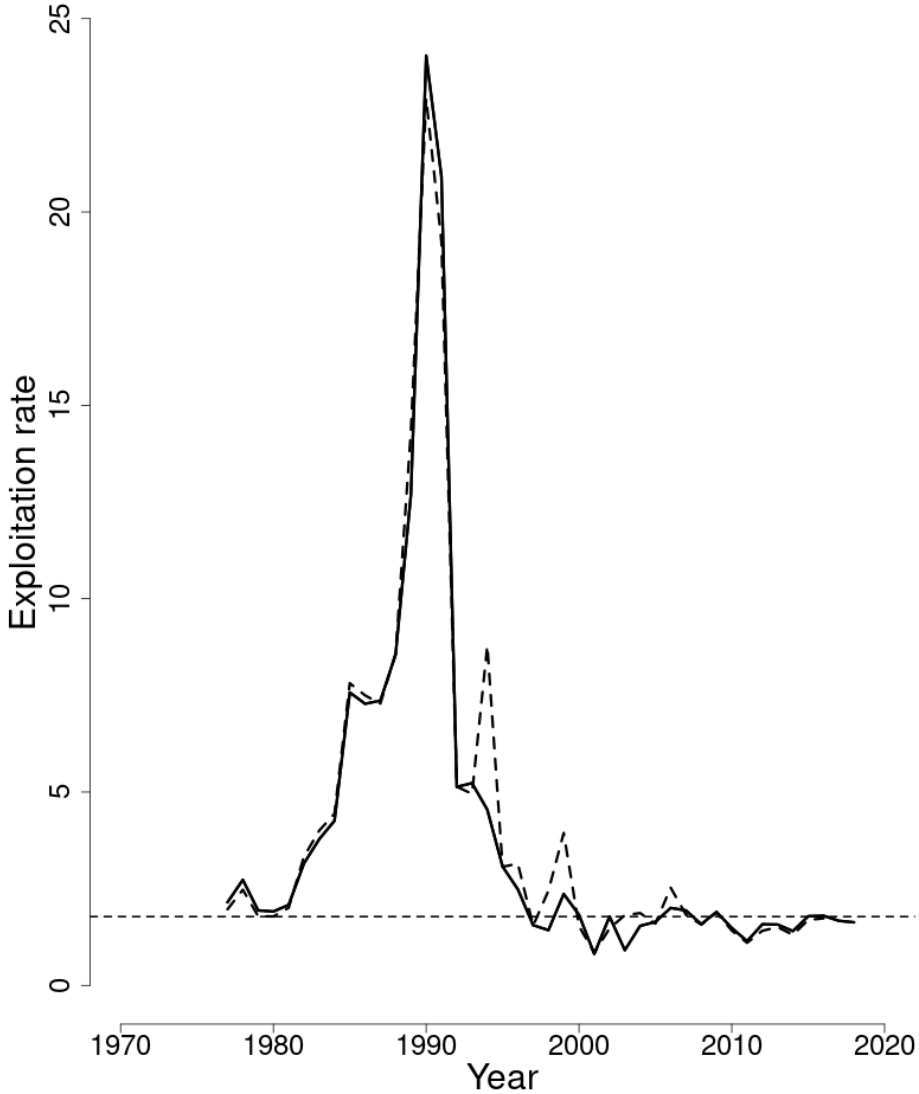

Figure 76: Trends in relative fishing mortality of southern New England - mid-Atlantic windowpane flounder between 1975 and 2018 from the current assessment using re-estimated discards (solid line), and the 2017 assessment (dashed line). The corresponding F_{MSY} *proxy*=1.78 is shown by the horizontal line.

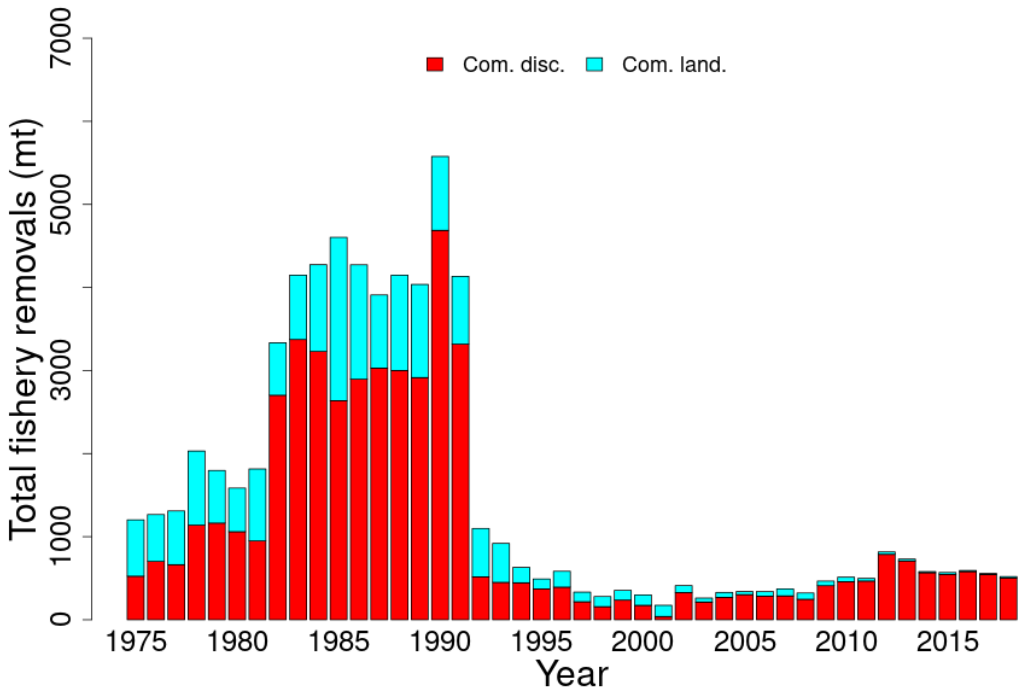

Figure 77: Total catch of southern New England - mid-Atlantic windowpane flounder between 1975 and 2018 by disposition (landings and discards).

