2024 June Management Track Peer Review Panel Report

John Wiedenmann (Chair, Week 1)[1], Kai Lorenzen[2], Alexei Sharov[3]
Paul Rago (Chair, Week 2)[4], Ed Camp[2], Yan Jiao[5], Katie Drew[6]

[1]Rutgers University
[2]University of Florida
[3]Maryland Department of Natural Resources
[4]NOAA Fisheries, retired
[5]Virginia Tech
[6]Atlantic States Marine Fisheries Commission

Executive Summary

Eight stock assessments were reviewed by the June 2024 Management Track peer review panels. Three of these were Level 2 Expedited Reviews: Atlantic Surfclam (Spisula solidissima), Black Sea Bass (Centropristis striata), and Butterfish (Peprilus triacanthus). The remaining five assessments were Level 3 Enhanced Reviews: Golden Tilefish (Lopholatilus chamaeleonticeps), Western Gulf of Maine Atlantic Cod (Gadus morhua), Eastern Gulf of Maine Atlantic Cod (G. morhua), Georges Bank Atlantic Cod (G. morhua), and Southern New England Atlantic Cod (G. morhua). Levels of review were as recommended by the Assessment Oversight Panel (Appendix A).

The Peer Review Panel (Panel) for Week 1 of the June 2024 Management Track Assessments met via webinar on June 18, and then in person in Woods Hole, MA with hybrid capabilities on June 20-21. The Week 2 Panel was conducted in person in Woods Hole, MA, although not all members of the Peer Review Panel could attend in person. The Panel was to determine whether the completed Management Track Assessment was technically sufficient to a) evaluate stock status, b) provide scientific advice, and c) successfully address the assessment Terms of Reference (Appendix C). Table 1 presents a list of the stocks, name of the lead analyst/presenters, and conclusions about stock status. Attendance at the meeting is provided in Appendix B with the Agenda shown in Appendix E.

The Panel thanks Kristan Blackhart (Population Dynamics Branch Chief) and Michele Traver (Assessment Process Lead) for their support during the meeting and expresses its appreciation to all of the assessment scientists who participated in the review. The presentations were uniformly excellent and responsive to prior suggestions from the Research Track, Management Track, and SSC recommendations. Extensive prior preparation allowed assessment scientists to be responsive to questions from the Panel. We also would like to thank the rapporteurs for taking extensive notes and faithfully capturing the lively discussions following each presentation.

Atlantic Surfclam

The Panel concluded that the TORs had been met and the assessment is sufficient for providing catch advice. The Panel discussed the time-varying growth parameters and whether it was appropriate to assume similar trends moving forward, but did not recommend any changes in time-varying growth for the next MT assessment. The Panel also discussed the possible benefit of estimating a time-varying natural mortality, and recommended this be investigated in the next MT assessment. Catchability may be either hyperstable or hyperdeplete, but the Panel did not make any recommendations regarding this for the next MT assessment. Finally, the Panel recommended developing figures to illustrate the effects of interactions between historically sustainable fishing mortality values and potentially declining natural mortality on future projections, pending capacity to estimate time-varying natural mortality.

Black Sea Bass

The Panel concluded that the TORs had been met and the assessment is sufficient for providing catch advice. This assessment is the first to implement the many changes introduced in the recently completed Research Track. Changes include use of a single model that allows for migration between northern and southern areas, estimation of fleet specific fishing mortality in both areas, and use of an integrated measure for relative abundance from the VAST model. The new projection methodology in WHAM accounts for the covariance structure of model parameters and is thought to provide a more valid characterization of future uncertainty. The stock is projected to remain above the overfished threshold throughout the short-term projection period. The estimated catch at Fmsy for 2025 is 6,193 mt and 4,916 mt for 2026. Weight-at-age of ages 2+ have shown decreasing trends in both the northern and southern regions since the mid-2010s. This is a concerning development, and further exploration of the biological reasons and potential impacts on population projections is recommended.

For the next MT the Panel recommended further exploration of the VAST model for age 1 indices, investigation of an age-dependent M and further consideration of using an informative prior for M. Longer term recommendations include the possibility of using a YOY index, and consideration of the effects of temperature on survival between age 0 and age 1. Potential relationships between parameterization of migration in the model and the derived VAST indices should be investigated, particularly with respect to potential dependencies on water temperature. Such refinements may improve the ability to forecast future recruitment.

Butterfish

The Panel concluded that the TORs had been met and the assessment is sufficient for providing catch advice. As in the previous MTA—the stock is at a high level of abundance (222% of SSB threshold) and lightly fished (3% of $F_{threshold}$). Updated biological reference points were slightly lower but not significantly different from previous MT estimates. The proposed change in the basis for the short-term projections was approved by the Panel and had little impact on the forecasted stock abundances and catch levels.

Several aspects of the current model and associated forecasts concerned the Panel. The F_{MSY} proxy, estimated at 5.1 in this assessment, results in extremely high estimates of projected catches of 60,475 mt in 2025 and 38,490 mt in 2026. These values are more than seven times the highest observed catch since the directed fishery reopened and comparable to estimates of SSB in 2024 (71,230 mt) and 2025 (39,203 mt). The high F_{MSY} proxy and high projected catch levels are due to high values of natural mortality, rapid maturation (60% mature by age-1) and delayed vulnerability to the fishery at age-3. Recruitment is assumed to be independent of SSB. Collectively these features reduce the impact of the fishery on stock dynamics.

The high value of F_{MSY} proxy is largely driven by the high value of M. However, the M estimate is dependent on the fixing of catchability at 0.2 in the earlier ASAP model based on selectivity experiments and definition of thermal habitat. A future assessment should explore whether a more traditional estimate of M based on longevity, life history, or weight-at-age would allow the model to fit q and estimate scale more plausibly. The next Management Track assessment for butterfish should reconsider the basis of current reference points for butterfish. Reference points based on total biomass instead of spawning stock biomass and forage-specific reference points using natural mortality should be considered.

