#7a

2020 Management Track Peer Review Panel Report [DRAFT]

Jean-Jacques Maguire[1] (chair), Richard Merrick[2], Patrick Sullivan[3] and Cate O'Keefe[4]

[1]Halieutikos inc., [2]NOAA Fisheries Service (retired), [3]Cornell University, [4]Fishery Applications Consulting Team, LLC

Executive Summary:

Thirteen groundfish and one scallop stock assessments were scheduled to be reviewed in the Autumn 2020 Management Track process. The Assessment Oversight Panel (AOP) reviewed the assessment plans and recommended that three assessments be direct delivery (Level 1): Ocean pout, Atlantic halibut and Northern silver hake. Of the remaining eleven, six were expedited reviews (Level 2) and five were enhanced reviews (Level 3). The eleven assessments with expedited or enhanced peer review included in this report are: 1) Atlantic wolffish, 2) Acadian redfish, 3) Atlantic sea scallops, 4) Northern window pane flounder, 5) Southern window pane flounder, 6) Georges Bank winter flounder, 7) Gulf of Maine winter flounder, 8) Southern New England Mid-Atlantic winter flounder, 9) Northern red hake, 10) Southern red hake, 11) Southern silver hake / Offshore hake.

Peer Review Panel Report:

The Peer Review Panel (PRP) for the September 2020 Management Track Assessments met via webinar on September 14 - 18, 2020. Attendance at the meeting is provided in Appendix A. The assessments were prepared under guidelines provided by the 2020 Assessment Oversight Panel (AOP). These guidelines provide a pathway for continuing development of previously accepted assessments for each species including incorporation of the most recent data and understanding of biology of the species being assessed.

We thank Russ Brown (Population Dynamics Branch Chief) and Michele Traver (Assessment Process Lead) for their support during the meeting. We thank the staff of the Population Dynamics Branch at NEFSC for the open and collaborative spirit with which they engaged the PRP. Our thanks extend not only to the analysts for each assessment, but also to the rapporteurs for taking extensive notes during the meeting and to staff of the New England Fishery Management Council, Atlantic States Marine Fisheries Commission, and NOAA Fisheries/Greater Atlantic Regional Fisheries Office who provide context and additional background. We also thank the other participants for helping make the meeting productive and collegial. Finally, the PRP thanks the staff at NEFSC for supporting the logistics during the meeting.

The PRP has suggestions for improvements that could be made for the next Management Track Assessments. With respect to information needs:

1. It is very helpful to have all background documents, information, and presentations available prior to the beginning of a stock's review. This should include the full AOP report and summary, documentation of the current assessment, documentation of the preceding assessment

(including peer review reports and relevant SSC reports), the most recent benchmark research track assessment (if different from the preceding), a table of the stock's status and reference points, and at least a draft version of the Powerpoint presentations.

2. It would be useful if changes between the previous method(s) and the currently proposed method were documented in assessment summary reports. For example, the northern windowpane report did not document updated AIM model output, and the red hake reports did not document the results of the Red Hake Stock Structure Workshop (a "Research Track" exercise).
3. Assessment update reports should match the requirements laid out in the Management Track Assessment Terms of Reference. For example, the analyst should list and respond to any review panel or SSC concerns relevant to the most recent prior assessments.

With respect to process:

1. The Panel should be provided with a clear summary of what each Management Track review level allows.
2. The implications of going to a plan B should also be explained. To that end, the Panel is concerned that rejection of a Plan A assessment, and acceptance of the Plan B approach, obligates the analyst to continue to use the Plan B approach until a research track assessment can be completed. It may be more expedient to allow the analyst to retable an improved Plan A assessment for a Level 3 review at the next assessment cycle.
3. It should also be made clear that the Panel is not expected to provide ad hoc management advice, but is to focus on reviewing the assessment and its results.
4. The NEFSC should consider allowing analysts to be cited as authors of their assessments.
5. An appendix should be added to the Management Track Assessment Peer Review Panel Report that compiles all relevant AOP background information, specifically the summaries of each stock's management track assessment proposal to the AOP and the Summary of the AOP Meeting.

Finally, the missing 2020 spring and fall surveys will create problems in the next set of assessments. As such, the next PRP should be made aware that these missing data will need to be handled in appropriate ways. A table or table(s) documenting survey completeness for the previous ten years should be provided in the background documents.

Stock	Lead Analyst/Presenter	Peer review conclusion
Atlantic Wolffish – Expedited review	Charles Adams	Concurs with the assessment that Atlantic wolffish are overfished and overfishing is not occurring.
Acadian Redfish – Expedited review	Brian Linton	Concurs with the assessment that Acadian redfish are not overfished and overfishing is not occurring.
Atlantic Sea Scallops – Enhanced review	Dvora Hart	Concurs with the assessment that Atlantic sea scallops are not overfished and overfishing is not occurring.
Northern Window Pane Flounder – Expedited review	Toni Chute	Concludes that northern windowpane overfished and overfishing status are unknown.
Southern Window Pane Flounder – Expedited review	Toni Chute	Concurs with the assessment that southern windowpane flounder are not overfished and overfishing is not occurring.
Georges Bank Winter Flounder – Enhanced review	Daniel Hennen	Concurs with the assessment that Georges Bank winter flounder are overfished and overfishing is not occurring.
Gulf of Maine Winter Flounder – Expedited review	Paul Nitschke	Concurs with the assessment that Gulf of Maine winter flounder overfished status is unknown and overfishing is not occurring.
Southern New England Mid-Atlantic Winter Flounder – Enhanced review	Anthony Wood	Concurs with the assessment that Southern New England/Mid—Atlantic winter flounder are overfished and overfishing is not occurring.
Northern Red Hake – Enhanced review	Toni Chute	Concurs with the SARC that northern red hake overfished and overfishing status are unknown, but not likely overfished and overfishing is not likely occurring.
Southern Red Hake – Enhanced review	Toni Chute	Concurs with the SARC that southern red hake overfished and overfishing status are unknown, but overfishing is not likely occurring.
Southern Silver Hake / Offshore Hake – Expedited review	Larry Alade	Concurs with the assessment that Southern silver/offshore hake are not overfished and overfishing is not occurring.

Atlantic Wolffish

The 2020 assessment update for Atlantic wolffish is an expedited review (Level 2 assessment) in accord with the decision at the 27 May 2020 meeting of the Assessment Oversight Panel (AOP). This was recommended because of the need to consider the cumulative effects of updated MRIP data, and a revised knife edge maturity at 52 cm from 50 cm.

This 2020 assessment is an update of the 2008 benchmark assessment and the 2017 operational assessment. This assessment updates commercial fishery catch data, research survey indices of abundance, the analytical SCALE assessment model and reference points through 2019.

The Peer Review Panel (PRP) concludes that the 2020 assessment update for Atlantic wolffish is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes. Retrospective adjustments were not made to the model results. Spawning stock biomass (SSB) in 2019 was estimated to be 676 (mt) which is 44% of the biomass target (SSB_{MSY} proxy = 1,543). The 2019 fully selected fishing mortality was estimated to be 0.005 which is 2% of the overfishing threshold proxy (F_{MSY} proxy = 0.2). The PRP concurs with the assessment that Atlantic wolffish are overfished and overfishing is not occurring.