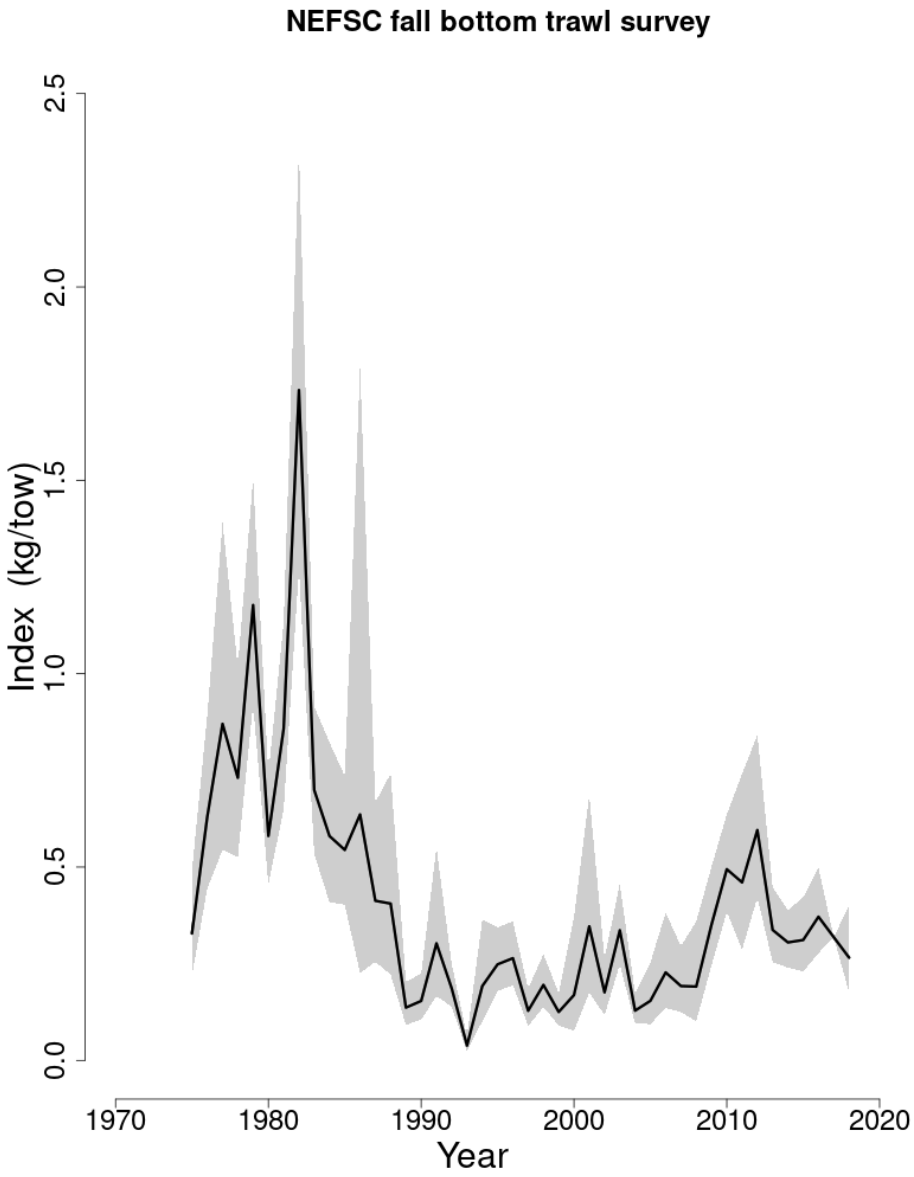

Figure 78: NEFSC fall bottom trawl survey indices in kg/tow for southern New England - mid-Atlantic windowpane flounder between 1975 and 2018. The approximate 90% lognormal confidence intervals are shown.

16 Appendices

16.1 Generic terms of reference for operational assessments

1. Update all fishery dependent data (landings, discards, catch at age, etc.) and all fishery independent data (research survey information) that had been used in the previous accepted assessment. Also, describe and present any new or revised data sets that are being used in the assessment[2].

2. (a) Estimate annual fishing mortality, recruitment, and stock size for the time series ("Plan A"). Include estimates of uncertainty, retrospective analyses (both historical and within-model), and bridge runs to sequentially document any changes from the previously accepted model to the updated model proposed for this peer review.

   (b) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management. "Plan B" will be presented for peer review only if the "Plan A" assessment were to not pass review.

3. Update the values of biological reference points (BRPs) for this stock.

4. (a) Recommend what stock status appears to be based on comparison of assessment results to BRP estimates.

   (b) Include qualitative descriptions of stock status based on simple indicators/metrics (e.g., age- and size-structure, temporal trends in population size or recruitment indices, etc.).

5. Perform short-term (3-year) population projections. The projection results should include an estimate of the catch at F_{MSY} or at an F_{MSY} proxy (i.e. this catch represents the overfishing level, OFL) as well as its statistical distribution (i.e., probability density function).

6. Comment on research areas or data issues to consider that might lead to improvements when this stock is assessed again in the future.

Source: NRCC. 2011. A new process for assessment of managed fishery resources off the Northeastern United States. Internal Report. With edits made by NEFSC on 6/3/2019.

[2] Major changes from the previous stock assessment require pre-approval by the Assessment Oversight Panel.

16.2 Summary of Assessment Oversight Panel Meeting

June 20, 2019 Woods Hole, Massachusetts The NRCC Assessment Oversight Panel (AOP) met to review the operational stock assessment plans for 14 stocks/species. The stock assessments for these stocks/species will be peer reviewed during a meeting from September 9-13, 2019.

AOP members

- Gary Nelson, Ph.D., Atlantic States Marine Fisheries Commission, Massachusetts Division of Marine Fisheries
- Jason McNamee, Ph.D., Chair NEFMC Scientific and Statistical Committee, RI Division of Environmental Management
- Paul Rago, Ph.D., member of the MAMFC Scientific and Statistical Committee, NOAA Fisheries (retired)
- Russell W. Brown, Ph.D., Northeast Fisheries Science Center, Woods Hole

Meeting participants

The participants in Woods Hole included: Chris Tholke, Liz Brooks, Greg DeCelles, Brian Linton, Jim Weinberg. George Maynard, Eric Hesse, Chris McGuire, Bill Duffy, Gary Shepherd, Steve Cadrin, Mike Simpkins, Mark Terceiro, Tara Trinko, Toni Chute, Anna Birkenbach, Brian Stock, Pierce McDonnell, Charles Perretti, Larry Alade, Lisa Hendrickson, Susan Wigley, Paul Nitschke, Tim Miller, Mike Palmer, Ariele Baker, Katherine Sosebee, Dan Hennen, Liz Sullivan, Emily Keiley, Jamie Cournane, and Michele Traver.

Remote participants via webinar included: Alex Hansell, Andrew Jones, Tony Wood, Chuck Adams, Daniel Caless, Karen Greene, Nichole Rossi, Rebecca Peters, William Gerencer, Melissa Sanderson, Steve Cadrin, Page Valentine, Peter Shelley, Tom Nies, Chris Kellogg, Robin Frede, Jessica Blaylock, Brett Alger, Alicia Miller, KB McArdle, Libby Etrie, Rich Bell, Raymond Kane, Jui-Han Chang, Gavin Fay, Mike Armstrong, and George Maynard

Meeting details

This meeting included implementation of the newly approved NRCC stock assessment guidance document. Three background documents were provided to the Panel: (1) an updated prospectus for each stock; (2) an overview summary all the salient data and model information for each stock; and (3) the NRCC Guidance memo on the Operational Assessments. The NRCC guidance memo was recognized as particularly relevant during the deliberations of the AOP. Prior to the meeting, each assessment lead prepared a plan for their assessments. The reports were consistent across species and reflected both the past assessment and initial investigations.

The meeting was held in the Meigs Room of the Marine Biological Laboratorys Swope Center in Woods Hole. The meeting began at 9:00 am. Approximately 31 people participated in Woods Hole and another 27 individuals participated via teleconference and Webinar.

The lead scientist for each stock gave a presentation on the data to be used, model specifications, evaluation of model performance, the process for updating the biological reference points, the basis for catch projections, and an alternate assessment approach if their analytic assessment was rejected by the peer review panel. In some cases, stocks were already being assessed using an "index-based" or "empirical" approach.

Common issues for multiple species

Population Dynamics staff reported on four common issues associated with multiple stock assessments: treatment of misreported catch, incorporation of survey catchability estimates from Northeast Trawl Advisory Panel conducted experiments, revised recreational catch estimates from MRIP, and incorporation of data collected using electronic monitoring.

Chris Legault presented background information on misreported catch, noting that while estimates are available by species, more detailed information such as statistical or stock area, market category and dates are not currently available. This precludes "correction" of misreported catch in stock assessments for the 2019 assessments. To accurately incorporate misreported catch, trip level detailed information would need to be corrected in catch databases. The timeline for correcting catch databases is unknown at this time.