Golden Tilefish

The Panel concluded that the TORs had been met and the assessment is sufficient for providing catch advice. A new assessment model based on WHAM, reviewed in the March 2024 Research Track Assessment, was applied in this assessment. Stock status has changed from the previous assessment; overfishing is now occurring. The final model selection process was thorough, well-reasoned and supported by model diagnostics.

The biggest source of uncertainty is the inclusion of random effects in the WHAM model. But this sensitivity cannot be isolated from the effects of the change in selectivity within the model beginning in 1987 where the model includes a sharp transition from a relatively flat-topped selectivity to a more dome-shaped selectivity. This feature, carried over from the ASAP model, implies a rapid change in fishery behavior and induces a large pool of "cryptic biomass" of fish that are not available to the current fishery. Independent empirical evidence for domed selectivity is lacking for Golden Tilefish. Assembling evidence about fishery history (spatial distribution, hook type and size, bait etc.) and more detailed analyses of recent pilot surveys should be a high priority in the near term. The fishery is largely supported by fish less than 10 years old, and periodic year classes appear to be reflected in the catch and CPUE analyses. The fishery is vulnerable if these year classes are in fact, products of newly mature fish rather than a larger, more stable cryptic spawning stock biomass.

The historical performance of the fishery with almost 25 years of nearly constant quotas is important to note. Frequent assessments, economic factors, individual quotas, and management interventions appear to have stabilized landings over this period and provided economic benefits. This does not diminish the risk of depending on episodic recruitment to support the fishery, and the potential rapid decline of spawning stock biomass if the dome-shaped selectivity pattern is not correct.

Western Gulf of Maine (WGOM) Atlantic Cod

The Panel concluded all of the ToRs have been met for this stock and the results are suitable for management advice. Based on the model results, the stock is overfished and overfishing is occurring. The Panel considered the assessment results in the most recent years and their extreme sensitivity to the 2023 spring BLLS data at length (see details in ToR 3 for WGOM cod below), and ultimately recommended a model that excluded this data point. This exclusion resulted in different terminal SSB and F estimates, but did not change the status determination. There was also discussion

about the large differences in the WHAM projections to calculate the SSB_{MSY} by fishing at $F_{40\%}$ (see ToR 4). Without clear knowledge of what was driving the different results in the projections, the Panel was not comfortable endorsing the proposed projection estimate of SSB_{MSY}., and instead approved the SPR 40% proxy approach to estimate SSB_{MSY}.

Eastern Gulf of Maine (EGOM) Atlantic Cod

The Panel concluded all but one of the ToRs have been fully met for this stock and the results are suitable for management advice. Based on the model results, the stock is overfished and overfishing is not occurring. The MT assessment included new discard estimates from a number of fisheries for the most recent four years from CAMS, and these estimates were dominated by discards in the lobster fishery. The Panel had serious concerns over the magnitude and representativeness of these estimates, and ultimately recommended a model that excluded the estimated discards from the lobster fishery (see ToR 1 and 3 for EGOM cod for more details). While the Panel acknowledged that this model was flawed due to the lack of a known source of mortality, they felt it was more appropriate than including only four years of likely overestimated discards. The Panel felt that this option was best because it kept the door open for the model to be updated in the next MT assessment with a longer time series of discard estimates from ongoing research that could not be completed in time for this assessment.

Georges Bank (GB) Atlantic Cod

The Panel concluded all of the ToRs have been met for this stock and the results are suitable for management advice. Based on the model results, the stock is overfished, and overfishing is not occurring. The Panel discussed the unusual selectivity pattern in the NEFSC fall survey (full selection of the youngest ages), without which the model would not converge. The Panel recommended exploration of the cause of this issue in future MT assessments, as well as splitting the NEFSC survey based on the Albatross and Bigelow time series.

Southern New England (SNE) Atlantic Cod

The Panel concluded all of the ToRs have been met for this stock and the results are suitable for management advice. Based on the model results, the stock is overfished

and overfishing is occurring. The Panel had concerns about the assumption that catch in the bridge year for projections (2024) was equal to catch in the previous year. Catches for this stock are primarily from the recreational fishery, which have greater uncertainty in their precision. Therefore, the Panel recommended that the most recent three-year average F be used in the bridge year calculation.

Table 1. Stocks reviewed at September 2023 Management Track Assessment Peer Review meeting

Level 2 – Expedited Review

Stock	Assessment Lead	Peer Review Panel Conclusion on Stock Status
Atlantic Surfclam	Daniel Hannen	Stock is not overfished and overfishing is not occurring
Black Sea Bass	Emily Liljestrand	Stock is not overfished and overfishing is not occurring
Butterfish	Charles Adams	Stock is not overfished and overfishing is not occurring

Level 3 – Enhanced Review

Stock	Assessment Lead	Peer Review Panel Conclusion on Stock Status
Golden Tilefish	Paul Nitschke	Stock is not overfished and overfishing is occurring
WGOM Cod	Charles Perretti	Stock is overfished and overfishing is occurring
EGOM Cod	Cameron Hodgdon	Stock is overfished and overfishing is not occurring
GB Cod	Amanda Hart	Stock is overfished and overfishing is not occurring
SNE Cod	Alex Hansell	Stock is overfished and overfishing is occurring

Level 2 – Expedited Reviews

Atlantic Surfclam

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

   This TOR was satisfactorily addressed.

The 2024 MT model updated fisheries data showed essentially constant and very low (<1% of landings) discards 2020-2023, with slight declines in landings, and relatively constant Landings Per Unit Effort (LPUE) compared to both the previous assessment and historical data.

TOR 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 2024 MT assessment used the restratified NEFSC clam survey, which focused on Atlantic surfclam in 2018, 2019, 2022, and 2023. The restratification did not appreciably change the survey abundance or length composition trends.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review

This TOR was satisfactorily addressed.