Atlantic Wolffish Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial landings data are primarily only available through 2010 after which possession was prohibited, and discard data from 2017, 2018 and 2019 were added to those used in the 2017 operational assessment. Similarly, recreational landings are available only through 2011. Recent MRIP data were used to update recreational landing data through 1981. Recreational discards for this stock were not included in the benchmark assessment.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. All three of the survey indices used in the benchmark assessment (NEFSC spring bottom trawl survey, NEFSC fall bottom trawl survey, MA DMF spring trawl survey) were updated through 2019. Few fish were captured in any of the three surveys over the past decade. Catch length frequencies were updated as well, although the only data available since 2011 has come from commercial fishery discards.

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*

   This TOR was satisfactorily addressed. The same SCALE model configuration used in the 2017 operational assessment was used in the 2020 update.

A bridge run was prepared to evaluate the impact of the revised MRIP recreational landings estimate for 1968-2016. Results using the original and revised estimates showed little difference. As a result, a final 1968-2019 assessment was conducted using the revised MRIP landings data. None of the various diagnostics showed significant differences between the predicted and observed values. The change in knife edge maturity had no influence on the trends being simply a scaling factor.
Though a moderate retrospective pattern was observed (Mohn's rho = 0.27 for SSB and -0.14 for F), no retrospective adjustments were made on the assessment. Values of Mohn's rho improved over the 2017 assessment.

A Plan B assessment was unnecessary because the SCALE assessment was accepted.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was satisfactorily addressed. The re-estimated BRPs are the following: F_{MSY} proxy = 0.200, SSB_{MSY} = 1,543 mt, and MSY = 218 mt. The 2019 fully selected fishing mortality was estimated to be 0.005 which is 2% of the overfishing threshold proxy (F_{MSY} proxy = 0.200). The stock is overfished but overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

   Due to the uncertainties in the assessment in general, the Northeast Data Poor Stocks Working Group concluded in 2009 that stock projections would be unreliable and should not be conducted.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   Several research needs were identified by the Peer Review Panel in the 2015 assessment:
   - Potential use of a likelihood profile to apply the criterion for a retrospective adjustment;
   - Further studies on growth parameters;
   - A tagging study to provide information on stock structure and movement; and
   - A study of post-capture nest site fidelity.

   There has been no progress on the use of a likelihood profile to apply the criterion for a retrospective adjustment. A recently completed S-K funded wolffish study included aging of wolfish and a manuscript on sex-specific growth parameters is in preparation. There was also a genetic component for which a manuscript is in preparation. Tagging data are being reviewed to see if they can be published as well. No progress has been made on a study of post-capture nest site fidelity.

   Finally, the issue of the use of the ocean pout calibration coefficient (to calibrate F/V Bigelow survey results to those of the F/V Albatross) will be addressed in the 2023 management track assessment.

Additional Recommendations

1. Evaluate longline survey data as an index for inclusion in the next assessment, or at least as a tool to evaluate the sensitivity of the assessment to changes in abundance.
2. Consider why the current SSB appears to be increasing, while the MSY and SSB_{MSY} appears to be decreasing.

Acadian redfish

The 2020 assessment for Acadian redfish is an expedited review (Level 2) update of the 2017 ASAP based operational assessment, as recommended by the Assessment Oversight Panel (AOP). This recommendation was made based on the addition of age data, new maturity data, investigation of the usefulness of adding a selectivity block and evaluation of the necessity to make a retrospective adjustment.

The first review by Peer Review Panel observed that the two stock size indices used in the ASAP model had been declining more steeply than the estimated biomass in the assessment. The Peer Review Panel considered rejecting the assessment on that basis, but given that the ASAP modelling did not show other problems, the analyst was asked to explore ways to better fit recent survey indices. The analyst found that altering the weighting of the various data sources provided a better fit to recent indices and improved the retrospective pattern. The Peer Review Panel accepted the base case assessment but cautioned that it may overestimate stock size as indicated by the sensitivity run where a different weighting scheme was used.

The Peer Review Panel concludes that the 2020 assessment update for Acadian redfish is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available for this stock for management purposes. Retrospective adjustments were made to the model results. The Peer Review Panel concurs with the assessment that Acadian redfish are not overfished, and overfishing is not occurring.

Acadian redfish Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial landings data were updated with 2017 – 2019 data added to the 1913 – 2016 time series used in the previous assessment. Catch at age data for 2017 was added to the 1969 – 1985 data used in the previous assessment. Total discards for 2017 – 2019 were added to those for 1989 – 2016 used in the previous assessment. Recreational catches and discards are not used in this assessment as agreed in the benchmark assessment. Total catches have varied between 3,900mt and 5,380mt during 2012 – 2019.

   The Peer Review Panel notes that age data have been collected for the entire period but those have not been processed. Additional age data for 1986 – 2016 and for years post 2017 would be likely to decrease uncertainty in the next assessment.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. The NEFSC autumn and spring bottom trawl surveys are used in the ASAP modelling. Both were updated to 2019. Indices at age were available for 1975 – 2019 for the autumn survey and 1975 – 1980, 1984 – 1990 and 2018-2019 for the spring survey.

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*

This TOR was satisfactorily addressed. The ASAP model used catches for 1913-2019 and ages 1 to 26+. Natural mortality was fixed at M = 0.05, selectivity was assumed to equal 1.0 for ages 10 and older. The addition of another selectivity block was found to be not warranted.

As there were no changes to the previous model, a bridge was not necessary. A plan B was prepared but was not needed as the assessment was accepted.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

This TOR was satisfactorily addressed. The re-estimated BRPs include: F_{MSY} = 0.038, B_{MSY} = 200,586mt, and MSY = 7,561mt. The most recent biomass estimate is near 400,000mt which is above B_{MSY}. The 2019 fishing mortality was estimated to be 0.014, which is lower than F_{MSY}. The stock is not overfished, and overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

Projections were carried out following accepted protocols assuming that 5,184mt would be caught in 2020 and setting fishing mortality in 2021-2023 equal to the F_{MSY} proxy of $F_{50\%}$. A retrospective adjustment was applied. Resulting catches at $F_{50\%}$ are 13,525mt for 2021, 13,235mt for 2022 and 12,990mt for 2023. The analytical team will try to complete sensitivity projections under the alternate weighting prior to the PDT and SSC meetings.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

Include additional age data: this was done, but must be continued to include more years.
Investigate effect of using female mean weight at age: this was done. The 2020 assessment uses female only weights at age, but a sensitivity was run using both female and male weights at age. The model predicts higher age 1 recruits when using weights from both sexes, but there is little difference in biomass and fishing mortality estimates.

Explore estimation of stock-recruit relationship internal or external to model: no progress made
Evaluate survey trends and how well they reflect abundance: no progress made
Explore data weighting scenarios to better reflect data quality: this was done during the Peer Review Panel meeting.

Additional Recommendations
The Peer Review Panel strongly recommends that the aging material collected be processed and be made available to be used in the next assessment. Exploration of data weightings should be continued, implying an enhanced review for the next assessment.

Atlantic Sea Scallop

The 2020 assessment for Atlantic sea scallops is an enhanced review (Level 3) update of the 2018 benchmark assessment, as recommended by the Assessment Oversight Panel (AOP), because of a new approach to modeling selectivity in reference points. The 2020 assessment focused on the two models used to determine stock status: CASA and SYM. The forward projection model, SAMS, was not reviewed as part of this assessment The Peer Review Panel was asked to review the mechanics and assumptions behind the Atlantic sea scallop assessment procedure, which employs the CASA model, in preparation for conducting 1) the stock assessment, 2) providing reference point estimates and 3) evaluating stock status. The responses to the Terms of Reference given below focusses on this element of the review. The last Benchmark review occurred in 2018 (SARC 65) at which time the stock was considered not overfished and overfishing was not occurring.