Tim Miller, a member of the Northeast Trawl Advisory Panel, presented information on comparison studies between survey trawls equipped with "chain sweeps" (assumed 100% efficiency) and rock hopper sweeps (gear used on NEFSC multispecies bottom trawl surveys). These experiments produced estimates of relative efficiency for targeted flatfish and other demersal species. Estimated sweep efficiencies were used to scale up survey abundance indices to swept area abundance and biomass estimates. For stocks where there were sufficient data, data are available for incorporation into current stock assessments and can be used for comparisons with assessment results.

Mark Terceiro presented information on incorporation of revised recreational catch information generated by the Marine Recreational Information Program (MRIP). The assessments in this review that will be affected by revised recreational catch estimates include cod, haddock and pollock. The Georges Bank winter flounder stock has an insignificant recreational catch and the two coastal stocks of winter flounder will not be assessed in the management track until 2020.

Paul Nitschke presented information on the prospects of incorporating Electronic Monitoring (EM) data into the 2019 management track assessments. Three Exempted Fishing Permits (EFPs) were issued to permit electronic monitoring beginning in calendar year quarter 4 of 2017 and extending through 2018. While these permits were issued for quota monitoring purposes, there are data being collected that may have value in terms of informing stock assessments. Considerable discussion occurred relative to incorporation of EM data including informative comments by both EM project principal investigators and participating industry members. These stakeholders advocated for the use of high quality EM data, which was viewed as being more accurate than self-reported VTR data.

The proportion of the catch in 2017 and 2018 that was monitored through EM is generally less than 5% for stocks being assessed in 2019. For these assessments, lead assessment biologists plan to remove total retention EM data from their analysis due to known differences in selectively that could bias results in terms of characterizing the overall fishery. While the AOP recognized the potential value of EM collected data for future assessments, it also supported the caution expressed by stock assessment leads in terms of more fully developing the data structures and databases required to incorporate these data.

Major recommendations for review of individual stocks

In general, the AOP approved the plans presented, but recommended several revisions to the review levels as summarized below:

Stock	Lead	Major Recommendations
American Plaice	Larry Alade	Level 2 - Expedited Review. Plan B - absolute biomass estimates based on catchability, derive exploitation rates.
GOM Haddock	Charles Perretti	Level 2 - Expedited Review. Plan B - Loess Smoothing
GB Haddock	Liz Brooks	Level 2 - Expedited Review. Plan B - Loess Smoothing
CC/GOM Yellowtail Flounder	Larry Alade	Level 2 - Expedited Review. Plan B - absolute biomass estimates based on catchability, derive exploitation rates.
SNE/MA Yellowtail Flounder	Larry Alade	Level 2 - Expedited Review. Plan B - absolute biomass estimates based on catchability, derive exploitation rates.
GB Cod	Chris Legault	Level 1 - Direct Delivery. Current Assessment is a Plan B. Alternate approach, recommend status quo management.
GOM Cod	Charles Perretti	Level 3 - Expanded Review. Plan B - Loess Smoothing
Northern Windowpane Flounder	Toni Chute	Level 2 - Expedited Review.
Southern Windowpane Flounder	Toni Chute	Level 1 - Direct Delivery.
Witch Flounder	Susan Wigley	Level 1 - Direct Delivery. Current Assessment is a Plan B. Alternate approach, recommend status quo management.
Halibut	Dan Hennen	Level 1 - Direct Delivery. Current assessment is a Plan B. Alternate approach not required.
GB Winter Flounder	Lisa Hendrickson	Level 2 - Expedited Review. Plan B - Loess Smoothing
Pollock	Brian Linton	Level 2 - Expedited Review. Plan B - Loess Smoothing
White Hake	Kathy Sosebee	Level 2 - Expedited Review. Plan B - Loess Smoothing

Individual Stock Discussion Summaries

American Plaice The AOP inquired about the impact of missing survey coverage and determined that the impact was likely minimal. The AOP questioned whether both the NEFSC spring and bottom trawls surveys would be used in the Plan B approach and the lead assessment biologist indicated that an average of the two surveys would be used. The assessment plan will exclude the inshore MADMF survey from the model because of diagnostic issues, as was done in the 2017 update. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (empirical biomass estimates).

Gulf of Maine (GOM) Haddock The AOP debated whether the inclusion of revised recreational catch estimates merited recommending a Level 3 review for this assessment. The discussion revealed that the majority of recreational catch is likely generated by the party/charter boat sector, where catch estimate revisions are less significant. It was also noted that missing survey coverage was likely to have a minimal impact because key missing survey strata (offshore strata 30) are not included in the assessment. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (Loess Smoothing).

Georges Bank (GB) Haddock The assessment lead recommended a Level 2 review based on an increasing retrospective pattern and trends in weights at age and selectivity at age due to large year classes which impact projections. The AOP questioned the planned length of the projections and it was noted that the New England Fishery Management Council was requesting 3 year projections. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supports the proposed Plan B approach (Loess Smoothing).

Cape Cod (CC)/GOM Yellowtail Flounder The AOP noted reservations about hindcast recruitment estimates and recommended that sensitivity analyses might be conducted including or excluding these estimates. The AOP noted that the recommended Plan B approach (empirical biomass estimates) was recommended because there were catchability estimates available for the NEFSC surveys. The AOP discussed whether it should be receive a Level 1 vs. a Level 2 review given the history of a retrospective patterns for this assessment. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (empirical biomass estimates).

Southern New England (SNE)/Mid-Atlantic (MA) Yellowtail Flounder The lead assessment biologist recommended as Level 2 review for this assessment, based on accounting for some data corrections, natural mortality changes, the possibility of new selectivity blocks and the prospects for retrospective adjustments. The AOP asked for a clarification of the changes in M and debated whether this constituted a Level 2 or Level 3 change. It was noted that the changes would be empirical including updating the data that drive the calculation and not introducing a new method. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (empirical biomass estimates).

Gulf of Maine (GOM) Cod The AOP panel debated the relative importance of changes in the recreational catch estimates, given the severe restrictions on recreational catch in recent years. It was noted that cod is an important bycatch species in the haddock recreational fishery and impacts on discard mortality could be significant and that older recreational catch data were also re-estimated. Given questions about the impact of the revised recreational catch estimates and the high level of stakeholder interest for this stock, the AOP concluded that the review for this stock should be Level 3 (expanded review). The AOP supported the proposed Plan B for this stock (Loess Smoothing).

Georges Bank (GB) Cod The lead assessment biologist recommended a Level 1 review for this stock based on the already employed Plan B approach and the fact that revised recreational catch estimates would have little or no impact on the approach. It was noted that the recreational catch from this stock is highly variable. The AOP noted that while there is considerable management interest in this assessment, there is little to be gained through a more formal review. The AOP concluded that the review for this stock should be Level 1 (direct delivery).

Windowpane Flounder (Northern and Southern stocks) The AOP had an extended discussion about review recommendations for these assessments, noting the relatively poor model fit for the northern stock. The AOP concluded that the review for the northern stock should be Level 2 (expedited review) and should be Level 1 (direct delivery) for the southern stock.

Witch Flounder The lead assessment biologist recommended a Level 1 review based on the simplicity of the Plan B method and the fact that the impact of missing survey coverage was minor. The AOP asked about the survey catchability estimates and it was noted that the new estimates will consider both length and day/night factors. The AOP asked if a new catchability coefficients are used, how would this affect the exploitation rate estimate. It was determined that the same 9-year average would be used simply updated with new values. The AOP concluded that the review for the northern stock should be Level 1 (direct review). The alternative plan is for constant management.

