Summary Report of the American Plaice Research Track Stock Assessments Peer Review

July 18-21, 2022

Report prepared by Panel members:

Yong Chen, SUNY Stony Brook, USA (Chair)

Massimiliano Cardinale, Swedish University of Agricultural Sciences

Steven Holmes, NIWA, New Zealand

Peter Stephenson, Department of Fisheries, West Australia

***Introduction***

The most recent (NEFSC 2019) assessment of the Gulf of Maine and Georges Bank American Plaice (*Hippoglossoides platessoides*; Fabricius 1780; commonly referred to as ‘dab’) stock was an operational assessment based on virtual population analysis. The 2019 assessment had the MADMF (Massachusetts Division of Marine Fisheries) inshore survey index excluded and retrospective pattern adjusted (NEFSC 2019). The 2019 assessment updated fishery catch data, research survey indices of abundance, and the VPA and reference points through 2018. Stock projections were updated through 2022. Based on this updated assessment, the stock status for Gulf of Maine and Georges Bank American Plaice was not overfished and overfishing was not occurring

Subsequently, the stock was the subject of a research track effort to identify a state-of-art stock assessment model to replace VPA and improve the quality of the assessment with work beginning in June 2021. A Working Group (WG) was created with staff from Northeast Fisheries Science Center, University of Massachusetts (UMASS) at Dartmouth, New England Fisheries Management Council, Gulf of Maine Research Institute, and the Canadian Department of Fisheries and Oceans (DFO). This 9-person WG (Chaired by Dr. Steve Cadrin, UMASS Dartmouth) met from June 2021 through June 2022 to review the literature to identify environmental drivers that might influence American Plaice stock dynamics; compile fishery, survey and life history data; configure three different modern stock assessment modeling frameworks (WHAM, SS, and ASAP); evaluate biological reference points for stock status determination; and provide short-term projections. The WG also reviewed all the research recommendations identified in previous reviews and decided if the recommendations had been addressed. The WG also evaluated possible Plan B options that can be used to provide catch advice in the case the recommended stock assessment is rejected. The Terms of Reference for the WG work on the Gulf of Maine and Georges Bank American Plaice are provided in Appendix 1.

This report summarizes the Panel’s consensus views on the peer review of the 2022 American Plaice Research Track Stock Assessment. The review Terms of Reference (TORs) are provided in Appendix 1. The final agenda for the meeting is provided in Appendix 2.

The Research Track Stock Assessment Peer Review Panel met via WebEx on July 18-21, 2022. The Panel was composed of three scientists selected by the Center for Independent Experts (CIE): Massimiliano Cardinale (Swedish University of Agricultural Sciences), Steven Holmes (NIWA, New Zealand), and Peter Stephenson (Department of Fisheries, West Australia). The Panel was chaired by Yong Chen (Stony Brook University), as a member of the New England Fishery Management Council’s Scientific and Statistical Committee. The instructions for research track peer reviewers are provided in Appendix 3 and the Performance Work Statement for CIE reviewers is provided in Appendix 4.

The Panel was assisted by Michele Traver (Chair, NEFSC’s Stock Assessment Workshop) and Russ Brown (Chief, NEFSC Population Dynamics Branch). Documentation was prepared by the American PlaiceWorking Group, and presentations were made by Steve Cadrin (WG Chair), Lisa Kerr, Jamie Behan, Paul Nitschke, Alex Hansell, Amanda Hart, Tim Miller, Dan Hennen, and Jamie Cournane. Members of the Working Group and public also provided valuable insights and discussion. Jason Boucher, Tony Wood, Russ Brown, Ben Levy, Brian Linton, Toni Chute, Laurel Smith, and Abigail Tyrell (all from the NEFSC) acted as rapporteurs throughout the meeting (see Appendix 5 for meeting attendees).

Prior to the meeting, assessment documents were made available to the Panel through the NEFSC Stock Assessment Data Portal (https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi\_report\_options.php). Panel members met with Michele Traver and Russell Brown before the meeting to review and discuss the meeting agenda, reporting requirements, meeting logistics and the overall process.