The Peer Review Panel concludes that the 2020 assessment update for Atlantic sea scallop is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available for this stock for management purposes. The Peer Review Panel concludes that Atlantic sea scallops are not overfished and overfishing is not occurring.

Sea Scallop Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards.*

   This TOR was satisfactorily addressed. Catch landings and discard data were provided by region. Landings are near record high levels.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. Sea scallop surveys included the dredge survey, the drop camera survey, which uses both video and digital still cameras, and the Habcam still camera survey. Neither the Habcam nor the drop camera surveys were fully completed on Georges Bank in 2019. Information from the two camera surveys was combined to cover all management areas. Indices from the three survey approaches track each other well although the Habcam survey was more sensitive to tracking the strong recruitment pulses representing the 2012 year class in Georges Bank and the 2013 year class in the Mid-Atlantic. The high recruitments were followed by high natural mortality in the recruiting year class that may be density dependent.

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*

This TOR was satisfactorily addressed. The CASA model, a size-based forward projection model, has been used since 2007. Growth was adjusted for observed slower growth in recent years and fishery selectivity periods for 2018 and 2019 were added to the Georges Bank Closed area model but no major structural changes were made to the CASA model. Current model output was compared to that from the 2018 benchmark assessment and the two assessments track each other well until the most recent years when the strong recruitment pulses followed by a period of high natural mortality on those recruits leads to an apparent and reasonably explainable retrospective pattern. Time varying natural mortality rates are estimated by the model using information from closed areas and sublegal elements of the population that provide information on natural mortality through observed declines in population density when no fishing mortality is present.

A Plan B assessment was not considered necessary as the assessment procedure was approved.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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 Stochastic Yield Model (SYM) was used for estimating the Biological Reference Points for Georges Bank and the Mid-Atlantic and combined whole stock. SYM combines per-recruit calculations with stock-recruit relationships to obtain probabilistic MSY-based reference point estimates. The assumed standard deviation of natural mortality in the SYM model was reduced for model runs this cycle. This had very little effect on the reference points but helped to stabilize the model. The revised Biological Reference Point estimates are: F_{MSY} = 0.61 and B_{MSY} = 102,657 mt meats. This compares to F_{MSY} = 0.64 and B_{MSY} = 116,766 mt meats from the 2018 assessment. The stock is not overfished and overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

No short-term projections were provided. Projections are prepared through a separate process by the scallop PDT as a direct delivery to the SSC using the current year's survey data. The 2020 survey data updates are not yet available.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

Research Recommendations from 2018 Benchmark

1. Further investigate methods for better survey coordination between the various survey programs, including survey design, timing, and standardized data formatting for easier sharing

The NEFMC is organizing a committee to facilitate coordination.

2. Investigate changes in dredge efficiency and saturation due to high scallop densities or high bycatch rates

No substantial progress has been made beyond that presented in the 2018 assessment. However, the issue has become less acute as densities have declined.

3. Analyze past juvenile scallop mortality events and develop better methods to model time-varying mortality in the assessment models

Some progress has been made – see the CASA models from this assessment.

4. Collect information needed for the management of the GOM fishery and development of appropriate reference points including biological parameters, fishery-independent surveys, and fishery-dependent data.

Survey and growth data are being collected from the GOM. However, it is difficult to estimate reference points due to the lack of a time series in the area.

5. Continue development of scallop ageing methods and examination of scallop growth processes including density dependent effects.

An RSA funded project is underway towards these goals.

6. Improve training of annotators used in optical surveys and develop standardized QA/QC procedures for data collected from imagery.

New QA/QC procedures are being developed for the NEFSC Habcam survey, and should be ready for implementation in 2021.

7. Investigate use of software for automated annotation of imagery from optical surveys.

Work towards this goal is underway; see https://www.fisheries.noaa.gov/feature-story/computers-now-see-animals-ocean-bottom

8. Investigate methods to better estimating biomass and abundance variances from Habcam optical surveys including development of Bayesian geostatistical methods.

A Bayesian geostatistical model is being developed.

9. Investigate and estimate current and historical unreported landings and effects of spatially heterogeneous fishing mortality on mortality estimates.

No progress has been made.

10. Develop a spatially-explicit methodology for forecasting the abundance and distribution of sea scallops by incorporating spatial data from surveys, landings, and fleet effort (aka GEOSAMS).

No progress has been made, although there are plans to develop a GEOSAMS model in the coming year.

11. Investigate and parameterize sub-lethal effects of disease, parasites, or discarding on mortality, growth, and landings.

There are ongoing projects investigating gray meats and nematode infections.

12. Revive and streamline previously-developed methods for interpreting VMS data.

No progress

13. Further refine and test methods for forecasting LPUE.

Some progress has been made developing spatial choice models for scallop fishermen, which would affect LPUE forecasts.

14. Continued investigation of discard mortality, particularly during warm water periods, by incorporating environmental data.

No progress.

15. Continue improvements of observer recordings for vessel fishing behavior including deck loading and shucking dynamics in responses to disease or poor scallop health.

Observer protocols were modified a few years ago to better track scallop health and meat condition.

16. Continue investigating the extent of incidental fishing mortality, particularly on hard bottom habitats.

No progress.

Research Recommendations from the SSC

1. Different growth rates found in different scallop harvesting areas, particularly the Nantucket Lightship region.

Shells have been collected from this area and aged. Analysis is in progress.

2. Further work to develop gonad-based estimates of SSB and reference points

No progress.

3. Runs of previous assessment model configurations to compare to new version of assessment.

No progress, although we routinely present an historical retrospective, which compares current and previous assessment models.

Additional Recommendations
In addition, the Panel was asked to evaluate a new approach being developed in the SYM model that models selectivity dynamically as a function of full recruitment fishing mortality. It was demonstrated that this dynamic selectivity model better characterizes selectivity as fishing mortality moves towards more extreme values and can result in more precise estimates of biological reference points when taken into account. The Peer Review Panel encouraged continued development of this approach and while it will not be used in this upcoming assessment cycle we look forward to its full implementation at a later date.

Northern Windowpane Flounder

The 2020 assessment for northern windowpane flounder is an expedited review (Level 2) update of the 2019 assessment, as recommended by the Assessment Oversight Panel (AOP). This recommendation was made because the AIM model used to assess the stock in previous assessments has been performing poorly and the Plan B approach was anticipated as a replacement method using new chainsweep study information. The 2020 assessment updated commercial fishery catch data and survey biomass indices to update the AIM model outputs, as well as applied an empirical approach based on the recent catchability study (Miller et al., 2020) to estimate swept-area biomass and annual exploitation rates.

The Peer Review Panel concluded that the AIM model should no longer be used as the basis for catch advice due to the lack of significance in the relationship between population response and fishing mortality. The Panel concluded that the Plan B approach based on estimated swept-area biomass calculated from survey catchability estimates specific to northern windowpane should be the basis of catch advice. This approach does not allow estimation of retrospective patterns, projections, or biological reference points; in the absence of reference points, overfished and overfishing status are unknown.

The Peer Review Panel concludes that the 2020 Plan B for Northern window pane flounder is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available for this stock for management purposes.

Northern Windowpane Flounder Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial landings data were updated through 2019, but possession of northern windowpane has been prohibited since 2010. Commercial discards are estimated from large and small mesh otter trawl gear and Limited Access and General Category scallop dredges and trawls. There are no recreational data for northern windowpane flounder. Total catch in 2019 was 43 mt, all of which was discards.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. Due to seasonal migration of northern windowpane flounder, only the NEFSC fall bottom trawl survey is included in the assessment. The survey index (kg/tow in Albatross units) and the swept-area biomass estimates applying northern windowpane specific catchability estimates for the Bigelow survey were updated through 2019. Survey length frequencies by proportion of total survey catch were updated through 2019 showing a bimodal distribution throughout the time series.