Halibut The AOP noted that there were no significant changes and that the assessment only involved updated inputs. One AOP member recused himself from the discussion since he had developed the assessment method. The AOP concluded that the review for this stock should be Level 1 (direct delivery). There is no Plan B approach for this stock.

GB Winter Flounder The lead assessment biologist recommended a Level 2 review based on planned updates to the discard time series and retrospective patterns that may require rho-adjustments. The AOP inquired about the prospects for incorporating survey catchability estimates. It was noted that these were not used in the last update because of significant differences in growth rates vs. the other winter flounder stocks, but would be investigated again. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (Loess Smoothing).

Pollock The lead assessment biologist recommended a Level 2 review based on revised recreational catch estimates and the potential for new selectivity time blocks in the model. The AOP questioned changing the functional form of selectivity and it was noted that there were two models, one with dome shaped selectivity and the other without. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (Loess Smoothing).

White Hake The lead assessment biologist recommended a Level 2 review based on that there are no changes to the model, only updated data. The AOP had questions concerning the application of age-length keys and why pooled age-length keys were required. The AOP noted that there were some questions about the construction of the catch at age and additional uncertainty about the impact of missing survey coverage on survey indices. The AOP concluded that the review for this stock should be Level 2 (expedited review) and supported the proposed Plan B approach (Loess Smoothing).

AOP process discussion and summary

The AOP discussed its application of the new stock assessment process. It was noted that the distinction between Level 2 and Level 3 reviews was less clear, but that this really represented guidance to the peer review process in terms of time allocation and level of detail for a given stock. The AOP appreciated the input provided by audience members including industry stakeholders and the one page stylized summaries provided for each stock.

In summary, the meeting was productive and an effective implementation of the new assessment planning document. The meeting concluded at 3:30 pm. The peer review panel will meet from September 9-13, 2019 to complete their review.

16.3 Description of New England and Mid-Atlantic Region Stock Assessment Process

Overview

The Northeast Regional Coordinating Council (NRCC) developed the enhanced stock assessment process described here with the goals of (a) improving the quality of assessments, (b) allowing more improvement to occur within the routine assessment process, and (c) providing more strategic and longer-term planning for research and workload management. The process described here lays out two tracks of assessment work: a management track that includes the more routine assessments but with more flexibility to make improvements than in the past, and a research track that allows comprehensive research and development of improved assessments on a stock-by-stock or topical basis. The process provides clear opportunities for input and engagement from stakeholders and research partners, and the process also provides a longer term planning horizon to carry out research to improve assessments on both tracks, but particularly the research track. A key aspect of this process is the NRCC's development and negotiation of long-term management track cycles for each stock (i.e., how often each stock is assessed and in what years) as well as a five-year research track schedule, which will be updated through time by the NRCC.

Roles and Responsibilities

Northeast Regional Coordinating Council The Northeast Regional Coordinating Council (NRCC) consists of members from the Atlantic States Marine Fisheries Commission (ASMFC), Greater Atlantic Regional Fisheries Office (GARFO), Mid-Atlantic Fishery Management Council (MAFMC), New England Fishery Management Council (NEFMC), and Northeast Fisheries Science Center (NEFSC). The NRCC fulfills several functions, and, in the context of stock assessments, the NRCC's primary roles and responsibilities focus on setting priorities and scheduling of assessments. With respect to assessment priorities, the NRCC (a) sets long-term (five-plus year) schedules for both the management and research track, (b) reviews and adjusts those schedules as needed, and (c) recommends priorities among complex management track assessments (i.e., assessments requiring expedited or enhanced peer reviews) in situations where more complex assessments are proposed than can be accommodated. Designated deputies from each NRCC member organization form the "NRCC Deputies" panel, which reviews and approves research track stock assessment working groups as well as external experts nominated to serve on management track or research track peer review panels.

Assessment Oversight Panel The Assessment Oversight Panel (AOP) consists of four members (a) the Chief of the Populations Dynamics Branch, NEFSC, or his/her designee, who serves as Chair of the AOP, (b) the Chair of the NEFMC SSC, or his/her designee, (c) the Chair of the MAFMC SSC, or his/her designee, and (d) the Chair of the ASMFC Assessment Science Committee, or his/her designee.

The primary responsibilities of the AOP are to (a) review and approve management track assessment plans in the context of guidelines for permissible changes under each level of management track peer review, (b), in the near term if they have not yet been developed and reviewed in a prior

assessment peer review, review and approve plans for any alternative (i.e., "Plan B") approach to be used if the peer review finds primary management track assessment is not suitable for providing management advice, (c) review and approve revisions to management track assessment plans developed in response to new data or based on advice from the AOP generated from review of the original plan, noting that any changes that would require upgrading or downgrading the assessment tier would require NRCC consultation; and (d) provide a summary report to the NRCC on an annual basis of AOP actions taken.

Assessment Oversight Panel meetings are open to the public. Council, Commission, and GARFO staff are welcome to participate, and those staff with lead responsibilities for stocks under consideration will be requested to serve as invited participants. At least one staff representative should participate from GARFO and each Council and Commission with stocks under consideration.

Northeast Fisheries Science Center Fish stock assessment scientists from the NEFSC support both management and research track assessments. NEFSC assessment scientists have primary responsibility for planning and carrying out management track assessments for all federally-managed stocks, as those assessments are conducted on a routine basis and require consistent capacity and expertise. As part of the management track process for stocks with NEFSC lead responsibility, NEFSC assessment scientists develop initial plans for assessments and alternatives (i.e., "Plan B") in advance of upcoming assessments and revise those plans if necessary in response to new data; where possible, alternative approaches should be developed in advance in prior research track assessments. NEFSC assessment scientists provide initial management track assessment plans for review by the AOP, which in turn reviews and provides recommendations to the NRCC. In unusual situations where more assessments are proposed for expedited and enhanced peer review than can be accomplished in the time available for peer review, then the NEFSC consults with the NRCC to determine which assessments to "downgrade" to a lower assessment level and peer review. NEFSC assessment scientists, as well as other NEFSC scientists and other federal, state, academic and other non-governmental scientists participate in research track assessments.

Atlantic States Marine Fisheries Commission ASMFC Technical Committee and Assessment Science Committee members may support both management and research track assessments. The ASMFC has primary responsibility for planning and carrying out management track assessments for several state-managed stocks, several of which require substantial NEFSC staff engagement and are managed according to the assessment process described here. As part of the management track process for jointly managed stocks with ASMFC lead responsibility, the relevant ASMFC Technical Committee develops initial plans for assessments and alternatives (i.e., "Plan B") in advance of upcoming assessments and revises those plans if necessary in response to new data. The Technical Committees' initial management track assessment plans are reviewed and approved by the Assessment Science Committee, which then provides those assessment plans to the AOP for its review and subsequent recommendations to the NRCC. In unusual situations where more management track assessments are proposed for expedited and enhanced peer review than can be accomplished in the time available for peer review, then the ASMFC consults with the NRCC to determine which assessments to "downgrade" to a lower assessment level and peer review. For ASMFC managed stocks that are scheduled following the process described here, ASMFC may opt to follow the AOP

and management track peer review process, or use traditional ASMFC planning and review processes, though care must be taken to coordinate with the management track process to avoid any work or review conflicts. ASMFC Technical Committee members, as well as NEFSC scientists and other federal, state, and academic scientists participate in research track assessments.