The meeting opened on 7:00 AM EST, Monday July 18, with welcoming remarks, comments on the agenda and NRCC process by Russ Brown, Michelle Traver, and Panel Chair Yong Chen. The first day of the meeting focused on presentations and discussion of the environmental effects (TOR 1), fishery data (TOR 2), and survey data (TOR 3) for the American Plaice 2022 research track assessment, the second day focused on the stock assessment model development, configuration and selection (TOR 4), the third day on the stock status determination (TOR 5), projection (TOR 6), recommendations (TOR 7) and Plan B (TOR 8). All Panel members contributed to this Summary Report, which was compiled and edited by the Panel Chair with assistance from the CIE Panelists, before submission of the report to the NEFSC. Additionally, each of the CIE Panelists will submit their separate reviewer’s reports to the Center for Independent Experts.

The scientific and statistical analyses conducted by the WG were thorough, informative, and of high quality. In addition to the American Plaice WG Report which provides detailed documentations of all the research activities, results and recommendations, the WG also provided 21 working papers which provide more details and background information for all the TORs. Their very clear and well-organized WG report, working papers, background information, and presentations made the Panel’s job much easier. The WG responded to all the questions and addressed the Panel’s requests for additional information during the review, and the whole process was transparent.

The Panel thanks all the members of the Working Group and participants in the Research Track Stock Assessment peer review for the large amount of work that went into the assessment and the collegial and constructive discussions with the Panel.

The specific comments on each ToR are provided below.

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

*The Panel concluded that this TOR has been fully addressed.*

The WG conducted a thorough and exhaustive literature review on the Gulf of Maine and Georges Bank American Plaice to identify key environmental drivers that might influence the stock dynamics. It also conducted regression analyses comparing key attributes important for stock assessment: recruitment, growth; with those potential environmental drivers.

The WG concluded that depth distribution of American plaice is influenced by season and temperature. The distribution of plaice shifts from deeper water in winter to more shallow water in spring/summer. Warming temperatures, however, appear to be causing a general shift to deeper waters and a possible contraction of preferred habitat.

Recruitment per spawner (R/SSB) seems related to temperature. Highest values of R/SSB coincide with years of extreme cold temperature. However, regression analysis of R/SSB against the temperature anomaly of the North Atlantic Oscillation (NAO) found a positive relationship between R/SSB and temperature.

Literature indicates warming temperatures should be associated with accelerated growth rates, earlier ages at maturity and reductions in body size of post maturity. Warmer temperatures are also expected to lead to higher rates of natural mortality, particularly at younger ages.

The WG identified southward shifting in the 1980s and 1990s, then northward in the 2000s and 2010s. Relationships were found to be related to changes in stock biomass and bottom temperature, but mechanistic relationships are not yet identified. This shift in distribution is likely to affect the availability of American Plaice to the surveys, in particular to the two inshore state surveys.

Fish weight at age, particularly in the fall NEFSC survey, went through a period of decline starting in the late 1990s and continuing through to the mid/late 2000s before stabilizing in the last decade.. The decline was likely caused by change in growth. Regression analysis suggested a negative relationship between condition and Atlantic Multidecadal Oscillation (AMO) temperature anomaly, but a positive relationship with bottom temperature anomaly.

The Panel commended the WG efforts to identify key environmental drivers and use the synthesized information to advise the work defined in other TORs.

The Panel recommends continued monitoring of phonological changes in movement, spawning and other life history processes given anticipated continuing warming temperature in the region.

The Panel recommends to derive age and time varying M based on estimated changes in size at age over time to be used in the assessment model.

The Panel also recommends to randomly draw historical Recruit per spawner in the projection (and estimation of biological reference points). The WG showed evidence suggesting a relationship between the stock-recruit relationship and water temperature. On the basis recruits per spawner should implicitly account for environmental influences, it would be useful, therefore, to test the performance of the assessment model using this metric rather than numbers recruited.

More research is needed to better understand relationships between environmental changes (trend and variability) and both recruitment dynamics and fish growth/condition. In particular, the seemingly contradictory results for the influence of bottom temperature and the Atlantic Multidecadal Oscillation (AMO) need to be resolved. The AMO can be measured in terms of temperature anomaly but its influences on plaice may be more indirect, e.g. through altered prey availability, altered larval transport or some other mechanism.

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

*The Panel concluded that this TOR has been fully addressed.*

The commercial catch (both landings and discards) appears to be well estimated. The assumed 100% discard mortality is justified by field experiments and observations. The landing age composition data show good cohort-tracking, which is consistent with that of the survey. The Panel considers estimating discard age composition using the survey age-length key applicable here because most discards are small plaice that are well selected in the survey. However, selectivity is different between surveys and fleets so that ideally the ageing data derived directly from discards would be more appropriate.