The 2024 assessment of the Atlantic surfclam (Spissula solidissima) stock is a management track update of the previous 2020 Management Track Level 3 report, which was conducted in Stock Synthesis (SS) and found the stock not to be overfished nor to be undergoing overfishing (NEFSC 2022). The 2020 MT assessment collapsed the previous 2-model structure to a single model with 2 areas. The 2024 MT update uses an updated version of SS (SS3.30.14), estimates more parameters, including time varying growth parameters, includes additional years of fisheries-dependent and independent survey data, and re-stratified input survey data. The model updates reference points through 2023 and projections through 2030.

The 2024 MT model estimates a single reproductive stock (i.e. single R0 parameter), with recruits allocated to Northern and Southern regions based on an estimated ratio. The Northern region consists of Georges Bank, and the Southern region consists of SVA:SNE. The two-region model shared productivity and natural mortality parameters (estimated and fixed) but generally estimated region-specific parameters for growth, selectivities (to fisheries and survey), and catchability. Some growth parameters (maximum length Lmax, and von Bertalanffy metabolic parameter K) were estimated as time-varying. Post-hoc linear model evaluations of these parameters showed declines in both parameters across regions, with the exception of the George's Bank K.

TOR 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 2024 MT update finds the Atlantic surfclam stock not to be overfished nor is overfishing occurring. The current $F_{Threshold}$ reference points are unchanged from the 2020 assessment and the SSB reference points (0, MSY, and Threshold) are all 2.7% lower. The 2023 EEZ landings of 10,653 are less than half of the quota of 26,218 (values in metric tons meat weight). The estimated spawning stock biomass (SSB) continued the decline shown in the 2020 assessment, and fell below the SSBTarget 2021-2023, though it remained above the SSB Threshold. Recent recruits and estimated F were relatively constant. The believed cause of the SSB decline is environmentally driven changes in natural mortality and potentially growth.

Based on this management track assessment, the overall stock was not overfished, and overfishing was not occurring.

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

   This TOR was satisfactorily addressed.

The 2024 MT model used SS3 to project the entire stock forward to 2030, assuming time-varying growth parameters remaining at their terminal year (2023) value, and an average recruitment. Spawning Stock Biomass was projected to remain greater than the $SSB_{Threshold}$ and the fishing mortality to remain lesser than the fishing mortality threshold for the entire projected period.

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

   This TOR was satisfactorily addressed.

The Panel discussed the time-varying growth parameters and whether it was appropriate to project those trends moving forward, but did not recommend any changes in the handling of time-varying growth for the next MT assessment. The Panel also discussed the possible benefit of estimating a time-varying natural mortality, and recommended this be investigated in the next MT assessment.
The Panel discussed the possibility of the commercial LPUE exhibiting either hyperstability or hyperdepletion, but did not make any recommendations regarding this for the next MT assessment.

Finally, the Panel discussed the fact that F has been below the F threshold for the entire time-series, and yet SSB has declined over this period, dropping below the SSB target in recent years. One hypothesis is an increase in M due to changing ocean conditions, but another possibility is that since the current F threshold is not mechanistically linked to the current SSB target, it may not be suitable for sustaining the population at the SSB target.

The Panel recommended

Investigating potential for estimating time-varying natural mortality, (e.g., as annual deviations, environmental interactions, or time-block effects).

Developing figures to illustrate the effects of interactions between historically sustainable fishing mortality values and potentially declining natural mortality on future projections, pending capacity to estimate time-varying natural mortality.

Black Sea Bass

The 2024 Management Track assessment for Black Sea Bass (*Centropristis striata*) stock is an update of the 2023 Research Track assessment (NEFSC 2023). This multi-WHAM assessment updates commercial and recreational fishery catch data, research survey indices of abundance, recreational catch per angler index of abundance, and reference points through 2023. Stock projections have been updated through 2026 with the catch of 2024 assumed to be the ABC catch.

The Panel concluded that the 2024 management track assessment update for black sea bass fulfilled the recommendations of the AOP, is technically sufficient to evaluate stock status and provide scientific advice and meets the Terms of Reference for the stock's assessment. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes.

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

This TOR was satisfactorily addressed.

The commercial and recreational landings were updated to 2023. There have been no changes in the algorithm used to calculate landings and discards when compared to the most recent research stock assessment (NEFSC 2023).

There is no length information on individuals in the unclassified market category since 2020. The catch at length for this market category for 2021-2023 is based on the information from before 2020, which introduces additional uncertainty according to the assessment team.

TOR 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 NEFSC and statewide surveys on relative abundance have been updated to 2023 in the VAST analysis. It was noted that the spring 2023 NEFSC bottom trawl survey only covered the Georges Bank area. The impact of this limited coverage is currently unknown, but it certainly introduces a significant source of uncertainty. The Recreational Catch per Angler (RecCPA) index has also been updated to 2023. Both the VAST index and the RecCPA show consistent trends over time, including the updated period.

Recruitment was informed by the VAST index, which includes age-1 data. All age-1 indices were incorporated into the VAST to derive the population trend and age compositions. The panel discussed the use of multi-age VAST and questioned whether the wide range of NEFSC trawl surveys might influence the recruitment signals from the estuary and shallow water surveys, where young-of-the-year (YOY) and age-1 fish are found.

Weight-at-age of ages 2+ have shown decreasing trends in both the northern and southern regions since the mid-2010s. This is a concerning development, and further exploration of the biological reasons and potential impacts on population projections is recommended.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review

> *This TOR was satisfactorily addressed.*

The base model is a single assessment model, a multi-WHAM, with two regions/stocks (north and south) and two fleets (recreational and commercial) in each region. Movement between the North and South stocks was included in the model structure. Two indices were used to calibrate the population dynamics: Rec CPA and VAST (see TOR1). Natural mortality is fixed at M = 0.4. External evidence from Cope and Hamel (2022) suggests that M = 0.4 aligns with its central tendency. The movement parameters were fixed and based on a previous ASAP model output.