Peer Review Panels Peer review panels are convened to review expedited (level 2) and enhanced (level 3) management track assessments and research track assessments. Peer review panels review the assessment(s) for technical merit and provide recommendations to the relevant Agency, Council(s), and or Commission on the whether the assessment should or should not be used for management. For management track assessments, the peer reviews will be conducted by a small panel of relevant SSC members with additional external experts if/as needed; reviewers will be nominated by the relevant Council(s) and/or Commission and confirmed by the NRCC Deputies. When nominating and confirming membership for management track peer reviews, consideration should be given to providing some continuity from one peer review to the next, to promote consistency in decisions across peer review panels. For research track assessments, peer reviews will likely, but not exclusively, be provided by the Center for Independent Experts (CIE). In some cases, it may be preferable to convene a research track peer review panel outside of the CIE process; in those cases, the relevant Council(s) and/or Commission will nominate panelists, which will be reviewed and confirmed by the NRCC Deputies. Consideration should be given to including SSC members in the peer review, including the possibility of having an SSC member chair the peer review; this approach has been helpful in the past to provide some continuity across the peer review and subsequent SSC review.

Scheduling Process

During 2016-2017, the NRCC developed a process for scoring and prioritizing stocks for both management and research track assessments, and the resulting information was used to inform the development of the initial management and research track schedules. The scoring and prioritization process built off of the process described in the National Marine Fisheries Service's "Prioritizing fish stock assessments". An NRCC working group evaluated the scoring process and factors recommended by the NMFS report, selected the factors that were most relevant to NRCC stock assessment scheduling, modified the factor descriptions and scoring rubrics, and added entirely new factors as needed. The working group then organized these factors into six categories: management needs, fishery importance, stock status and trend, ecosystem importance, assessment information, and stock biology. The resulting scoring factors are described in [insert scoring document as link or appendix]. Briefly, and generally speaking, NRCC working group members scored each stock within their jurisdiction for each factor[3], and then those scores were averaged across all members for each factor, averaged across all factors for each category, and then averaged across categories for each stock, resulting in one overall score for each stock. A different suite of factors was used to calculate the final score for management track vs research track assessment priorities, and a few factor or category scores were provided independent of the overall score because they were deemed particularly important for developing assessment schedules.

[3] NMFS working group members scored all stocks; GARFO scored factors related to management and regulations, and NEFSC scored factors related to science. The Councils and Commission scored their respective stocks.

With the resulting scores as information, the NRCC working group developed initial strawman schedules for both management and research tracks. Those strawman schedules, prioritization scores, and other information were used by the NRCC to develop an initial five-year schedule of research track assessments and an initial schedule of management track assessments, with each management track assessment assigned a starting year and a certain cycle or periodicity ranging from annual management track assessments to 6-year intervals between management track assessments. The resulting schedules were informed, but not driven, by the prioritization scores; final decisions regarding the schedules were made through NRCC negotiation.

In order to maintain a five-year research track schedule each year, as what had been the fifth year becomes the fourth year, the NRCC will consider the existing research track schedule, research track scores, and other information and identify which stocks or topics should be addressed in the new fifth year of the schedule. The NRCC will also consider any changes to the existing research or management track schedules as needed. In the absence of changes, the management track schedule will continue with the same periodicity for each stock.

The prioritization scores developed for both research and management tracks in 2016-2017 may degrade in terms of relevance over time. When the NRCC feels those scores are no longer relevant for informing scheduling discussions, the scoring process will be conducted again to provide fresh scores to inform the scheduling process. Because the scoring process is laborious, the NRCC anticipates refreshing the scores on an infrequent basis, perhaps once every 5-7 years.

Management Track Process

Management track assessments are designed to provide routine, scheduled, updated advice to directly inform management actions. Management track assessments are designed to be simpler, quicker, and more efficient than research track assessments. However, the management track provides some flexibility to allow assessments to improve over time by building off the previously accepted assessment, without requiring a research track assessment for every step along the way. The modifications allowed within the management track are intended to provide the analyst with the flexibility needed to improve the science and update a previously accepted assessment when issues arise or new data become available.

Management Track and Peer Review Levels The flexibility in management track assessments allows for different levels of complexity and extent of changes that can be applied when conducting a management track assessment. These different levels of complexity and extent of changes, in turn, call for different levels of peer review and public engagement. For consistency sake, the levels of peer review, extent of public engagement and changes allowed under each management track level are described below. Generic terms of reference for management track assessments are also provided below.

When developing the list of permissible changes, it was recognized that all possible changes that would warrant consideration could not be anticipated given the evolving nature of science and assessment methods. Consequently, the following lists represent specific changes that are permitted under each level but should not be considered exhaustive. If a change proposed by an analyst is not

detailed below, the AOP will determine whether the modification is permissible and which level of peer review would be required.

During and prior to the assessment planning stage, stakeholders will be able to provide input on all assessments. During the "input" phase of management track assessments (described below), NEFSC, ASMFC and NRCC partners will work together to engage with stakeholders, academic and state partners to solicit new data and ideas for any and all levels of upcoming management track and research track assessments. Additional stakeholder engagement would occur during the public comment periods of the AOP meeting (described below) where the assessment plans presented by NEFSC and ASMFC analysts will be reviewed. Opportunities for public engagement during assessment reviews are specific to the assessment level and are described below.

Data Updates In some cases, data updates may be requested by a Council or Commission between scheduled Management Track assessments. Data updates are just that, summaries of new data that have become available since the last Management Track assessment. Data updates do not involve rerunning any assessment model and in most cases do not provide a formal update of stock status. The NEFSC is actively working to automate much of the assessment data processing, with the goal of being able to provide standardized data updates through an automatic reporting system. Previously, some requested data updates were quite extensive and required data processing and manipulation that would be challenging to automate, and in some cases those requested data updates required as much work as what would be considered a Level 1 assessment in the current process. In addition to cases needing additional work beyond updating available data, cases where data must be acquired from sources outside of the NEFSC (e.g. state index datasets) may take additional efforts and may not be possible in a data update framework. If such extensive data examinations are requested in the future, they would need to be added to the Management Track schedule to account for the workload requirements. However, requests for standardized, automated data updates would not need to be added to the Management Track schedule because they could be provided at very low cost in terms of staff time. During the, hopefully short, timeframe while NEFSC develops the automated data update system, any data update requests will need to be negotiated through the NRCC.

Standardized, automated data updates are not formally considered as Management Track assessments and do not undergo any peer review, just normal quality assurance and control procedures. The intent of data updates is to provide reassurance that multi-year specifications set based on the most recent Management Track assessment are still appropriate, without requiring a new assessment. Such updates are most useful when they are formally accounted for within a fishery management plan with clear decision rules on what action should be taken if a data update implies a strong change in stock status. Without such decision rules, data updates may just highlight a concern that cannot be addressed without a formal management track assessment, which would require adding an assessment to the schedule on short notice, or waiting for the next scheduled assessment.

Level 1: Direct delivery A level 1 management track assessment is essentially a simple update of the previously approved assessment with new data. This level of assessment update will be delivered directly from the NEFSC to the appropriate Council or Commission technical body (e.g., SSC)

and will not undergo peer review beyond that conducted by those technical bodies. Furthermore, although there will be opportunities for public input on assessments in advance during the input phase described below, there will be limited opportunity for public engagement during the assessment review, which will occur during the public comment period of the technical body's meeting. Given the limited peer review and public engagement, only minor changes, such as those detailed below, are permissible.