Landings per unit of effort (LPUE) was standardized by the WG to develop a possible fishery-dependent stock size index. The standardization considered fishing location, season, vessel tonnage, depth, and price. Bottom temperature was not considered. The standardized LPUE was found to have a similar temporal trend to the NEFSC survey indices. An analysis of LPUE vs Model-estimated SSB (WHAM 29F4 base run) was done during the review. This analysis shows a weak hyperstability pattern. The standardized LPUE was explored in the WHAM, but was not selected in the base run based on the prescribed model selection protocol developed by the WG.

The Panel recommends continued development of fishery logbook reporting and development of a fishery-dependent abundance index which can be used in the assessment process. A standardized Landings Per Unit Effort (LPUE) could possibly include bottom temperature or SST in the standardization.

The Panel recommends continuing development of the Electronic Monitoring program and at-sea observer monitoring to quantify the discards.

The Panel suggests the otoliths collected by at-sea observers be processed as these would complement the sometimes small numbers of otoliths collected in the factory otolith sampling.

The Panel recognizes that the otolith reading procedures adopted for this species are current best practice, including excellent documentation of procedures, rolling reference samples, and reconciliation of results. Despite this, the reader agreement is not good. The Panel recommends that current procedures be continued. In addition, the age-length keys derived from the fleet landings and surveys should be routinely compared and explanations sought for any observed differences (like spatial coverage, length of tow, gear type).

The assessment models use catches from 1960 but there are recorded catches of American Plaice prior to this date. The Panel recommends that historical catches be reconstructed together with any size and age data or survey information.

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

*The Panel concluded that this TOR has been fully addressed.*

The WG explored different ways to take account for the changes from survey vessel Albatross to Bigelow in the NEFSC fall and spring surveys including having a single time series with calibration and also considering two separate surveys. The model analyses suggest that the Albatross and Bigelow survey data should be treated as two separate time series for better diagnostics and model performance.

Two state surveys, MADMF and ME-NH conducted in both spring and fall, were evaluated by the WG. These two state surveys have no age-length key. The GOM-GB American plaice stock has been found (both by fishermen and in the analysis of literature and survey data) to shift to offshore in deeper waters in recent years. This makes the two inshore state surveys less reliable in indexing American plaice stock dynamics in recent years. The MADMF survey was removed from the 2019 management track stock assessment.

The NEFSC survey age composition data show good cohort tracking and the strong year classes identified in the NEFSC surveys are the same as those suggested in the fishery catch-at-age data (i.e., Year classes 1987, 1993, 2004, and 2013), suggesting that the NEFSC surveys can capture the stock dynamics well.

A spatiotemporal model, Vector Autoregressive Spatio-Temporal (VAST), was used to develop model-based survey indices, which can integrate all three surveys (NEFSC, MADMF and ME-NH surveys). VAST indices were dominated by the NEFSC surveys because they cover much larger areas than the inshore surveys. Using the VAST model, the WG conducted a counterfactual analysis to evaluate impacts of different environmental variables on abundance indices for improved understanding of interactions of environmental variables and abundance indices.

The Panel recommends continuing exploring VAST to include federal and state surveys for the development of integrated abundance indices that cover the entire stock area. The Panel also recommends monitoring changes in the effective area occupied to test the hypothesis of fish moving out of the state survey area. Effective area occupied should be considered alongside spatial maps showing the modeled fish density by year and a map of bottom depth. This would potentially maximize insight into how decreases (increases) in the area occupied metric translate to changes in density in different regions of the stock area.

The Panel notes that the data suggest old and large plaice are not fully selected by the survey, but tend to be fully selected in the fishery. The explanation for such a difference in selectivity for old/large American plaice remains unknown.

The Panel recommends an analysis of the fleet and survey data at a higher spatial resolution to evaluate possible differences in stock availability and/or gear catchability between the survey and fishery and the effect of not accounting for spatial considerations when combining ALK and length compositions to derive numbers at age in the catch and in the survey.

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

*The Panel concluded that this TOR has been fully addressed.*

The Panel considers that the data used within the presented assessment models are generally appropriate and data uncertainty sufficiently acknowledged. The models used to conduct the data preparation for the assessments are suitable for the available data as well as the data series are adequate to support the assessment models used. The choice of the various analytical tools used to derive the data is well justified in the background documents presented. The models (i.e., WHAM, Stock Synthesis, VPA and ASAP) used to assess the American plaice stock are appropriate, robust and in general properly configured, and in line with standard practices.