The Panel's discussion primarily centered around the choice of M, the application of VAST in developing the relative abundance index, and the movement assumptions in both the base model and the projections. These key model components, highlighted in the research track review, were the focus of the Panel's attention. Sensitivity analyses on M and the inclusion of VAST surveys suggest that the rationale for using M may need further justification or potentially estimating M. The use of VAST should be further explored, particularly regarding how the age-1 fish trend is influenced by the spatial scale and resolution of each survey. The impact of the absence of the 2023 NEFSC trawl survey in the southern region may be investigated in the future.

Retrospective patterns were minor, and retrospective adjustments for terminal year estimates were not needed.

After considerable discussion, the Panel recommended future exploration of using age-varying M, and time blocks for movement assumptions based on the VAST and catch composition changes observed.

TOR 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 spawning stock biomass in 2023 was estimated to be 24,572 mt, corresponding to 219% of the biomass target (SSBmsy proxy = 11,225 mt). The 2023 fully selected fishing mortality was estimated to be 0.82, corresponding to 77% of the overfishing threshold proxy (Fmsy proxy = 1.071).

Based on this management track assessment, the overall stock was not overfished, and overfishing was not occurring.

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

   This TOR was satisfactorily addressed.

Short-term projections were conducted in the WHAM model, which propagates uncertainty in the fleet selectivity, abundance at age including recruitment and environmental covariate. The annual fishery selectivity, maturity-at-age, and mean weights-at-age used in the projections represent the most recent 5-year averages, following the decisions made at the 2023 research track. Mean recruitment for the projections represented the average from 2000 onward.

The stock is projected to remain above the overfished threshold throughout the short-term projection period. The estimated catch at Fmsy for 2025 is 6,193 mt and 4,916 mt for 2026.

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

   This TOR was satisfactorily addressed. The assessment team provided sensitivity runs recommended by the research track review except for the age-dependent M.

The review panel has several research recommendations for potential improvement or clarification of the assessment:

Recommendations to be considered for the next management track review

   Explore the use of age-dependent M, and compare the options of fixed constant M and age-dependent M based on AIC, likelihood, and data fitting. Future exploration of using an informative prior for M is suggested.

   Investigate the trend of age-1 indices and the one from VAST to determine whether it is worthwhile to develop a separate recruitment index.

Recommendations to be considered for mid-long term

   Explore the possibility of developing a YOY index to aid age-1 projections in the future given its potential high benefit in better predicting short-term population size. Currently, YOY indices or black sea bass less than 15 cm are excluded

from all indices. The assessment team noted that the first winter might be a bottleneck, which could disrupt the correlation between YOY and age-1. Consideration of using YOY and environmental factors such as bottom temperature to inform age 1 is encouraged.

Consider estimating region-specific stock status to avoid regional overfishing.

Explore the influence of assuming changes in north-south movement based on the observed timing of catch composition in VAST, and in the spring and fall catch, as an initial step to test whether black sea bass movement between north and south regions changes over time due to water temperature variations.

Consider different environmental conditions or alternative periods in projection. Discussed options include basing projections on mean environmental conditions, terminal year environmental conditions, average recruitment over a period of years, long-term average recruitment, and recent period recruitment.

Butterfish

The 2024 Management Track assessment of butterfish (Peprilus triacanthus) was an update of the 2022 Research Track assessment. The assessment was elevated from the Level 1 (Direct Delivery) 2022 Management Track assessment to a Level 2 (Expedited Review) for 2024, due to a change in the recruitment stanza used for reference points and projections. A change in the recruitment stanza usually requires a Level 3 (Enhanced Review), according to the NRCC's Stock Assessment Process guidance document. However, the change to the stanza itself was minor, with the first year of the most recent recruitment regime changing from 2011 to 2010, and the effect on the reference points and projections was negligible, resulting in a less than 2% change in the estimates of the SSB reference point and MSY compared to using the original regime definition. Therefore, the AOP recommended a Level 2 review for butterfish for the 2024 Management Track assessment.

> ***The Panel concluded that the 2024 Management Track assessment for butterfish met all the TORs and represented the BSIA for this species.***

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

> ***This TOR was satisfactorily addressed.***

The commercial landings and discards at age were updated to 2023.

TOR 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.

Survey indices for fall and spring NEFSC bottom trawl surveys were updated through 2023 and 2022, respectively. The NEFSC Spring survey was treated as missing for 2023. NEAMAP fall and spring survey data were updated through 2023.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review

   This TOR was satisfactorily addressed.

The WHAM model was updated through 2023. There were no substantive changes.

TOR 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 2024 Management Track assessment found that butterfish were not overfished and not experiencing overfishing, with SSB in 2023 being 222% of the SSB threshold and F in 2023 being 3% of the F threshold. This was the same status determination as the 2022 Management Track assessment. The estimates of the F_{MSY} proxy, $SSB\text{-}_{MSY}$ proxy, MSY, and median recruitment in the most recent recruitment regime were all slightly lower than the 2022 Management Track assessment, but not significantly different. F has been below the F threshold and SSB has been above the SSB target for the entire time-series of the assessment.

The change to the recruitment stanza was reasonable and did not change stock status stock status. The SSB_{msy} proxy decreased 1.6% from 38,036 to 37,437.

Based on this management track assessment, the overall stock was not overfished, and overfishing was not occurring.

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

This TOR was satisfactorily addressed.

The Panel shared the concerns about the F_{MSY} proxy raised by the 2022 Research Track Review Panel and the SSC during the 2022 Management Track Assessment. The F_{MSY} proxy for butterfish was estimated at 5.1 in this assessment, resulting in extremely high estimates of catch for 2025 and 2026 from the short-term projections. Catch was estimated at 60,475 mt in 2025 and 38,490 mt in 2026, more than seven times the highest observed catch since the directed fishery reopened and comparable in magnitude to the estimates of SSB in 2024 (71,230 mt) and 2025 (39,203 mt). The high F_{MSY} proxy and high projected catch levels are due to the high values of natural mortality used in the model and the fact that butterfish mature very quickly, with approximately 60% mature by age-1, but are only fully vulnerable to the fishery at age-3. Both the reference point calculations and the projections assume that recruitment is unrelated to SSB, meaning the same level of recruitment will enter the population each year to replace the biomass lost due to the fishery. As a result, according to the model, the fishery has little impact on the spawning stock biomass and subsequent recruitment and must operate at extremely high levels to reduce SSB to 50% of unexploited SSB (the definition of the current F_{MSY} proxy).