- Model that has been updated with revised data, with minor changes (such as small adjustments to data weights, fixing parameters estimated at bounds, correcting minor errors in previous model)
- Incorporation of updated data from recent years in the estimation of biological information (growth, maturity, length-weight relationship)
- Evaluating effects of delayed seasonal surveys or missing strata on fishery-independent measures of abundance
- If adding or revising data reveals problems in model performance, analyst should identify concerns that may need further analyses and/or review
- Standard QA/QC procedures employed by the NEFSC

Level 2: Expedited review A level 2 management track assessment can involve a little more flexibility for deviations from the previously accepted assessment, but that flexibility is limited to allow for efficient peer review of multiple assessments in one peer review meeting, similar to what previously had been carried out for groundfish operational assessments for the NEFMC. Level 2 assessments will undergo a formal, but expedited (1-2 hour maximum), peer review by a small panel of SSC members from the relevant Council(s), along with additional external experts if desired, before submission to the appropriate Council or Commission technical body. In addition to opportunities for public input on assessments in advance, opportunities for public engagement will occur during the public comment periods of the public review meeting and the subsequent meeting of the Council or Commission technical body. Given the moderate level of peer review and engagement, level 2 assessments will generally use the same assessment structure and data as the previously accepted assessment, but some changes are permitted (detailed below) that warrant review by an external body. In this level, the cumulative impacts of the number of changes should also be considered; any individual change may be minor, but if there are several changes, the overall impact could be substantial and may warrant shifting an assessment to level 3 and providing enhanced peer review. Changes permitted in level 2 assessments include those noted in level 1, and:

- Updated discard mortality estimates, when based on peer-reviewed experimental evidence
- Evaluating effects of delayed seasonal surveys or missing strata on fishery independent measures of abundance if significant analysis is required to characterize the effects
- Recalibrated catch estimates (e.g., transition to Marine Recreational Information Program, area allocation tables, conversion factors (whole to gutted weight))

- Simple changes, corrections, or updates to selectivity, including but not limited to:
  - Changes to most recent selectivity stanza
  - Changes to historical selectivity stanza if they are corrections or reinterpretations of previously used block timeframes
- Retrospective adjustment to management metrics following established retrospective adjustment protocols Technically, when either the rho-adjusted SSB or F (point estimate / (1 + Mohn's rho)) falls outside the 90% confidence interval of the terminal year estimate, the retrospective adjustment is applied for both status determination and to the starting population for projections.
- Adjustment of method for estimating biological information (growth, maturation, sex ratio, changes to length-weight relationships, etc.), when based on methods developed with sufficient peer review or justification for its use
- Calculate new values for the existing BRPs

Level 3: Enhanced review A level 3 management track assessment will permit more extensive changes than a level 2 assessment and therefore requires a more extensive peer review (one-half to a one full day). The flexibility in level 3 provides an opportunity to make progress within the management track toward the Next Generation Assessments envisioned in the Stock Assessment Improvement Plan, by including more detailed spatial, temporal, environmental and species interactions within existing model frameworks. It is important to note, however, that full achievement of Next Generation Assessments will likely require research track efforts as well. As in level 2 assessments, public engagement opportunities will occur during the public comment periods of both the public review and the subsequent meeting of the Council or Commission technical body, as well as during the input phase of the assessment process as described below.

Level 3 assessments will be reviewed by a small panel of SSC members from the relevant Council(s) as well as additional external experts as needed; any external reviewers outside of the SSCs will be nominated by the Council or Commission and confirmed by the NRCC Deputies. Given the enhanced peer review, changes to most assessment elements, with the exception of stock structure, would be permitted in level 3 assessments; however, cumulative impacts should be considered when making a determination between the changes permissible within the "enhanced review" level and changes that would require switching to the research track process. Changes permitted in level 3 assessments include those noted in levels 1 and 2, and:

- Inclusion of new or alternate interpretations of existing indices
- Changes to estimation method of catchability, including but not limited to:
  - Empirical estimations
  - Changes in habitat/availability/distribution on catchability
  - Use of informed priors on catchability in a model
- Updating of priors based on new research if done on a previously approved model

- Recommend significant changes to biological reference points, including but not limited to:
  - Change in the recruitment stanza
  - Number of years to include for recent means in biological parameters
  - Suggestions of alternate reference points if based off a similar modeling approach (e.g. age-based, length-based, etc.)
- Updating of historical selectivity stanzas
- Changing recruitment option used, meaning using a stock-recruitment relationship, or cumulative distribution function, etc.
- Changes to selectivity functional form (i.e. such as a new selectivity model) if supported by substantial empirical evidence.
- Changes to fleet configuration
- Changes to natural mortality (M)
- New modeling framework, if the new framework was evaluated during a previous research track topic investigation, and the species in question was one of the examples evaluated. Through research track topics focused on methods, new models could be implemented in parallel with an accepted model and provide a basis for eventual shift to a new model through a level 3 management track assessment. This would allow model evolution, technical innovations, and testing without the penalty of forgoing research on stock dynamics until a new Research Track process is scheduled.

Management Track Assessment Terms of Reference Generic Terms of Reference (TORs) for assessment updates that will be used directly for management (Management Track assessments) are provided below. They include the TORs necessary for updating the necessary input data (catch and survey), assessment model, biological reference points and short-term projections but do not include the research-oriented TORs that are included in Research Track assessments.

1. Estimate catch from all sources including landings and discards.
2. Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).
3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) as possible (depending on the assessment method) for the time series using the approved assessment method and estimate their uncertainty. Include retrospective analyses if possible (both historical and within-model) to allow a comparison with previous assessment results and projections, and to examine model fit.
   (a) Include bridge runs to sequentially document each change from the previously accepted model to the updated model proposed for this peer review.
   (b) Prepare a "Plan B" assessment that would serve as an alternate approach to providing scientific advice to management if the analytical assessment were to not pass review

4. Re-estimate or update the BRP's as defined by the management track level and recommend stock status.
5. Conduct short-term stock projections when appropriate.
6. Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.

Management Track Process and Logistics

Step 1: Input Throughout the year data come in and new ideas are generated. As part of the new management track assessment process, the NEFSC and ASMFC will work with NRCC partners and others to engage with stakeholders, academic and state partners to solicit new data and ideas. This engagement strategy will involve ongoing, regular two-way communications with stakeholders and partners using a variety of approaches, which could include, but not be limited to, social media and web interactions as well as face-to-face stakeholder engagement meetings convened by NRCC members or hosted by stakeholder groups. The engagement strategy will adapt as needed to improve two-way communications, but at a minimum will involve biannual engagement efforts to provide updates on the most recent management and research track assessments and to seek input on upcoming assessments. This engagement will solicit input on all levels and types of assessments, but will particularly focus on research track assessments where there are not only more opportunities for change and improvement but also opportunities for joint research planning and direct collaborative research efforts with stakeholders and partners, which the NRCC is particularly interested in fostering. All input received will be provided to the assessment leads to support development of their assessment plan. Six months or more in advance of a scheduled management track assessment, the NEFSC or ASMFC assessment lead for the stock compiles available input and does initial exploratory work to determine how complex the next management track assessment should be in terms of new data streams or model changes incorporated.