The Panel commends the WG effort to explore a large number of runs with different configurations and assumptions, and to employ different modeling platforms to evaluate the robustness of the WHAM primary stock assessment model.

However, the process of selection of the final model is difficult to follow and several aspects of the process that leads to the “best case” model configuration used to provide advice, as described in the American Plaice Research Track Working Group report, can be improved. The Panel considers that the assessment Team should decide a priori the criteria for model selection, present them upfront, i.e. they could be clearly listed at the beginning of the model selection process. Those criteria should be used in combination (and not in isolation) to compare and select models and referred to each time a model is selected or discarded, including navigating between different model configurations and their pruning. Criteria for model selection and pruning should not be based on derived quantities such as SSB or reference points but should be centered on diagnostics that allow comparison between models with different weights of the model components and different data sources. Thus, AIC is not recommended, while Mohn´s rho, quantitative analysis of residuals and MASE should be preferred.

Objective criteria as above should be augmented by first principles. First principles are particularly useful to build base case scenarios from which model exploration should be derived. Establishing a base scenario would also facilitate navigation between the different model configurations by external readers. Alternative model configurations should be based on hypothesis testing well defined for each model.

Following comments from the Panel, the WG evaluated their model selection process and provided during the review a protocol (Appendix 6) the WG used for selecting the base run which could be further developed and used in future stock assessments.

Several WHAM model configurations and even different model platforms achieve comparable performances in the terms of model diagnostics but different stock status in terms of depletion (e.g., SS\_Model\_Run\_BASE14fixFleet). Thus, the Panel recommends that in the future, an ensemble of different plausible configurations and model platforms be selected and their performance evaluated using a comprehensive diagnostic criteria agreed beforehand for American plaice. This is particularly important, in the context of providing probabilistic statements of stock status, given the uncertainties in the data used as input, the key biological parameters, and environmental processes.

The Panel recommends further work to be conducted with the aim to achieve the integration of key environmental covariates in the stock assessment models;

Also, the Panel recommends that further work is conducted to use VAST based survey indices in the stock assessment models.

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

*The Panel concluded that this TOR has been fully addressed.*

The WG examined possible stock recruitment relationships and found that the recruitment is negatively related to SSB and no recruitment model can be developed, which justifies the use of SPR0 as MSY proxies. Like other groundfish species in this region, FSPR40% and SSBFSPR40% are used as the proxies for FMSY and BMSY. Based on a careful evaluation of changes in recruitment and other biological and fishery processes for this fishery, the WG chose to use recruitment from the entire time series and the most recent 5-year estimates of selectivity and observed weight at age to calculate SSBF40%. The Panel considered this approach is technically and biologically sound and supported the WG decisions for the estimation of SDC.

The WG estimated F40% being 0.43 and SSBF40% 18,000 mt. The stock is not overfished and overfishing is not occurring. The Panel supports the conclusion.

The Panel notes that stock status would be consistent among the models and among different WHAM configurations, suggesting that the stock status determination conclusion is robust.

The Panel recommends to investigate the use of R/SSB, instead of recruitment, in the calculation of SSBF40%, on the grounds this may remove some possible density-dependent influence on recruitment. This is also consistent with the finding in TOR 1 that R/SSB is significantly related to bottom temperature. These values would still be randomly drawn from the whole time series.

The Panel notes that the final WHAM model does not assume a SR relationship. Although this is considered not relevant at the current conditions, a SR would be needed for forward projections as MSE and for deriving limiting reference points. Also, an SR is important when the stock is depleted and a recovery plan is envisaged. In terms of the SPR target levels and how FSPR relates to FMSY, for SPR fraction = 0.4 FSPR exceeds FMSY at steepness levels below 0.65. Thus, given the assumed best estimate of steepness being less than 0.65, there are some risks associated with an FSPR40%.

Given the large change of bottom temperature in the stock area and its potential impacts on American plaice life history processes, the Panel recommends that dynamic BRPs with consideration of environmental covariates be explored in future research.