   The 2017 NEFSC fall bottom trawl survey had a reduced number of sampled stations in some strata. The strata with incomplete sampling were examined relative to the full survey time series and only a minor effect on the overall biomass index was detected.

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*

   This TOR was satisfactorily addressed. The AIM model was updated with commercial catch and NEFSC fall bottom trawl survey data through 2019 but continued to perform poorly as seen in previous update assessments in 2017 and 2019. The model previously suggested that the stock was not responding to very low catches and estimated increasing B_{MSY} proxies and decreasing F_{MSY} proxies. The 2020 updated proxies were marginally improved from the 2019 estimates but the relationship of biomass replacement to relative F remained uninformative. For these reasons, the AIM assessment was rejected by the Peer Review Panel.

   The Plan B approach is an empirical method to estimate swept-area biomass and annual relative exploitation rates based on the recent catchability study specific to northern windowpane flounder (Miller et al., 2020). Catch efficiency was estimated annually for the Bigelow time series (2009-2019) and the mean of those estimates was applied to the prior survey time series (1975-2009).

Exploitation rates are expressed as a percent of the estimated biomass removed by the fishery (catch/biomass) for each calendar year.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was partially addressed. The mean exploitation rate for the years 1995-2001 (1.67%) was used as an F_{MSY} proxy. The time period corresponds to the period used to generate an MSY estimate for the AIM model. Northern windowpane flounder is currently in a rebuilding plan with an $F_{rebuild}$ of 70% F_{MSY}, resulting in a target exploitation rate of 1.17%. The mean exploitation rate for the most recent 3-year running average is 1.04%, with a terminal estimate of 0.34%. Based on this F_{MSY} proxy, overfishing is not occurring. The Peer Review Panel did not recommend continued use of the AIM-based F_{MSY} proxy due to the mismatch in assessment methods and time series of exploitation rates exceeding the proxy in nearly all years. In the absence of agreed reference points, the Panel concluded that stock status is currently unknown due to the empirical assessment approach but noted that recent exploitation rates have been very low.

5. *Conduct short-term stock projections when appropriate.*

   There were no projections made for the northern windowpane flounder stock. The Peer Review Panel noted that recent exploitation rates have been constrained by management actions, specifically a no possession regulation starting in 2010. They recommended exploration of exploitation rates over different time periods with consideration of current stock conditions. The choice of exploitation rate has important implications for scientific advice to management.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   The 2019 Peer Review Panel rejected the updated F_{MSY} estimate from the AIM model and recommended continued use of the 2017 reference points. They recommended reconsideration of the entire assessment approach. Additionally, they noted the large amount of uncertainty associated with discard estimates for some fleets and the potential impact on this assessment due to the prohibition on landings.

   The 2020 AIM model update was rejected and the Plan B approach using the recent catchability study was applied. The change in method is an improvement and addresses the previous Panel's concern about the uninformative relationship of biomass replacement to relative F.

Additional Recommendations

Include an Appendix to the 2020 report summarizing the inputs/assumptions/steps used to develop the Northern windowpane flounder empirical approach.

Appropriate exploitation rates should be further explored in the next assessment.

Southern Windowpane Flounder

The 2020 assessment for southern windowpane flounder is an expedited review (Level 2) update of the 2019 assessment, as recommended by the Assessment Oversight Panel (AOP). This recommendation was made based on the potential to rescale the survey indices to swept-area biomass estimates. The 2020 assessment updated commercial fishery catch data, survey indices of abundance, the AIM model outputs, and reference points through 2019.

The Peer Review Panel concluded that the AIM model is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available for this stock for management purposes. The mean NEFSC fall bottom trawl survey index from the most recent three-year moving average (2017-2019) was 0.288 kg/tow, which is higher than the $B_{threshold}$ value of 0.097 kg/tow and higher than the B_{MSY} proxy value of 0.195 kg/tow. The 2019 relative fishing mortality was estimated to be 1.210 kt per kg/tow, which is lower than the F_{MSY} proxy of 1.708 kt per kg/tow. The Peer Review Panel concurs with the assessment that southern windowpane flounder is not overfished and overfishing is not occurring.

Southern Windowpane Flounder Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial landings data were updated through 2019, but possession of southern windowpane has been prohibited since 2010. Commercial discards are estimated from large and small mesh otter trawl gear and Limited Access and General Category scallop dredges and trawls. There are no recreational data for southern windowpane flounder. Total catch in 2019 was 374 mt.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. The NEFSC fall bottom trawl survey is the only index included in the assessment because it is considered more stable for southern windowpane than the spring survey. The survey index (kg/tow in Albatross units) was updated through 2019. The NEFSC fall bottom trawl survey was unable to cover any of the southern windowpane stock strata in 2017. The estimate for 2017 was imputed by averaging the mean survey biomass per tow values from 2016 and 2018 by stratum to calculate a stratum-weighted index. Information from the Northeast Area Monitoring and Assessment Program (NEAMAP) survey was qualitatively compared to the NEFSC survey time series and showed a similar value to the 2017 imputed index.

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*

This TOR was satisfactorily addressed. The same AIM model configuration used in the 2019 operational assessment was used in the 2020 update. The AIM model performs well for this stock as indicated by the significant relationship between population response and fishing mortality. Bridge runs were not needed because only the data inputs for the model were updated. The Plan B assessment was reviewed but not recommended because the AIM model assessment was accepted.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was satisfactorily addressed. The re-estimated BRPs include: F_{MSY} proxy = 1.708 kt per kg/tow, B_{MSY} proxy = 0.195 kg/tow, and MSY proxy = 333 mt. The most recent three-year biomass index was estimated to be 0.288 kg/tow, which is above the B_{MSY} proxy. The 2019 relative fishing mortality was estimated to be 1.210 kt per kg/tow, which is lower than the F_{MSY} proxy. The stock is not overfished and overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

   There were no projections made for the southern windowpane flounder stock.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   The 2019 assessment for southern windowpane flounder was a direct delivery (level 1) assessment that was not reviewed. The SSC did not express concerns with the 2019 assessment.

Additional Recommendations
The Peer Review Panel noted that the AIM model has performed well for the southern windowpane stock, but not the northern stock. They recommended future analyses to determine the mechanism driving the performance of this modeling approach.

Georges Bank Winter Flounder

The 2020 assessment update for Georges Bank winter flounder received a Level 3 Enhanced Review in accord with the decision at the 27 May 2020 meeting of the Assessment Oversight Panel (AOP). The lead analyst proposed to the AOP to transition the current MSY biological reference points (calculated from the model stock-recruitment relationship) to proxy-based reference points ($F_{40\%}$, $SSB_{40\%}$) to match the Gulf of Maine winter flounder stock and recommendations of a panel review in 2019. The AOP discussed the potential impact of changing reference points given that the stock is in a rebuilding plan and recommended that the old method should also be calculated for continuity. The AOP agreed that

the Level 3 Enhanced Review recommended by the lead analyst is appropriate given the proposed change to reference points.

The 2020 assessment of the Georges Bank winter flounder stock is an update of the existing 2019 operational VPA assessment which included data for 1982-2018. This assessment updates commercial fishery catch data, research survey biomass indices, and the analytical VPA assessment model and new $F_{40\%}$ and $SSB_{40\%}$ reference points proxies through 2019. Stock projections have been updated through 2023.