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

This TOR was satisfactorily addressed.

In 2022, the SSC chose to use F=2/3M to determine the OFL; this was the F_{MSY} proxy definition from the 2014 SAW/SARC assessment (NEFSC 2014). The Panel recommended that short-term projections be done using F=2/3M and presented in the "Special Comments" section of the short report to provide more context for the current $F=F_{MSY}$ proxy projections.

The Panel also recommended that the next Management Track assessment for butterfish include reconsideration of the current reference points for butterfish, which would require a Level 3 Enhanced Review. Although the 2022 Research Track Peer Review Panel accepted the use of the $F_{MSY}=F_{50\%SPR}$ proxy as the overfishing threshold, the SSC's concerns about that reference point meant that the OFL for 2023 and 2024 was calculated using a different F value. If the SSC continues to have concerns about the scale of the current F_{MSY} proxy, providing a comprehensive evaluation of alternative reference point options will allow them to make a more informed decision than simply defaulting to the F_{MSY} proxy from NEFSC (2014). Options that could be considered include but are not limited to reference points based on total biomass instead of spawning stock biomass and forage-specific reference points such as those recommended by Pikitch et al. (2012).

The Panel briefly discussed the issue of the catchability parameter (q). The model used the estimate of M from NEFSC (2014), which was estimated by the statistical catch-at-age model ASAP. However, to get the current model to produce reasonable estimates of scale in this assessment, q had to be fixed at 0.2, the value used in NEFSC (2014), which was derived from habitat modeling and survey efficiency work. This was identified in the 2021 Research Track and the 2024 Management Track as a significant source of uncertainty, and the Panel agrees with previous recommendations to explore this issue further. While additional empirical work on survey catchability could address this recommendation, a reevaluation of the M estimate used in the model would also be informative and easier to accomplish. A future management or research track assessment should explore whether a more traditional estimate of M based on longevity, life history, or weight-at-age would allow the model to fit q and estimate scale more plausibly.

The Panel also briefly discussed the stability of the recruitment stanza definitions, and suggested a retrospective analysis be conducted to evaluate how often the start year of the most recent stanza changed with the addition of new years of data. Overall, the change in the recruitment stanza had a minimal effect on the assessment results. However, the analysts were fairly confident that this approach could detect a significant

regime change with very few years of data, meaning that future assessments might see more significant changes in recruitment stanzas if this analysis is updated with every management track update.

Pikitch, E., Boersma, P.D., Boyd, I.L., Conover, D.O., Cury, P., Essington, T., Heppell, S.S., Houde, E.D., Mangel, M., Pauly, D., Plagányi, É., Sainsbury, K., and Steneck, R.S. 2012. Little Fish, Big Impact: Managing a Crucial Link in Ocean Food Webs. Lenfest Ocean Program. Washington, DC. 108 pp. Available online at: https://www.lenfestocean.org/en/news-and-publications/published-paper/little-fish-big-impact-a-report-from-the-lenfest-forage-fish-task-force

Northeast Fisheries Science Center (U.S.) (2014). 58th Northeast Regional Stock Assessment Workshop (58th SAW) assessment report. http://doi.org/10.7289/V5KP8043

Level 3 – Enhanced Reviews

Golden Tilefish

The 2024 Management Track assessment Golden Tilefish (Lopholatilus chamaeleonticeps) is the first assessment following the 2024 Research Track assessment. This assessment was implemented using the state-space modeling software known as the "Woods Hole Assessment Model" (WHAM). Unlike most NEFSC assessments, there are no fishery-independent surveys of relative abundance and aging data from landings are not available for all years. Age 1 and 2 fish are infrequent in the landings. Relative abundance is estimated from trends in commercial CPUE using a Generalized Linear Model with individual vessel effects. Sparse data, the complex life history and significant transitions in fishery management measures created challenges for the RTWG and the RT Peer Review Panel regarding model identification. The RTWG selected a preferred model but the RT Peer Review Panel did not fully endorse this model and there was insufficient time to craft an acceptable alternative. The RT Peer Review Panel provided extensive guidance on methods to identify an acceptable model for the MT assessment. In response, the NEFSC relied on a team of experienced assessment scientists to develop a biologically plausible model. Owing to the relative paucity of data, the proposed assessment model incorporates a subset of state-space model features that have been recently used in NEFSC assessments. A

considerable fraction of the MT Review Panel's discussion centered on the process of model identification. Details of these discussions are provided below.

>*The Panel concluded that the 2024 management track assessment update for Golden Tilefish fulfilled the recommendations of the AOP, is technically sufficient to evaluate stock status and provide scientific advice and meets the Terms of Reference for the stock's assessment. The assessment represents Best Scientific Information Available (BSIA) for this stock for management purposes.*

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

> ***This TOR was satisfactorily addressed.***

Commercial landings were updated through 2023. Historically the landings data excluded longline discards, non-target commercial discards, and recreational landings. These removals were added during the RT assessment and updated through 2023. Collectively these removals constitute less than 7% of the total catch in 2023 and much less in most years. Recreational landings averaged about 3% of the commercial landings. Within the model the new catch did not significantly rescale overall abundance.

Prior to 2013 a substantial fraction of the landings was unclassified in some years. This creates additional uncertainty in the catch at age estimates. Through 2017 a pooled age length key was used; in 2021 age-specific keys were used for the first time. In 2024, age length keys were developed for 2009 onward.