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

*The Panel concluded that this TOR has been fully addressed.*

The WG used the WHAM Run 29F4 model configuration to produce integrated projections accounting for uncertainty in all the estimated parameters, 2020 abundance at age and future recruitment, including uncorrelated process variance in survival and recruitment. The short-term projection was made based on the entire time series of recruitment, recent 5-year selectivity and observed weight at age, and maturity at age from the entire time series. This is consistent with the revised BRP calculation in TOR 5. The WG made the projection through 2022 for 4 scenarios (i.e., F at F40% OFL, 75%F40%=ABC, F2019=status quo, and F =0). Because there is little retrospective pattern in the WHAM 29F4 run, no retrospective adjustment is necessary.

The Panel supports the projection made by the WG. Since the WHAM 29F4 model configuration is likely to be used in the 2022 management track assessment to provide catch advice, the Panel raised an issue of possible impact of missing 2020 survey data when the 2020 and 2021 data are updated in the upcoming management track assessment. Such an impact should be carefully evaluated.

The Panel recommends close monitoring of the recruitment dynamics in the future given changing thermal habitat and its potential impacts on recruitment. The Panel also noted that using the whole time period of recruitment may under-estimate recruitment in the projection because of increased warming environment for the stock and a negative relationship between R and temperature.

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

*The Panel concluded that this TOR has been fully addressed.*

The Panel commends the WG for an excellent and exhaustive examination of all previous research recommendations. The WG concluded that all the previous research commendations had been addressed except ageing samples for the MADMF inshore survey.

The WG developed the following new research recommendations listed in their priority:

1. Continue to monitor shifts in distributions of plaice, particularly depth and environmental covariates on catchability.
2. Exploration of spatiotemporal integration of federal and state surveys should continue.
3. Investments are needed to streamline the estimation of commercial catch and promote reproducibility of estimates.
4. Consider deriving discards from electronic monitoring when an integrated catch monitoring system is developed.
5. As the Gulf of Maine scallop fishery expands, it should be included in discard estimation.
6. Archived otolith samples should be processed (state surveys, at-sea observers, 1975-1979).
7. The relationship between recruitment and ocean temperature should continue to be monitored.
8. Time-varying natural mortality, possibly with environmental covariate should be explored.
9. Methods should be developed to compare models with and without environmental covariates.
10. If the proposed assessment approach does not meet the standards of peer review, an alternative model should be developed to integrate information from catch, age composition and indices.

The Panel supports these new research recommendations developed by the WG. Additional research recommendations suggested by the Panel include:

1. Conduct further work to achieve the integration of key environmental covariates in the stock assessment models.
2. Explore an ensemble modeling approach to incorporate different plausible configurations and model platforms selected and weighed by a comprehensive diagnostic against performance criteria agreed beforehand to provide stocks status and management advice for American plaice.
3. Continue developing LPUE which can be used in the assessment process, possibly including bottom temperature in LPUE standardization.
4. Continue developing Electronic Monitoring program and Observer monitoring to quantify the discards and biological information.
5. Explore dynamic BRPs with consideration of environmental covariates given the large change of thermal habitat in the stock area and its potential impacts on American plaice life history processes.
6. Continue developing built-in diagnostic tools for WHAM (e.g., plots of MASE).
7. Consider developing an age- or size-dependent M. The current value used is a mean value over all ages/sizes but it is highly likely that M is greatest on the youngest fish. Miss-specification of M by size may lead to biased estimates of selectivity and hence BRPs. One common approach is to scale the Lorenzen weight based Ms to the overall mean derived from meta-analyses.
8. Further work on the development of VAST based survey indices in the stock assessment models.

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

*The Panel concluded that this TOR has been fully addressed.*

The WG explored various approaches for a Plan B to provide scientific advice to managers if the proposed assessment approach is rejected. The WG concluded that index-based methods, which are common fell-back assessment methods in this region when the formal quantitative stock assessments are rejected, are not suitable for GOM-GB American plaice because older plaice are not fully selected in the surveys and fishery catch is not directly comparable to survey biomass estimates for a relative exploitation rate estimate.

The WG recommends that ASAP (Run 43) for Plan B for the future management track assessment. The Panel supports the WG recommendation to use ASAP (Run 43) as the Plan B approach for the upcoming management track assessment. However, for future management track assessments, the Panel recommends that other model-based stock assessment be further explored (e.g., different configurations for WHAM) if the proposed assessment is rejected.

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

This TOR is not applicable.