The Peer Review Panel (PRP) concludes that the 2020 assessment update for Georges Bank winter flounder is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes. Spawning stock biomass (SSB) in 2019 was estimated to be 4,061 mt. The 2019 fully selected fishing mortality (F) was estimated to be 0.088. However, the 2019 point estimate of SSB and F, when adjusted for retrospective error (0.57% for SSB and -0.34% for F), are outside the 90% confidence intervals of the unadjusted 2019 point estimates. Therefore, the values used in the stock status determination were the retrospective-adjusted values of F_{2019} = 0.133 which is 37% of the overfishing threshold (F_{MSY} = 0.358), and SSB_{2019}=2,587 mt, which is 36% of the biomass target for an overfished stock (SSB_{MSY} = 7,267 mt, with a threshold of 50% of SSB_{MSY}).

The PRP concurs with the assessment that Georges Bank winter flounder stock is overfished but that overfishing is not occurring.

Georges Bank Winter Flounder Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial (US and Canadian) landings and discard data from 2019 were added to those used in the 2019 operational assessment.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. All three of the survey indices used in the benchmark assessment (NEFSC spring bottom trawl survey, NEFSC fall bottom trawl survey (lagged forward one year and age), Canadian spring trawl survey) were updated through 2019 (DFO through 2020). Commercial catch at age, and catch weight at age data from 2019 were added to those used in the 2019 operational assessment.

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*

This TOR was satisfactorily addressed. The same VPA model configuration used in the 2019 operational assessment was used in the 2020 update. However, there was concern about the absence of the spring 2020 NEFSC trawl survey (not conducted on Georges Bank because of Covid-19) and the effect of a new stock assessment analyst. As a result, two bridge runs were prepared to compare with the 2020 assessments results (i.e., rerunning the 2019 assessment with the new analyst's data decisions, rerunning the 2019 assessment without the spring 2019 NEFSC survey, and the 2020 assessments). Results suggest neither had a significant impact on the assessment results.

As in the 2019 assessment, there was a major retrospective pattern (Mohn's rho = 0.57 for SSB, -0.34 for F, and 0.45 for recruitment). Because the 2019 point estimate of SSB and F, when adjusted for retrospective error, were outside the 90% confidence intervals of the unadjusted 2019 point estimates it was necessary to retrospective-adjust both point estimates.
A Plan B assessment was not evaluated because the VPA assessment was accepted.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

This TOR was satisfactorily addressed with the PRP supporting the use of $F_{40\%}$ as the F_{MSY} proxy. Spawning stock biomass (SSB) in 2019 was estimated to be 4,061 mt. The 2019 fully selected fishing mortality (F) was estimated to be 0.088. The retrospective-adjusted values used in the stock status determination were F_{2019} = 0.133 which is 37% of the overfishing threshold (F_{MSY} = 0.358), and SSB_{2019}=2,587 mt, which is 36% of the biomass target for an overfished stock (SSB_{MSY} = 7,267 mt, with a threshold of 50% of SSB_{MSY}). The stock was overfished but overfishing was not occurring in 2019.

5. *Conduct short-term stock projections when appropriate.*

This TOR was satisfactorily addressed. Short-term projections of biomass were derived by sampling from a cumulative distribution function of recruitment estimates (1982-2018 year classes) 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 (2015-2019). An SSB retrospective adjustment factor of 0.637 was applied in the projections. The 2020 estimated catch was from the Plan Development Team (PDT) and 2021-2023 catches were projected from the F_{MSY} proxy ($F_{40\%}$). The PRP notes that recruitment from the 2019 year class is likely to be underestimated.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

The AOP commented that the completed Plan A operational assessment is appropriate for assessing stock status. However, the AOP was concerned about the reference point definitions and

recruitment assumptions in projections. Specifically, using a fixed steepness value may not be appropriate and a F40% MSP F_{MSY} proxy might be a more stable and reliable estimator.

- This was done in this assessment with $F_{40\%}$ recommended as an appropriate proxy.

The residual pattern in the stock-recruitment relationship indicates that recent recruitment has been weaker than expected. Alternative projections should be considered that assume future recruitment will be similar to recent recruitment.

- Sensitivity analyses were conducted to evaluate various recruitment scenarios

There is poor tracking of cohorts in many 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.

- This remains an issue but would need to be pursued via the Research Track process

Explorations regarding the source of the retrospective pattern and recent poor recruitment for this stock.

- No progress to date

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.

- No progress to date

Additional Comments

1. The PRP acknowledged the utility of having additional data provided on how the $F_{40\%}$ was calculated.
2. PRP agrees that future analysis of the stock could be improved using a model that incorporates statistical fits to commercial length and age composition.

Gulf of Maine Winter Flounder

The 2020 assessment update for Gulf of Maine winter flounder is an expedited review (Level 2) in accord with the decision at the 27 May 2020 meeting of the Assessment Oversight Panel (AOP). The AOP discussed the issue of changing to a two-year average of biomass and whether the changes may be significant enough to warrant the elevation of the proposed level 2 review to a level 3, but ultimately recommended an expedited review.

This assessment of the Gulf of Maine winter flounder stock is an update of the existing 2017 area-swept operational assessment to include 2018-2019 catch and survey data.

Using the length-based Yield Per Recruit (YPR) relationship from SARC 52 and an M = 0.3, the $F_{MSY} = F_{40\%}$ = 0.31. The 2019 30+ cm exploitation rate was estimated to be 0.052 which is 23% of the overfishing exploitation threshold proxy (E_{MSY} proxy = 0.23). Biomass (30+ cm mt) in 2019 was estimated to be 2,862 mt.

The Peer Review Panel (PRP) concludes that the 2020 assessment update for Gulf of Maine winter flounder is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes. It concurs that Gulf of Maine winter flounder's overfished status is unknown and overfishing is not occurring.

Gulf of Maine Winter Flounder Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial landings and discard data were estimated for 2009-2019 by adding 2018 and 2019 data to those used in the 2017 operational assessment. Recent MRIP data were used to update recreational discard and landings for 2009-2019.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. All three of the survey indices used in the 2017 operational assessment (NEFSC fall bottom trawl survey, Massachusetts Division of Marine Fisheries fall trawl survey, and Maine-New Hampshire fall trawl survey) were updated for 2009-2019. Additional tows were also available from the NEFSC's twin trawl experiment for revised estimates of Q.

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*

   This TOR was satisfactorily addressed. The assessment was based on the 30+ cm Survey Area-Swept Calculation used in the 2017 Operational Assessment (developed in SARC 52 (2011)) and also used in the 2014 and 2015 Operational assessments. The major difference was in the re-estimated Q value (fall = 0.709, spring = 0.623), both of which are lower than the Q from the 2017 survey (Q = 0.866). The impact of this change in Q would be a relative increase of swept area biomass compared to that which would be calculated from the previous Q.
   A Plan B assessment was unnecessary because the area swept approach was accepted.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was satisfactorily addressed. Using the length based YPR from SARC 52 and an M=0.3 the $F_{MSY} = F_{40\%} = 0.31$. The 2019 30+ cm exploitation rate was estimated to be 0.052 which is 23%

of the overfishing exploitation threshold proxy (E_{MSY} proxy = 0.23). Biomass (30+ cm mt) in 2019 was estimated to be 2,862 mt.