Step 2: Assessment planning Following data input and exploration, and based on the explicit management track guidelines, the assessment lead proposes to the AOP the extent of assessment changes to be explored and the associated level of peer review. The assessment lead also provides proposals for assessment complexity under lower levels of peer review, to provide options for consideration. In the case of ASMFC led stock assessments, this initial proposal is developed by the relevant Technical Committee and reviewed by the Assessment Science Committee before being proposed to the AOP. The resulting assessment plans should indicate what input was considered and how it will be addressed, included or excluded, in the assessment; this provides the explicit connection between public or other input and the assessment plan.

Step 3: AOP and NRCC review After data have arrived and exploration has occurred, the AOP is convened to provide technical review of the proposed management track assessment plans for the upcoming year. For any assessment proposed for level 2 or 3 peer review, the AOP considers the changes suggested (and "Plan B" if not previously vetted by a research track or prior management track assessment) and approves those changes (and Plan B) and applies the peer review level

guidelines to confirm the level of peer review for the most complex proposed version of assessment (i.e., levels 2-3 above).

At the completion of the AOP review, the NEFSC, which manages the logistics of the peer review process, reviews the AOP approved suite of assessments to ensure that the peer review logistics are feasible. In unusual situations where more assessments are proposed for expedited and enhanced peer review than can be accomplished in the time available for peer review, the NEFSC consults with the NRCC to determine which assessments to "downgrade" to a lower assessment level and peer review. The resulting recommendations from the AOP, modified if needed and approved by the NRCC, are then implemented by the NEFSC and ASMFC assessment leads.

Step 4: Assessment conducted This step may include several phases. First, each assessment lead evaluates any new data that have arrived since they developed the original proposal for assessment complexity and level (see step 2). If any changes to the approved assessment plan are needed in response to new data, the assessment lead proposes those revisions. If those proposed revisions could result in changes in the peer review level, then the AOP provides technical review and applies the management track peer review guidelines to determine the appropriate level of peer review, likely via conference call or virtual meeting. In unusual cases where such changes could result in substantive changes to the overall suite of planned peer reviews, the NRCC would be consulted with respect to priorities. The assessment leads then carry out the management track assessment within the scope of the approved assessment plan for each stock.

Step 5: Peer review Expedited and enhanced (levels 2 and 3, see above peer review levels) management track peer reviews are scheduled and convened, as described below, seeking to combine peer reviews as appropriate for efficiency and to optimize the ability to provide timely peer reviewed results to as many fishery management action processes as feasible. Outputs of peer reviews are provided as expeditiously as possible to the appropriate Council or Commission technical bodies and then to the Councils and/or Commission to inform management action (Step 6 in the management track process flow chart). These outputs will be provided in the form of summary reports and will address the assessment terms of reference (see above). For the usual situation where multiple management track assessments are reviewed at one time, the summary reports would likely be compiled as chapters in one overall summary report, and the peer review comments and recommendations would likely be incorporated within each chapter. In all cases, associated data and analytical details will be accessible. Early in the implementation of this process, the NRCC will develop and approve standard report templates for each level of management track assessment (and data updates).

General Timing of Management Track Process Two management track peer reviews for level 2 and 3 assessments will be conducted each year to accommodate the variation in fishing year among stocks and minimize the time lag between the final year of the assessment model and the subsequent implementation of new specifications. Each peer review could include both level 2 and level 3 assessments, and the peer review panel would be composed appropriately with SSC members from the relevant Council(s) and any additional experts as needed. For the majority of stocks, the fishing year starts at the beginning of January or May. Consequently, a peer review will be conducted during the beginning of September for those stocks with fishing years around May 1 and another peer

review will be held at the end of June to accommodate stocks with fishing years beginning around January 1 (see table below). This timing is designed to ensure that products from the assessment review can be provided in time to meet the associated management timelines. Assessment models examined during the September peer review will incorporate data through the end of the previous year. For the suite of stocks that undergo peer review in June, it will be difficult to incorporate fishery catches through the end of the previous year due to timing constraints of data availability; it is likely that assumptions may need to be made for the terminal year catch. Assessment reviews for transboundary stocks carried out under the auspices of the Transboundary Resources Assessment Committee will continue to be scheduled based on bilateral negotiation.

Level 1 management track assessments will be delivered directly to the appropriate Council or Commission technical body and are not evaluated as part of the two peer reviews. If desirable, some level 1 assessments can be prepared and delivered throughout the year according to the Councils' and Commission's current delivery schedules. If, upon incorporating the most recent year of data, a level 1 assessment needs to be upgraded to a higher level that requires peer review, delivery of the assessment will be delayed until the next peer review, typically resulting in a delay of weeks to a few months. In such situations, the relevant Council or Commission would be consulted to discuss the needed changes and the resulting delay. In some situations, changes may be required to provide valid scientific advice to management. In others, the changes may be needed to provide improvements to the quality of the advice, in which cases the relevant Council or Commission may prefer to maintain the original delivery timeline while sacrificing the improvement. Furthermore, as the management track schedule comes into effect and workloads, timing, and demands shift, one way to enhance the efficiency of the process may be to simplify the delivery system to have most or all level 1 assessments coincide with the timing of the peer reviews, eliminating the need for some additional consultation and sacrifices.

Species or FMP	Beginning of Fishing Year	Management track peer review
Tilefish	November 1	End of June
Northern Shrimp	December 1	End of June
Bluefish	January 1	End of June
Mackerel/Squid/Butterfish	January 1	End of June
Fluke/Scup/Black sea bass	January 1	End of June
Surf clam / Ocean quahog	January 1	End of June
Atlantic herring	January 1	End of June
Striped bass	January 1	End of June
River herring / Shad	January 1	End of June
Red crab	March 1	End of June
Scallop	April 1	Beginning of September
Spiny dogfish	May 1	Beginning of September
Monkfish	May 1	Beginning of September
Groundfish (NE multispecies)	May 1	Beginning of September
Hakes (Small mesh multispecies)	May 1	Beginning of September
Skates	May 1	Beginning of September
American Lobster	July 1	Beginning of September

Fishing year and peer review dates for each species or fishery management plan (FMP)

Research Track Process

Research Track Assessments and Topics Research track assessments and topics are complex scientific efforts focused either on (a) assessments of individual stocks with comprehensive evaluation of new data streams and model changes or (b) research topics that apply to assessments of several stocks. Generally speaking, applied scientific efforts in the fish stock assessment arena lie along a continuum from "research" to "research track" to "management track," with each step informing the next and getting closer to directly informing management decisions. Generic "research" may be designed to inform the research track, but typically is not designed to directly inform the management track. Research track efforts, on the other hand, are designed to directly inform future management track assessments, but may not immediately inform management decisions. Research track efforts can inform management track assessments by, among other things, (a) direct examination and development of an assessment or (b) tackling analytical, data, or other issues facing multiple assessments.

Research Track Process and Logistics

Step 1: Research Topic and Assessment Development Initial research track topics and assessments are developed and proposed to the NRCC via individual NRCC members. These proposals can derive from ideas or recommendations proposed to or developed by Councils or Commission, through ideas or proposals developed by NEFSC or ASMFC scientists, or through ideas or proposals submitted through the NEFSC or GARFO. NRCC member organizations will work together to develop effective stakeholder engagement processes to solicit ideas (see Management Track Step 1 above for more on input), which in turn could develop into research assessment or topics that would be proposed by one or more NRCC members. These proposals are then evaluated through the scheduling process described above.