***Appendix 1. Terms of reference for the 2022 American plaice research track stock assessment.***

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

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

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

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

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

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

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

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

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

***Appendix 2. Final agenda for the American plaice research track stock assessment peer review meeting.***

**American Plaice**

**Research Track Assessment Peer Review Meeting**

**July 18-21, 2022**

WebEx link:

<https://noaanmfs-meets.webex.com/noaanmfs-meets/j.php?MTID=mf97d5121d96d26f36e88243f0dd9e013>

Meeting number (access code): 2763 669 5649

Meeting password: mP4vVXESd74

JOIN BY PHONE

+1-415-527-5035 US Toll

**DRAFT AGENDA\* (v. 7/12/2022)**

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

Monday, July 18, 2022

| **Time** | **Topic** | **Presenter(s)** | **Notes** |
| --- | --- | --- | --- |
| 7 a.m. - 7:15 a.m. | Welcome/Logistics  Introductions/Agenda/Conduct of Meeting | Michele Traver, Assessment Process Lead  Russ Brown, PopDy Branch Chief  Yong Chen, Panel Chair |  |
| 7:15 a.m. - 7:45 a.m. | Introduction and Overview | Steve Cadrin |  |
| 7:45 a.m. - 8:45 a.m. | TOR #1 | Lisa Kerr and Jamie Behan | Environmental Effects |
| 8:45 a.m. - 9 a.m. | Break |  |  |
| 9 a.m. - 11 a.m. | TOR #2 | Steve Cadrin | Fishery Data |
| 11 a.m. - 11:15 a.m. | Break |  |  |
| 11:15 a.m. - 12:15 p.m. | TOR #3 | Paul Nitschke and Alex Hansell | Survey Data |
| 12:15 p.m. - 12:30 p.m. | Summary/Discussion | Review Panel |  |
| 12:30 p.m. - 12:45 p.m. | Public Comment | Public |  |
| 12:45 p.m. | Adjourn |  |  |

Tuesday, July 19, 2022

| **Time** | **Topic** | **Presenter(s)** | **Notes** |
| --- | --- | --- | --- |
| 7 a.m. - 7:05 a.m. | Welcome/Logistics | Michele Traver, Assessment Process Lead  Yong Chen, Panel Chair |  |
| 7:05 a.m. - 8 a.m. | TOR #3 cont. | Paul Nitschke and Alex Hansell | Survey Data |
| 8 a.m. - 9:30 a.m. | TOR #4 | Amanda Hart, Tim Miller, Steve Cadrin, Dan Hennen and Alex Hansell | Assessment Models |
| 9:30 a.m. - 9:45 a.m. | Break |  |  |
| 9:45 a.m. - 11:45 a.m. | TOR #4 cont. | Amanda Hart, Tim Miller, Steve Cadrin, Dan Hennen and Alex Hansell | Assessment Models |
| 11:45 a.m. - 12 p.m. | Break |  |  |
| 12 p.m. - 12:30 p.m. | TOR #4 cont. | Amanda Hart, Tim Miller, Steve Cadrin, Dan Hennen and Alex Hansell | Assessment Models |
| 12:30 p.m. - 12:45 p.m. | Summary/Discussion | Review Panel |  |
| 12:45 p.m. - 1 p.m. | Public Comment | Public |  |
| 1 p.m. | Adjourn |  |  |

Wednesday, July 20, 2022

| **Time** | **Topic** | **Presenter(s)** | **Notes** |
| --- | --- | --- | --- |
| 7 a.m. - 7:05 a.m. | Welcome/Logistics | Michele Traver, Assessment Process Lead  Yong Chen, Panel Chair |  |
| 7:05 a.m. - 8 a.m. | TORs # 5 and #6 | Steve Cadrin  Paul Nitschke and Jamie Cournane | Reference Points  Projections |
| 8 a.m. - 9 a.m. | TOR # 7 | Steve Cadrin | Research Recommendations |
| 9 a.m. - 9:15 a.m. | Break |  |  |
| 9:15 a.m. - 10:45 a.m. | TOR #8 and Near Term Plans | Steve Cadrin | Alternative Assessment Approach |
| 10:45 a.m. - 11 a.m. | Break |  |  |
| 11 a.m. - 11:15 a.m. | Summary/Discussion | Review Panel |  |
| 11:15 a.m. - 11:30 a.m. | Public Comment | Public |  |
| 11:30 a.m. - 12 p.m. | Key Points/Follow ups/Panel Wrap ups | Review Panel |  |
| 12 p.m. | Adjourn |  |  |

Thursday, July 21, 2022

| Time | **Topic** | **Presenter(s)** | **Notes** |
| --- | --- | --- | --- |
| 7 a.m. - 12 p.m. | Report Writing | Review Panel |  |