Based on this updated assessment, the Gulf of Maine winter flounder stock's overfished status is unknown and overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

   This TOR was satisfactorily addressed. However, projections are not possible with area-swept based assessments. The Peer Review Panel agrees that catch advice be based on 75% of $E_{40\%}$ (75% E_{MSY} proxy) using the most recent two years of information from fall surveys for the biomass estimate and catch advice.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   a. Additional studies on federal and state survey gear efficiency would be useful. For example, quantifying the degree of herding between the doors and escapement under the footrope and/or above the headrope for state surveys is warranted.
   b. Studies quantifying winter flounder abundance and distribution among habitat types and within estuaries could improve biomass estimates.
   c. A moving average approach to estimating catch advice (rather than based on a single year) should be considered to stabilize catch advice. This was completed in this assessment.

Additional Recommendations

1. Consider using both spring and fall surveys to provide catch advice in the next assessment.
2. Evaluate using Q = 1.0 for the two state surveys in the next assessment.

Southern New England Mid-Atlantic winter flounder

The 2020 assessment for Southern New England Mid-Atlantic winter flounder is an enhanced review (Level 3) update of the 2017 ASAP operational assessment, as recommended by the Assessment Oversight Panel (AOP). This recommendation was made because of changes to the selectivity blocks and selectivity form, the inclusion of new MRIP data, changes to the reference points and possible inclusion of the NEAMAP data.

The Peer Review Panel concludes that the 2020 ASAP assessment update for Southern New England Mid-Atlantic winter flounder is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes. Retrospective adjustments were not made to the model results, but the retrospective adjusted value for SSB and F were close to the 90% confidence interval. In the previous assessment, the stock was considered overfished but overfishing was not occurring. In the current assessment, the F MSY proxy (F 40%) = 0.284, SSB MSY = 12,322 mt and ½ SSB MSY (SSB threshold) = 6,161 mt. F 2019 /F MSY = 27% and SSB 2019 /SSB threshold = 60%. The Peer Review Panel concurs with the assessment that the stock is overfished but that overfishing is not occurring.

And in the table under TOR 4 the SSBMSY value needs to be changed to 12,322 mt.

Southern New England Mid-Atlantic winter flounder Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

   This TOR was satisfactorily addressed. Commercial and recreational landings and discards data were updated through 2019. Total catch in 2019 was 310 mt, a third of which were commercial discards and two thirds commercial landings. Total catches have been less than 1,000 mt since 2009 except in 2013.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. Fishery independent indices of stock sizes for 1981-2019 and ages 0-7+ were used. In total, twelve indices were used, including two for recruits. Surveys generally showed declining stock sizes with much lower values since the early 2000s compared with previous years. The Massachusetts Division of Marine Fisheries age 0 survey showed variability without clear trend.

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*

   This TOR was satisfactorily addressed. Bridge runs were made for each change. Adding a third selectivity block from 2010 resulted in very similar selectivities for the three blocks. Assuming flat topped selectivity rather than dome-shaped reduced the biomass estimate and increased (marginally) the retrospective. Recreational catches were small and have little influence on the ASAP results. The NEAMAP survey was included but did not produce large changes in estimates.
   A plan B was prepared but was not necessary.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was satisfactorily addressed. In previous assessments, MSY reference points were calculated based on a stock and recruitment relationship with recent recruitments being consistently and significantly below predicted values. In addition, most other groundfish stocks assessed by the NEFSC use $F_{\%SPR}$ to estimate reference points. $F_{40\%}$ values were estimated:

	2017	2020
$F_{MSY}/F_{40\%}$	0.34	0.284
$SSB_{MSY} – SSB_{40\%}$	24687	12322
$MSY – MSY_{40\%}$	7532	3906

5. *Conduct short-term stock projections when appropriate.*

   This TOR was satisfactorily addressed. Short-term projections were made following standard protocols, without retrospective adjustment, assuming a catch of 251 mt in 2020 and fishing at $F_{40\%}$ in 2021-2023.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   All recommendations directly related to the assessments have been implemented. The main research recommendations for stock suggest additional studies on maximum age, maturity, movement, localized stock structure and environmental influence on recruitment. Considerable progress has been made on some of these topics since the last benchmark assessment and much of this research continues.

   There has been new research investigating maturity at the science center which can be used to update the maturity ogive during the next research track.

   A 2020 publication out of SUNY Stony Brook details work on otolith micro-chemistry that reveals new information on localized stock structure. A simulation study could be carried out to investigate the impacts on overall stock dynamics and the current stock assessment.

   An environmental model for this stock has been developed and is presented in a 2018 publication (Bell et. al 2018). This model and indices were updated for this assessment cycle. However, in order to fully investigate and possibly shift to a new assessment model a research track assessment will be needed.

Additional Recommendations
The Peer Review Panel notes, as had been done in previous reviews, that recruitment had been declining throughout the period and was currently very low. As for several other stocks under the purview of the NEFSC it would be helpful to evaluate if the previously observed high recruitment are possible, i.e. is it simply a matter of building back SSB and recruits will follow, or are there other factors at play. If the productivity of the resource(s) has decreased, it would be helpful to adjust reference points accordingly. This would be unlikely to change fisheries yield much but would be more realistic in terms of setting expectations.

Northern red hake

The 2020 assessment for northern red hake is an enhanced review (Level 3) of approaches described in the 2020 Red Hake Stock Structure Research Track assessment. This recommendation was made

because the AIM model used to assess the stock in previous assessments was rejected by the Research Track Stock Assessment Review Committee (SARC) and a new assessment approach was not recommended. The SARC recommended using new chainsweep study information for northern red hake to estimate swept-area biomass but did not recommend an approach to determine BRPs. The 2020 assessment updated commercial and recreational fishery catch data and survey biomass indices. The Peer Review Panel reviewed an empirical approach based on the recent survey catchability study to estimate swept-area biomass and annual relative exploitation rates. This approach has been applied and peer-reviewed for flatfish stocks. The Panel concluded that the updated swept-area biomass estimates provide qualitative information about stock trends, but the relative exploitation rates should not be used as BRP proxies and do not provide a basis for scientific advice. The Panel concurs with the SARC that the exploitation rates are currently low, and that overfishing is not likely occurring. Additionally, recent survey estimates indicate that the population is at a relatively high level and it is unlikely that the stock is overfished.

The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes.

Northern Red Hake Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

This TOR was satisfactorily addressed. Commercial and recreational landings data were updated through 2019. Recreational catch was based on uncalibrated MRIP data for the full time series. Commercial discards are estimated from several gear types with the majority attributed to small and large mesh otter trawl. Total catch in 2019 was 236 mt, of which was 110 mt was discards.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

This TOR was satisfactorily addressed. The NEFSC spring and fall bottom trawl survey indices (kg/tow in Albatross units) and the swept-area biomass estimates applying northern red hake specific catchability estimates for the Bigelow survey were updated through 2019. A sensitivity analysis was conducted using the NEFSC spring bottom trawl survey only and showed only minor differences in swept-area biomass or relative exploitation rates compared to the combined survey estimates.

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*

This TOR was not met. The 2020 Red Hake Stock Structure Research Track assessment SARC rejected the AIM model for northern red hake. The SARC recommended use of swept-area biomass estimates based on the chainsweep study for northern red hake and reviewed an alternative method for calculating reference points based on spawning potential ratio (SPR), but concluded that *"there was sufficient uncertainty in the sensitivity of reference point estimates to various assumptions made that the reference point estimates should not be used for management advice for red hake at this time."* The SARC recommended additional analyses for the SPR approach and noted that methods currently used for other data-limited stocks in the region could be explored for both northern and southern red hake. The SARC did not recommend an assessment method.