Step 2: Working group(s) Once a research track assessment or topic is scheduled, NEFSC and/or ASMFC assessment lead(s) are assigned and reach out to stakeholders, academics, and NRCC and management partners, etc., and consult existing sets of research recommendations (e.g., from past assessments or Council or Commission research priorities) to identify research needs to inform a given research track effort. This outreach effort could include formation of a working group or steering committee to carry out the outreach, or that working group or steering committee could be formed after the initial outreach and focus primarily on developing the plan for the research track effort.

Given the potential long-term nature of research track efforts, in some cases a steering committee to guide work may be established initially. The purpose of such a steering committee would be to identify research needs and provide guidance for the research that is undertaken, to ensure that the eventual research outputs are useful and able to be considered within the eventual research track assessment or topic. Given that purpose, members of a steering committee should be recognized experts in fields of study relevant to the priority research needs for a given research track assessment or topic; this could include federal, state, and academic scientists as well as industry or non-governmental experts engaged in developing or guiding cooperative research studies. Membership

of a steering committee could be somewhat dynamic and change through time for longer term research track efforts, as research progresses and different expertise is needed to provide research guidance. Steering committee members would be nominated by NRCC members as well as solicited through public outreach; steering committee membership would be reviewed and confirmed by the NRCC Deputies, with a focus on ensuring that all members have significant, relevant expertise. Care should be taken to avoid any perceived or real conflicts of interest, for example if steering committee members advocate for research that would be conducted by their host institution. A steering committee chair would be nominated and approved by the NRCC Deputies from the suite of steering committee members, and that chair would guide the overall work of the steering committee and seek to avoid conflicts of interest.

For stock-specific research track assessments, a formal stock assessment working group will likely be convened in addition to, or instead of, a broader steering committee. Those working groups would be formed following the process established for past Stock Assessment Workshop working group protocols.

Research track working groups, both topical and stock-specific, will be tasked with developing and implementing the research plan and terms of reference based on scoping. The research plan should indicate which outputs will be applied, and how, to future management track assessments and/or management actions. This is most critical for research topics, where the terms of reference at the start should clearly indicate what outputs will inform future management track assessments, and how they would do so. For stock specific research track assessments, consideration should generally be given to development of alternative approaches to providing management advice if a research track or future management track assessment should be deemed unsuitable for use in management, i.e., development of "plan B" assessment advice approaches. In most, if not all cases, such "plan B" approaches would be evaluated by the peer review panel after the panel completed its review of the research track assessment; "plan B" approaches should be considered as backup plans for any future problems with an assessment, not an alternative to the developed research track assessment, unless of course that assessment is rejected for use in management advice. In situations where a "plan B" approach has been developed and approved through a research track peer review, the expectations are that approach would be applied in future management track assessments as a backup, and the AOP would not need to repeat the review and approval of that "plan B" approach.

Step 3: Research Once the research plan and terms of reference are established, the steering committee and/or working group guides and/or carries out the necessary research and compiles the results to inform the research track effort, incorporating public planning, data, and analytical meetings as appropriate. In some cases, funding, staff, or other resources may limit research efforts, and, in those cases, the steering committee or working group should set priorities and ensure the most critical research is accomplished. When resources are limiting, the steering committee or working group should also inform the NRCC, whose members may be able to seek out additional resources to support the required work.

In order to promote an effective and innovative research track, topics and stock-specific assessments in this track typically will be carried out over longer time frames and with fewer requirements for using the most recent data, etc. In the two-track approach, the research track is intended to be the opportunity for extensive and comprehensive research and analysis, so it is helpful to remove timing

constraints as much as possible. This is different than the management track, which is very much driven by the need to meet specific management timelines and apply the most recent data feasible. As appropriate and feasible, the research and management track schedules will be designed to have management track assessments for specific stocks immediately follow research track assessments for those stocks, which allows for the comprehensive and innovative research to occur with less limitations but ensures immediate application of the research results with the inclusion of the most recent data in a management track assessment.

Step 4: Comprehensive peer review Research track peer reviews are considered to be "comprehensive" peer reviews, in contrast to the expedited and enhanced peer reviews carried out for management track assessments. These reviews generally require 1.5-4 days and are intended to consider all aspects of the research topic or stock-specific assessment and provide advice on the validity of the research and analyses conducted as well as provide recommendations as to whether the outputs are suitable for use in future management track assessments and/or to inform future management actions. Typically, but not exclusively, peer review panels would be provided through the Center for Independent Experts (CIE) and would include at least one relevant SSC member to provide continuity with later Council, Commission, and SSC reviews and actions. As mentioned previously, in some cases it may be preferable to convene a research track peer review panel outside of the CIE process; in those cases, the relevant SSCs, NEFSC, and/or ASMFC Assessment Science Committee will nominate panelists, which will be reviewed and confirmed by the NRCC Deputies.

Outputs of research track peer reviews are provided as expeditiously as possible to the NEFSC and/or ASMFC Assessment Science Committee for use in future management track assessments. These outputs will be provided in the form of an assessment summary report, a peer review report, and a comprehensive assessment document that covers the full suite of work carried out. The peer review report could either be one panel report, or a compilation of individual peer review reports along with a summary panel report. Working group papers, associated data, and background materials will be accessible if needed. If immediate management action is required based on the outcomes of a research track assessment, the outputs also will be provided to the appropriate Council or Commission technical bodies and then to the Councils and/or Commission to inform management action.

Step 5: Translate to Management In many cases, research track outputs will be incorporated into future management track assessments, as indicated in the relevant initial research plan. In some cases, research track outputs may also be used to directly inform immediate management actions. This would typically occur when research track outcomes indicate important or urgent changes in stock status that require immediate attention; otherwise, the expectation is that it usually will be more appropriate to take the research track outcomes and apply those with updated data in the next scheduled management track assessment to inform future management action.

16.4 Onsite participants

2019 Groundfish Operational Assessments Attendee List

Name	Affiliation
Alicia Miller	NEFSC
Andy Beet	NEFSC
Ariele Baker	NEFSC
Benjamen Frey	NEFSC
Brian Linton	NEFSC
Brian Stock	NEFSC
Brooke Wright	SMAST
Cate O'Keefe	MADMF
Charles Adams	NEFSC
Charles Perretti	NEFSC
Chris Legault	NEFMC SSC Reviewer
Dan Caless	GARFO
Dan Hennen	NEFSC
Debra Duarte	NEFSC
Emily Keiley	GARFO
Fred Serchuk	NEFMC SSC
Gareth Lawson	Conservation Law Foundation
Gary Shepherd	NEFSC
Greg Decellis	MADMF
Greg Power	GARFO
Jackie O'Dell	Northeast Seafood Coalition
Jamie Cournane	NEFMC Staff
Janelle Morano	Cornell University
Jessica Blaylock	NEFSC
Jim Weinberg	NEFSC
Kathy Sosebee	NEFSC
Larry Alade	NEFSC
Libby Etrie	NEFMC
Lisa Hendrickson	NEFSC
Liz Brooks	NEFSC
Liz Sullivan	GARFO
Maggie Raymond	Associated Fisheries of Maine
Mark Terceiro	NEFSC
Matt Cieri	MEDMR
Melanie Griffin	MADMF
Michael Palmer	NEFSC
Michele Traver	NEFSC
Pat Sullivan	NEFMC SSC Reviewer
Paul Nitchske	NEFSC
Rebecca Peters	Maine DMR
Rich McBride	NEFSC
Russell Brown	NEFSC
Scott Steinback	NEFSC
Shelly Dawicki	NEFSC
Steve Cadrin	Reviewer; U. Mass
Susan Wigley	NEFSC
Tim Miller	NEFSC
Tyler Pavlowich	NEFSC