TOR 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.***

Presently there are no fishery independent indices of abundance used in the assessment. Three pilot surveys have been conducted since 2017 and a fourth is planned for 2025. These data have not been fully analyzed, particularly with respect to hook selectivity. These indices will be included in the next Management Track assessment.

Indices of abundance are based on CPUE from the commercial longline fishery. From 1976 to 1982 indices were based on analyses by Turner (1986). From 1972 to 1994 the dealer weigh-out data were used for CPUE estimation. From 1991 onward, a GLM model based on VTR data was used to standardized CPUE data. The GLM model was updated with data through 2023.   The conversion from historical VTR estimates to the new CAMS methodology generally increased the CPUE slightly but the proportional increase was consistently low across all years

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review

> ***This TOR was satisfactorily addressed.***

The final WHAM model was developed by a team led by Paul Nitschke and colleagues at NEFSC. Model identification followed the guidance given by the RT Peer Review Panel. Scenarios tested included setting selectivity for age 1 and 2 fish to zero prior to 1989 and setting full selectivity to one for ages 6 to 9. After 1989, age 1 selectivity was set to zero and full selectivity was set to 1 at age 5. This parameterization proved to be unstable. Importantly, inclusion of random effects (RE) on numbers at age, recruitment and selectivity led to a stock that was overfished and overfishing was occurring. If RE on selectivity were included, the model estimates a common selectivity pattern across all years with full recruitment at age 6 and a modest amount of doming at older ages. When RE on selectivity were eliminated and the original stanzas of recruitment were retained, the model predicted that the stock was experiencing overfishing but not overfished.

The WHAM depictions of the resource status differed sharply from the ASAP formulation in 2020 wherein the stock was not overfished and overfishing was not occurring.   Detailed investigation of various model performance tests within WHAM

suggested that earlier conclusions of stock status were optimistic. The final model included RE on numbers at age and recruitment but not selectivity. The best fitting model retained features from the model with RE on selectivity wherein selectivity on age 5 and 6 was set to one. This parameterization reduced the slope of the dome for ages 7 and above, thereby retaining a feature estimated under the model with full RE on selectivity.

TOR 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 biomass and fishing mortality reference points changed substantially in this assessment. Both changes appear to be due to the change in selectivity pattern. In the previous ASAP model full selectivity occurred at age 5 and selectivity for older fish dropped sharply. In the current WHAM model, selectivity of older fish is fixed at 1 for both ages 5 and 6; selectivity on older fish drops less sharply. This implies that the force of mortality is extended over a broader range of older fish, thereby reducing the Fmsy proxy from 0.261 to 0.211. These same changes reduce the potential accumulation of biomass in the older fish resulting in a 31% decline in the SSBmsy from 10,995 in 2021 to 7,537 mt in 2024.

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

This TOR was satisfactorily addressed.

Short-term projections were conducted in WHAM and were consistent with the final model structure. When the population is exploited at Fmsy proxy=0.211, projections suggest a rapid rebuilding of the stock from its current biomass of 58% SSBmsy to greater than SSBmsy of 7,537 mt by 2027. The rapid increase appears to be driven by highly uncertain age-1 recruitment estimates in 2021 and 2022.

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

This TOR was considered satisfactorily addressed. *The assessment team provided sensitivity runs recommended by the Research Track peer review panel except for the age-dependent M.*

We, the Management Track peer review panel has several research recommendations for potential improvement or clarification of the assessment:

Recommendations to be considered for the next management track review

The Review Panel agrees with the assessment that the biggest source of uncertainty is the inclusion of random effects in the WHAM model. This sensitivity however, cannot be isolated from the effects of the change in selectivity within the model beginning in 1987 where the model includes a sharp transition from a relatively flat-topped selectivity to a more dome shaped selectivity. This feature, carried over from the ASAP model, implies a rapid change in fishery behavior and induces a large pool of "cryptic biomass" of fish that are not available to the current fishery. Dome-shaped selectivity patterns in some fisheries are well supported by empirical evidence such as comparative gear experiments and fishery-independent surveys. Unfortunately, direct evidence is lacking for Golden Tilefish.

Comments from industry suggest a basis for the shift in selectivity stanzas as it was noted that transition from J hooks to circle hooks has reduced the catch rate of large fish. It was also noted that standard circle hooks are less likely to capture larger fish because of their bony jaws. The Review Panel noted that direct empirical evidence of this process should be investigated. Historically the fishery used J hooks baited with herring. J hooks with squid were used before the more recent introduction of circle hook baited with squid.

The selectivity pattern of three recent fishery independent pilot surveys over the likely full range of Golden Tilefish has not revealed strong evidence of dome-shaped selectivity. In 2017 the first pilot survey used three hook sizes. No marked differences in catch rates of large fish were observed for the largest hooks. The 2020 and 2023 surveys did not use the largest hook size but instead used the two hook sizes: the standard commercial hook and a smaller hook. The Review Panel strongly suggested a more quantitative analysis of selectivity for all three survey years to see if evidence of reduced selectivity of large fish may be present.

Industry noted that many areas that were fished historically are now within marine protected areas or are otherwise inaccessible to fishing. Deep water habitats are difficult to fish. The Review Panel encouraged the development of maps of historic fishing areas to help support this assertion.

The final model selection process was thorough and well-reasoned. Following the guidance of the RT Review Panel, the model was evaluated with respect to several measures of internal consistency rather than comparison with previous perceptions of model status. In addition to standard residual analyses, goodness-of-fit measures included one-step ahead residuals, jitter analysis, AIC, retrospective pattern, and self tests. Nonetheless, the historical performance of the fishery with almost 25 years of nearly constant quotas is important to note. Frequent assessments, economic factors, individual quotas, and management interventions appear to have stabilized landings over this period. This does not diminish the risk of depending on episodic recruitment to support the fishery and the potential collapse of possible spawning stock biomass if the dome-shaped selectivity pattern is not correct.