The Peer Review Panel reviewed a proposed Plan B approach based on an empirical method to estimate swept-area biomass and annual relative exploitation rates based on the recent catchability study specific to northern red hake (Miller et al., 2020). Catch efficiency was estimated annually for the Bigelow time series (2009-2019) and the mean of those estimates was applied to the prior survey time series (1981-2009). Exploitation rates are expressed as a percent of the estimated biomass removed by the fishery (catch/biomass) for each calendar year. The Panel concluded that the updated swept-area biomass estimates provide qualitative information about stock trends, but the relative exploitation rates do not provide a basis for scientific advice.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was partially addressed. BRPs could not be estimated from the proposed Plan B approach. In the absence of agreed reference points, the Panel concluded that stock status is currently unknown. The Panel reviewed the updated biomass estimates and relative exploitation rates and concluded that the exploitation rates are currently low, and that overfishing is not likely occurring. Additionally, recent survey estimates indicate that the population is at a relatively high level and it is unlikely that the stock is overfished.

   Reference points that were applied in the previous assessments were based on survey indices (kg/tow). These reference points could be evaluated for application to the updated swept-area biomass estimates and potential use in management if they were converted to swept-area biomass.

5. *Conduct short-term stock projections when appropriate.*

   There were no projections made for northern red hake. The Peer Review Panel noted that recent exploitation rates have been constrained by management actions that were based on the rejected AIM model.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   The 2020 Red Hake Stock Structure Research Track assessment SARC made several recommendations for further evaluation of the proposed SPR-based assessment method. The SARC

noted that the SPR-based reference points could be suitable for red hake and that the 40% proxy level for F and SSB was reasonable. They suggested the following analyses:

- A catch curve analysis on the survey data could be used to estimate M in recent years;
- Exploration of the sensitivity of the knife-edge selectivity assumption; and
- Expansion of the time series of recruitment estimates over longer periods and evaluation of the sensitivity of the $SSB_{40\%}$ estimates to different recruitment time series.

The SARC also noted that decoupling between fishing pressure and population trends has been observed for other stocks in the region (e.g., Georges Bank yellowtail flounder) and suggested that methods currently used for setting catch advice for other data-limited stocks could be explored for red hake.

Additional Recommendations

The Peer Review Panel recommended additional analysis on the proposed SPR-based assessment method, as described by the SARC. They noted that due to the Research Track and Management Track process, there is not a currently accepted assessment method for the red hake stocks and no basis for scientific advice at this time. The Panel recommended a subsequent review process for a newly developed red hake assessment.

Appropriate exploitation rates should be further explored in the next assessment.

Southern red hake

The 2020 assessment for southern red hake is an enhanced review (Level 3) of approaches described in the 2020 Red Hake Stock Structure Research Track assessment. This recommendation was made because the AIM model used to assess the stock in previous assessments was rejected by the Research Track Stock Assessment Review Committee (SARC) and a new assessment approach was not recommended. The SARC recommended using new chainsweep study information for southern red hake to estimate swept-area biomass but did not recommend an approach to determine BRPs. The 2020 assessment updated commercial and recreational fishery catch data and survey biomass indices.

The Peer Review Panel reviewed an empirical approach based on the recent survey catchability study to estimate swept-area biomass and annual relative exploitation rates. This approach has been applied and peer-reviewed for flatfish stocks. The Panel concluded that the updated swept-area biomass estimates provide qualitative information about stock trends, but the relative exploitation rates should not be used as BRP proxies and do not provide a basis for scientific advice. The Panel concurs with the SARC that the exploitation rates are currently low, and that overfishing is not likely occurring. Additionally, southern stock survey indices are near the lowest in the time series, and the Panel agrees with the SARC that overfished status is unknown.

The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes.

Southern Red Hake Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards*.

This TOR was satisfactorily addressed. Commercial and recreational landings data were updated through 2019. Recreational catch was based on uncalibrated MRIP data for the full time series. Commercial discards are estimated from several gear types with the majority attributed to small mesh otter trawl. Total catch in 2019 was 1,889 mt, of which was 1,239 mt was discards.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. The NEFSC spring and fall bottom trawl survey indices (kg/tow in Albatross units) and the swept-area biomass estimates applying southern red hake specific catchability estimates for the Bigelow survey were updated through 2019. A sensitivity analysis was conducted using the NEFSC spring bottom trawl survey only and showed only minor differences in swept-area biomass or relative exploitation rates compared to the combined survey estimates.

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*

   This TOR was not met. The 2020 Red Hake Stock Structure Research Track assessment SARC rejected the AIM model for southern red hake. The SARC recommended use of swept-area biomass estimates based on the chainsweep study for southern red hake and reviewed an alternative method for calculating reference points based on spawning potential ratio (SPR), but concluded that *"there was sufficient uncertainty in the sensitivity of reference point estimates to various assumptions made that the reference point estimates should not be used for management advice for red hake at this time."* The SARC recommended additional analyses for the SPR approach and noted that methods currently used for other data-limited stocks in the region could be explored for both northern and southern red hake. The SARC did not recommend an assessment method.

   The Peer Review Panel reviewed a proposed Plan B approach based on an empirical method to estimate swept-area biomass and annual relative exploitation rates based on the recent catchability study specific to southern red hake (Miller et al., 2020). Catch efficiency was estimated annually for the Bigelow time series (2009-2019) and the mean of those estimates was applied to the prior survey time series (1981-2009). Exploitation rates are expressed as a percent of the estimated biomass removed by the fishery (catch/biomass) for each calendar year. The Panel concluded that the updated swept-area biomass estimates provide qualitative information about stock trends, but the relative exploitation rates do not provide a basis for scientific advice.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

   This TOR was partially addressed. BRPs could not be estimated from the proposed Plan B approach. In the absence of agreed reference points, the Panel concluded that stock status is currently unknown. The Panel reviewed the updated biomass estimates and relative exploitation rates and concluded that the exploitation rates are currently low, and that overfishing is not likely occurring. Additionally, southern stock indices are near the lowest in the time series, and the overfished status is unknown.

   Reference points that were applied in the previous assessments were based on survey indices (kg/tow). These reference points could be evaluated for application to the updated swept-area biomass estimates and potential use in management if they were converted to swept-area biomass.

5. *Conduct short-term stock projections when appropriate.*

   There were no projections made for southern red hake. The Peer Review Panel noted that recent exploitation rates have been constrained by management actions that were based on the rejected AIM model.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

   The 2020 Red Hake Stock Structure Research Track assessment SARC made several recommendations for further evaluation of the proposed SPR-based assessment method. The SARC noted that the SPR-based reference points could be suitable for red hake and that the 40% proxy level for F and SSB was reasonable. They suggested the following analyses:
   - A catch curve analysis on the survey data could be used to estimate M in recent years;
   - Exploration of the sensitivity of the knife-edge selectivity assumption; and
   - Expansion of the time series of recruitment estimates over longer periods and evaluation of the sensitivity of the $SSB_{40\%}$ estimates to different recruitment time series.

   The SARC also noted that decoupling between fishing pressure and population trends has been observed for other stocks in the region (e.g., Georges Bank yellowtail flounder) and suggested that methods currently used for setting catch advice for other data-limited stocks could be explored for red hake.

Additional Recommendations

The Peer Review Panel recommended additional analysis on the proposed SPR-based assessment method, as described by the SARC. They noted that due to the Research Track and Management Track process, there is not a currently accepted assessment method for the red hake stocks and no basis for scientific advice at this time. The Panel recommended a subsequent review process for a newly developed red hake assessment.

Appropriate exploitation rates should be further explored in the next assessment.

Southern Silver Hake / Offshore hake

The 2020 assessment for southern silver hake/offshore hake is an expedited review (Level 2) update of the 2017 assessment as recommended by the Assessment Oversight Panel (AOP) based on missing 2017 survey data and potential data imputation approaches. The last Benchmark review occurred in 2010 (SARC 51) when the ASAP model was considered and rejected and the empirical approach was adopted.