***Appendix 3. Instructions to research track peer reviewers.***

(Based on: 2011 Generic Operational Assessment Process White Paper, 2011, Description of New England and Mid-Atlantic Region Stock Assessment Process, 2018, and NEFSC edits. v.01/24/2022)

The Peer Review is to determine whether the completed research track assessment is technically sufficient to (a) evaluate stock status determination, (b) evaluate new data streams and/or model changes and (c) successfully address the assessment Terms of Reference. The Peer Review Panel may determine that the proposed research track assessment approach has not worked; if so, the alternative backup approach to the assessment will be reviewed.

Conduct of the Meeting:

* The Meeting Chair is responsible for:

○ the conduct of the meeting

○ ensuring that meeting participants are provided with opportunities to provide input

○ ensuring that the Peer Review panel drafts a report that evaluates stock status determination, evaluates new data streams and/or model changes and whether the assessment Terms of Reference are addressed.

* The Peer Review meeting is a public meeting and opportunities are provided for input from participating scientists and interested stakeholders. The Meeting Chair is responsible for providing opportunities for public participation, while balancing the need for the reviewers to complete their work
* The Assessment Process Lead (currently Michele Traver) and the Branch Chief of the Population Dynamics Branch (currently Russell Brown) are staff who can provide guidance relative to process or policy related questions. These staff are also responsible for meeting logistics and support.
* The Northeast Fisheries Science Center provides a rapporteur who will take meeting notes. Meeting notes are a reference for use by the Peer Review panel for their deliberations and report writing but have no official standing and are not included in the meeting reports.
* The panel has the option to meet privately (without other participants) for short periods of time to discuss issues related to their review. Private meetings should include all panel members including the Chair and may or may not include the Assessment Process Lead and/or the Population Dynamics Branch Chief at the discretion of the panel. Private meetings are announced to meeting participants in the interest of transparency.

Report Guidance:

For each stock assessment, the assessment reports:

* should address whether each stock assessment TOR was completed successfully
* should make clear whether the proposed assessment approach was accepted, or whether the backup assessment approach was recommended
* should identify major sources of uncertainty in the stock assessment and comment on the qualitative descriptions of stock status based on simple indicators/metrics in the assessment report
* can also make recommendations for improving the assessment in the future, which could be considered in determining stocks or topics for future research track assessments. It is helpful if these research recommendations are prioritized.
* If the panel rejects the model presented, please indicate the reason why and make recommendations.

***Appendix 4. Performance Work Statement for CIE reviewers for the American plaice research track stock assessments.***

Performance Work Statement (PWS)

National Oceanic and Atmospheric Administration (NOAA)

National Marine Fisheries Service (NMFS)

Center for Independent Experts (CIE) Program

External Independent Peer Review

*American plaice Research Track Peer Review*

Background

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

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

Scope

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

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

Requirements

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

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

Tasks for Reviewers

● Review the background materials and reports prior to the review meeting

o Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.

● Attend and participate in the panel review meeting

o The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers

● Reviewers shall conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.

● Each reviewer shall assist the Peer Review Panel (co)Chair with contributions to the Peer Reviewer Summary Report

● Deliver individual Independent Reviewer Reports to the Government according to the specified milestone dates

● This report should explain whether each research track Term of Reference was or was not completed successfully during the peer review meeting, using the criteria specified below in the “Tasks for Peer Review Panel.”

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

● During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.

● The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

● During the peer review meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the Peer Review Panel chair shall identify or facilitate agreement among the reviewers for each research track TOR.

● If the panel rejects any of the current BRP or BRP proxies (for BMSY and FMSY and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.

● Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for Peer Review Panel chair and reviewers combined:

Review the Report of American plaice Research Track Working Group.

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

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

Foreign National Security Clearance

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

Place of Performance

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

Period of Performance

The period of performance shall be from the time of award through August 5, 2022. Each reviewer’s duties shall not exceed 14 days to complete all required tasks.