Golden tilefish are a long-lived species with a poorly understood life history. Tilefish do not reach maximum size until about 15 yrs old. The fishery is largely supported by fish less than 10 years old, and periodic year classes appear to be reflected in the catch and CPUE analyses. This Review Panel noted the vulnerability of the population if these year classes are in fact, products of newly mature fish rather than a less vulnerable cryptic spawning stock biomass. MT Review Panel suggests that age distribution of the SSB under the final model be compared to age distribution under alternative models, particularly the RE model for selectivity and the initial model with RE for numbers (IID) and selectivity fixed to one for age 5 only.

Recommendations to be considered for mid-long term

Explore the implications of age-varying natural mortality M. Based on first principles of size dependent predation, the likelihood of a constant M over all age classes is considered low
(Lorenzen, 1996). At a minimum, different natural mortality rates for Age 1 and 2 vs older fish should be considered. Several species in the Northeast have recently introduced age-dependent M (e.g., bluefish). Evidence from the literature (Deroba and Schueller 2013) suggests that variations in M across years may be more consequential for determination of stock status.

Experiences with the current model and the data to support it suggest that testing of this feature will come at the expense of some other factor such as random effects on recruitment or selectivity. The next assessment should investigate this possibility and its interaction with use of RE for selectivity.

The MT Review Panel concurred with the NEFSC scientists that a single model is unlikely to fully capture the uncertainty in this assessment. In fact, the MAFMC's SSC explicitly incorporates this principle when selecting an appropriate level of stock uncertainty for deriving an ABC. However, there are presently no operational guidelines for handling ensembles of model to craft management advice. For example, there are no guidelines for identifying the number of plausible candidate models and the relative importance of each in the overall characterization of the resource. Moreover, the Council process is not presently configured to handle the complexity of choosing among multiple models. More formal consideration of decision trees and risk analyses would be needed before ensemble models could be considered.

Atlantic Cod Assessment Overview

The Panel reviewed four level three management track (MT) stock assessments for Atlantic cod. These assessments were previously reviewed as part of the research track (RT) process, where the historical two stock structure (Gulf of Maine, or GOM, and Georges Bank, or GB) was split into four stocks: GB, Southern New England (SNE), Eastern Gulf of Maine (EGOM), and Western Gulf of Maine (WGOM) cod. The RT review accepted the assessment results fully for GB and WGOM cod. The assessments for SNE and EGOM were partially accepted based on issues with the recreational LPUE and lobster pot discard estimates, respectively, that needed addressing.

Overall the Panel was pleased with the work conducted by the analysts for the current MT assessments. The analysts conducted thorough updates to their respective models under tight time frames and provided clear presentations about the model results and justifications for model selection. However, the Panel had some concerns with each of the models for different reasons, which will be detailed in the stock specific reports below.

All four stocks used the Woods Hole Assessment Model (WHAM), with varying degrees of complexity. For three of the stocks, the backup model was a simplified version of the model that basically mimicked ASAP, and would include a retrospective adjustment if appropriate. The justification for this backup was a key finding of the Index-Based Methods Working Group (IBMWG) RT assessment. The IBMWG simulation tested data limited methods that could be used for data rich stocks that had assessments rejected due to strong retrospective patterns. This work showed that an ASAP model with a retrospective adjustment was often superior to the other data-limited methods for setting catch advice. While the Panel supports this finding, it questions the applicability across all stocks as a backup. First, two of the stocks are now limited in their overall data, and are not data rich. Second, differences in the two MT models for SNE and EGOM cod from an ASAP-like model were small, such that simplifying the model would not likely produce differences in the results. Finally, not all assessments are rejected due to retrospective patterns, and a simplified WHAM model might not address the issues behind a rejection. The Panel recommends that careful consideration be given to the fallback option and how different it is from an ASAP model, and also that a Plan-C approach be developed in advance in case an assessment is rejected for reasons other than a retrospective pattern.

The Panel has some concerns about the projections in WHAM and recommends careful evaluation of the results and what could be driving them. For WGOM cod there was an exploration of the long-term, standard deterministic WHAM projections at $F_{40\%}$, which

resulted in the long-term, equilibrium biomass being around 50% lower than SPR-based estimate of SSB_{MSY}. Stochastic projections were also conducted in WHAM that carry forward the random effects, and this resulted in the long-term, equilibrium biomass being around 66% higher than the SPR-based estimate. There is an upcoming RT assessment based on projection methodology, and the Panel recommends that SPR-based reference be used for all stocks (cod and beyond) until there is a better understanding of what is going on in WHAM between the deterministic and stochastic projections.

The Panel also had concerns about the scale of estimates coming out of some of the models. For three of the four stocks, catches were close to the total estimated SSB in recent years. Knowing this would be an issue of debate, prior to the meeting the Panel requested estimates of swept area biomass from the NEFSC survey and exploitable biomass to get a better sense of scale. Unfortunately, these estimates were not provided. It was stated that exploitable biomass is not currently an output in WHAM, and that although it could be calculated manually, there was concern about doing so in a timely manner with confidence because the inclusion of random effects complicates the calculations. Regarding swept area biomass, it was noted that this has not been suggested for use for "roundfish," and is only used for red hake and some flatfish species ([Miller et al., 2023](#)). It was also noted that zeros in recent years for multiple stocks would also be problematic. The Panel agrees that swept area biomass is not suitable as a true measure of total biomass (i.e., catchability < 1) for cod because some are likely able to avoid capture. However, it is potentially a useful measure of a lower limit of cod biomass. Thus, if swept area estimates were substantially higher than the assessment model estimates, that would give reason to be more skeptical of the model results. In contrast, if the estimates were comparable or lower, that would suggest that the model estimates were at an appropriate scale.

The Panel also wondered if the scale of estimates was a topic of discussion at the RT track assessment review. While the information (SSB, F, R, and catch) is all there in the RT reports, they are largely in the form of figures. In contrast, the MT assessment summary reports present the last 10 years of these estimates in table format right at the beginning (Table 1 in every MT assessment), which makes it easier to see scale issues which may highlight a concern. While the Panel understands that the focus of the RT assessments is not to focus on management advice, presenting that information more clearly and up front in the RT assessments could help identify potential issues with the model.