The Peer Review Panel concludes that the 2020 assessment update for southern silver hake / offshore hake is technically sufficient to evaluate stock status and provide scientific advice. The assessment represents Best Scientific Information Available for this stock for management purposes. The Peer Review Panel concurs with the assessment that the stock is not overfished and that overfishing is not occurring.

Southern Silver Hake / Offshore Hake Terms of Reference (TOR)

1. *Estimate catch from all sources including landings and discards.*

   This TOR was satisfactorily addressed. Catch, landings (foreign and domestic) and discard data for the southern silver Hake / offshore hake complex were provided by year and by fleet. Landings in 2019 were 5,400mt and well below the time series average of 32,670mt (1955-2019) and a more recent time series average of 6,198mt (2000-2019). The southern and offshore silver hake are combined as identification based on morphology can be difficult. There is an algorithm for separating the two species out by length category, but the proportion of offshore hake catch based on survey proportions is negligible (averaging about 4% by weight).

   In 2017, only 20% of the total southern silver hake stock area was surveyed due to ship mechanical difficulties that led to a delay in the survey. The analyst considered several imputation procedures for use to fill in the missing data gaps, including the AMELIA II method, but in the end used a simple 3-year running average to smooth the time series, with the 2017 year dropped out for runs in which that year was missing (thus 2-year averages). The review panel believes this approach provides a workable solution to the missing data problem until something more robust can be created. Future work on the feasibility of data imputation is encouraged as missing data will likely be an issue in future assessments.

2. *Evaluate indices used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.).*

   This TOR was satisfactorily addressed. Silver hake are observed in both the fall and spring NEFSC bottom trawl surveys, but the spring survey indicates movement of the population towards the shelf edge and so only the fall survey is used as an index. A length to age binning approach is used to help characterize age compositions, but cohort-by-cohort trends are difficult to detect using these data and partly due to high predation on young of the year (Smith, B.E. 2020. Consumption estimates of red hake and silver hake at various stages for northern and southern stocks of the Northeast Continental shelf. Working paper). Nevertheless, there appears to have been some recent strong recruitment events in 2018 and 2019.

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*

    The empirical approach (running average of NEFSC fall survey means) was used, consequently no direct estimates of fishing mortality, recruitment or stock biomass were expected. When the ASAP assessment was rejected in 2017, this was effectively the Plan B assessment that has since become plan A. No plan B was prepared.

4. *Re-estimate or update the BRP's as defined by the management track level and recommend stock status. Also, provide 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.).*

    This TOR was satisfactorily addressed. Proxy biomass reference points were based on the arithmetic average of the NEFSC fall survey index (1973-1982). The exploitation reference point is based on an exploitation index calculated as the ratio of total catch to the averaged fall survey index (note the difference in units between numerator and denominator in this calculation). Given this approach it appears that the stock is not overfished and overfishing is not occurring.

5. *Conduct short-term stock projections when appropriate.*

    This is an empirical approach; no short-term projections were made.

6. *Respond to any review panel comments or SSC concerns from the most recent prior research or management track assessment.*

    There appear to be no ongoing recommendations for changes to this approach.

Additional recommendations
The basis for the existing BRP (1973-1982) should be investigated to ensure if it is still applicable to current conditions.

The NEFSC should devote effort towards developing a quantitative analytical assessment approach that can address the information content and outstanding issues associated with modeling this stock.

Additional research should be conducted to better address missing data values in the survey data time series as discussed above. For future reference, a Bayesian spatio-temporal model that uses a Conditional Autoregressive (CAR) spatial model to account for spatial correlation in fish density and

using a state-space model over time to account for temporal population dynamics implemented using R-INLA might be considered. (See for example, Blangiardo, M., and Cameletti, M. (2015). Spatial and Spatial-temporal Bayesian Models with R-INLA. Wiley, Chichester, UK. 308pp.)

Finally, if applicable, the method used here to derive biological reference points might be considered for broader application in other index-based assessments when appropriate.

Appendix A. September 2020 management track peer review meeting attendees.

Key:
ASMFC - Atlantic States Marine Fisheries Commission
NEFSC - Northeast Fisheries Science Center
NEFMC - New England Fisheries Management Council
MADMF - Massachusetts Division of Marine Fisheries
MEDMR - Maine Department of Marine Resources
SMAST - School of Marine Science and Technology, Univ. of Massachusetts, Dartmouth
GARFO - Greater Atlantic Regional Fisheries Office
NOAA - National Oceanic and Atmospheric Administration

Panel
J-J Maguire - Chair
Catherine O'Keefe - Reviewer
Richard Merrick - Reviewer
Pat Sullivan - Reviewer

Attendees and Presenters
Russ Brown - NEFSC
Michele Traver - NEFSC
Alex Hansell - MADMF
Alejandro Gonzalez
Alicia Miller - NEFSC
Andy Applegate - NEFMC Staff
Andrew Jones - NEFSC
Andrew Ray
Brian Linton - NEFSC
Brian Stock - NEFSC
Burton Shank - NEFSC
Carolina Bastidas - MIT Sea Grant
Chad Keith - NEFSC
Charles Adams - NEFSC
Charles Perretti - NEFSC
Charles Keith - NEFSC
Chris Kellogg - NEFMC Staff
Chris Legault - NEFSC
Chris Tholke - NEFSC
Dan Hennen - NEFSC
Dave McElroy - NEFSC
Dave Rudders - VIMS
Drew Minkiewicz - Kelly Drye & Warren LLP

Dustin Colson Leaning - ASFMC
Dvora Hart - NEFSC
Elizabeth Fairchild - University of New Hampshire
George Lapointe - GARFO
Georgette L -
Halle Berger - University of Connecticut
Jamie Cournane - NEFMC
Jaz Bonnin -
Jeff Kaelin - Lund's Fisheries
Jennie Rheuban - Woods Hole Sea Grant
Jennifer Couture - NEFMC
Jessica Blaylock - NEFSC
Jon Deroba - NEFSC
Jonathan Duquette - NEFSC
Jonathan Peros - NEFMC Staff
Jui-Han Chang - NEFSC
Juliet Simpson - MIT Sea Grant
Kaitlyn Clark - VIMS
Katherine Sosebee - NEFSC
Kelly Whitmore - MADMF
Kyle Molton - GARFO
Larry Alade - NEFSC
Libby Etrie - NEFMC Member
Lisa Hendrickson - NEFSC
Liz Sullivan - GARFO
Louise Cameron - Northeastern University
M Smith -
Maggie Raymond - Associated Fisheries of Maine
Mark Terceiro - NEFSC
Matthew Cieri - MEDMR
Megan Ware - MEDMR
Melissa Errend - NEFMC Staff
Michael Bergman - NEFSC
Nancy McHugh - NEFSC
Nicole Charriere - NEFSC
Pat Thames - NOAA
Paul Nitschke - NEFSC
Rebecca Peters - MEDMR
Rich Powell - NEFSC
Richard McBride - NEFSC
Robin Frede - NEFMC Staff
Samuel Asci - NEFSC
Shannah Jaburek - GARFO
Spencer Talmage - GARFO
Steve Cadrin - SMAST
Susan Wigley - NEFSC
Tara Trinko Lake - NEFSC
Tom Nies - NEFMC Executive Director
Toni Chute - NEFSC
Toni Kerns - ASMFC
Tony Wood - NEFSC

Travis Ford - GARFO
Z. Aleck Wang - Woods Hole Oceanographic Institute