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

|  |  |
| --- | --- |
| Within 2 weeks of award | Contractor selects and confirms reviewers |
| Approximately 2 weeks later | Contractor provides the pre-review documents to the reviewers |
| July 18-21, 2022 | Panel review meeting |
| Approximately 2 weeks later | Contractor receives draft reports |
| Within 2 weeks of receiving draft reports | Contractor submits final reports to the Government |

\* The Peer Reviewer Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

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

# ● Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact

Michele Traver, NEFSC Assessment Process Lead

Northeast Fisheries Science Center

166 Water Street, Woods Hole, MA 02543

Michele.Traver@noaa.gov

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

***Appendix 5. Attendees for July 18-21 American plaice research track peer review meeting.***

**American Plaice Research Track Peer Review Attendance**

**July 18-21, 2022**

NEFSC - Northeast Fisheries Science Center

GARFO - Greater Atlantic Regional Fisheries Office

NEFMC - New England Fisheries Management Council

SMAST - University of Massachusetts School of Marine Science and Technology

GMRI - Gulf of Maine Research Institute

MADMF - Massachusetts Division of Marine Fisheries

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

*Yong Chen - Chair*

*Steven Holmes - CIE Panel*

*Peter Stephenson - CIE Panel*

*Massimiliano Cardinale - CIE Panel*

Russ Brown - NEFSC, Population Dynamics Branch Chief

Michele Traver - NEFSC, Assessment Process Lead

Alex Dunn - NEFSC

Alex Hansell - NEFSC

Alicia Miller - NEFSC

Amanda Hart - SMAST

Angela Forristall - NEFMC Staff

Charles Adams - NEFSC

Charles Perretti - NEFSC

Chris Kellogg - NEMFC Staff

Cole Carrano - SMAST

Dan Hennen - NEFSC

David McCarron - MADMF (retired)

Jackie ODell - Executive Director of Northeast Seafood Coalition

Jamie Behan - GMRI

Jamie Cournane - NEFMC Staff

Jason Boucher - NEFSC

Kathy Sosebee - NEFSC

Libby Etrie - NEFMC Member

Lisa Kerr - GMRI

Mark Alexander - Asst. Director (retired) of the Fisheries Division, Connecticut Dept. of Energy & Environmental Protection

Mark Terceiro - NEFSC

Max Grezlik - SMAST

Paul Nitschke - NEFSC

Robin Frede - NEFMC Staff

Steve Cadrin - SMAST

Tim Miller - NEFSC

Tony Wood - NEFSC

**Appendix 6. Model Selection Procedure for American Plaice Research Track 2022**

The plaice WG considered WHAM model variants (alternative ‘runs’) based on structural aspects of greatest relevance to plaice, as identified by Terms of Reference or recommendations from previous assessments. In particular, environmental effects (ToR1), index selection (e.g., 2019 decision to exclude MADMF index and recommendation to consider separate Albatross and Bigelow indices; NEFMC 2020 recommendation to consider fishery CPUE), earlier start year (NEFSC 2002 recommendation), time varying selectivity, various random effect structures, and different age composition likelihoods.

1. Initial model acceptance initially focused on the requirement that the model converge on a solution. This convergence criterion eliminated some of the variants considered (e.g., notably run 37E with estimation of selectivity at age for multi-survey VAST stock indices did converge).
2. Model validation then focused on residual analyses. Non-random residual patterns for the inshore state surveys and calibrated Albatross-Bigelow series were used to justify excluding state surveys and splitting Albatross and Bigelow surveys as separate indices. In addition to conventional residuals, one-step ahead residuals were also used to judge model fit.
3. AIC was used to compare candidate models that were fit to the same data, fit the data well, and assumed the same statistical distributions and therefore had comparable likelihoods. AIC was similar among candidate runs but lowest the run used for status determination and projections (29F-4).
4. We examined retrospective patterns for all candidate model runs and measured retrospective inconsistency as Mohn’s rho for spawning stock biomass and fully selected fishing mortality. All runs using the revised natural mortality assumption had similarly high retrospective consistency (rho<0.1).
5. We evaluated prediction skill of all candidate model runs using error of forecast values. Mean absolute scaled error (MASE) was similar among candidate runs but was lowest for the run used for status determination and projections (29F-4).
6. Self-tests were conducted on the three candidate runs (29F2, 29F4, 29F5). Candidate runs performed similarly in self-tests.

With criteria 1-6 generally being similar among the three candidate runs, run 29F4 was selected to present results, status determination and short-term projections, because it had the best retrospective consistency, AIC, prediction skill, and estimation performance for spawning stock biomass with 100% convergence in self-tests. In summary, the WG conducted model selection initially using traditional convergence and residual diagnostics for age-based assessments as well as some more recently developed diagnostics to determine the three candidate runs and the run selected for status determination and projections.