Related to the RT reviews, the Panel feels that it would be helpful for there to be some continuity between the RT and MT review processes. While having different reviewers is

preferred, one option could be that the MT panel (or at least the chair of the panel) is selected prior to the RT review, and be required to sit in on the review process to get a better understanding of the focus of the discussion and any issues that are determined.

Finally, the Panel discussed some issues that were beyond the scope of the MT process, but could be addressed in future RT assessments. Some of the issues identified are related to the current stock structure. Two stocks with ample data have been split into four stocks, two of which still have ample data (GB and WGOM), and two that are limited in the available information (EGOM and SNE). While there are reasonable justifications for the four stock structure, both biological and practical, the limited information for EGOM and SNE cod may cause future issues in the assessments. For example the SNE stock relies heavily on a CPUE index in the recreational fishery. Severe restrictions in this fishery could limit the amount of data available for this index, and the assessment overall.

The Panel also expressed concerns about the inconsistency of approaches for specifying M in the model, as some relied on fixed M-at-age and others relied on age-specific M values. Concerns were discussed over the impact on model fits, but also on the reference points.

Western Gulf of Maine Atlantic Cod

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

> ***This TOR has been met.***

Commercial and recreational landings and discards were updated through 2023 in the model. Recreational fishing is estimated to account for a substantial share of removals, with recreational discards contributing about 20% and recreational landings about 12% over the update period. The panel noted that recreational removals are estimated from the MRIP survey and are subject to greater uncertainty than commercial removals. Updated commercial discards from CAMs included new fisheries, including longline, handline, scallop dredge, surf clam dredge, and lobster pots. Discards from the newly included fisheries accounted for a very small proportion of overall removals from the WGOM stock and had a negligible impact on assessment result (as shown by a sensitivity run). The panel discussed the uncertainty and possible bias in the CAMS lobster pot discard data but decided against removing the data from the WGOM

assessment given the very low sensitivity of results to inclusion of CAMS discard estimates in general and CAMS lobster pot discards specifically.

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

This TOR has been met.

The assessment used updated data from the spring and fall NEFSC bottom trawl survey, the spring and fall Maine / New Hampshire survey, and the NMFS the Maine / New Hampshire age-0 survey, the MADMF spring survey, the NEFSC co-operative bottom longline survey (BLLS), and the industry-based survey (IBS). The panel discussed the spatial coverage of surveys and noted that only the NEFSC bottom trawl survey provides fairly comprehensive coverage of the WGOM stock area including areas 521 and 526 to the south of the GoM proper where almost half of the catches from the newly defined WGOM stock are taken. The ME/NH and MADMF surveys focus on state waters and sample recruits as well as up to ages 4 and 3, respectively. The BLLS covers federal waters in the northern part (only) of the WGOM stock area and sampling rough (untrawlable) as well as smooth (trawlable in principle) bottom. The BLLS was designed to increase sampling of several data-poor and depleted stocks specifically associated with rocky habitat, while also enhancing data collection for some data-rich stocks already seemingly well sampled on the bottom trawl survey. The panel noted that the most survey indices were declining gradually or stable over the most recent decade while the BLLS indices (both spring and fall) declined rapidly over the same period. The reasons for these quantitatively different behaviors are not known and have not been systematically explored. While all indices provide useful information about the WGOM cod stock, differences in spatial coverage and habitat sampled as well as quantitatively conflicting trends should be considered when evaluating the representativeness of different indices as measures of relative abundance of the overall stock.

Several surveys had notable gaps in the update period. Most surveys (except for the ME/NH fall and BLLS fall surveys) were paused in 2020 due to Covid-19. In addition, the spring NEFSC bottom trawl survey in 2023 was only partially completed, deemed non-representative and dropped from the assessment; the MADMF survey caught zero cod in 2022 and 2023 and these years were treated as missing data by the model; and the IBS has not been conducted since 2019.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review.

> ***This TOR has been met with the inclusion of changes to the model recommended by the panel.***

Estimates of annual fishing mortality, recruitment and spawning stock biomass (but not total biomass) and estimates of their uncertainty were provided for the time series using the assessment method approved in the RT. Model diagnostics showed overall good fits to the indices and age compositions, low retrospective patterns (no retrospective adjustment needed) and passed the jitter analysis and self-test. The overall patterns in fishing mortality, recruitment and spawning stock biomass were similar to those estimated in the RT for the overlapping period.

The panel noted and considered at length some concerning patterns in the most recent years (years added since the RT). In the proposed base run, catch in 2023 was 90% of the estimated SSB, and with the assumption of bridge year catch would have been 206% of the estimated SSB in 2024. Model estimates of swept area and exploitable biomass were requested by the panel to better gauge removals relative to estimated biomass but could not be provided (see summary discussion for additional details). Estimated F averaged 0.4 from 2019 to 2022, then spiked to 0.77 in 2023, representing an 80% increase in F from the previous year (0.42). Based on the time series of recruitments presented in Table 1 of the assessment summary (since 2014), the 2018 - 2022 year classes were above average in all but one year (2019), and represented four of the five highest recruitments. Also, the base run Jan-1 estimate of SSB in 2024 was 320 mt, which would mean that the WGOM cod is currently comparable in size to the SNE and EGOM stocks

The panel noted that the terminal patterns of annual biomass removals close to or exceeding spawner biomass and dramatically increasing F estimated in the base run were concerning and seemingly unrealistic in part due to information about this fishery

which did not indicate a rapid increase in fishing effort or dramatic decline in availability. Leave-one-survey-out sensitivity runs, leaving out the 2023 data for each of the surveys (not the whole survey, just the terminal year) were provided in the MT overview presentation. Results in the most recent years were shown to be highly sensitive to the 2023 spring BLLS data: Removal of this one year of BLLS data resulted in a drastically lower estimate of terminal F and a higher estimate of terminal SSB (Figure 1).
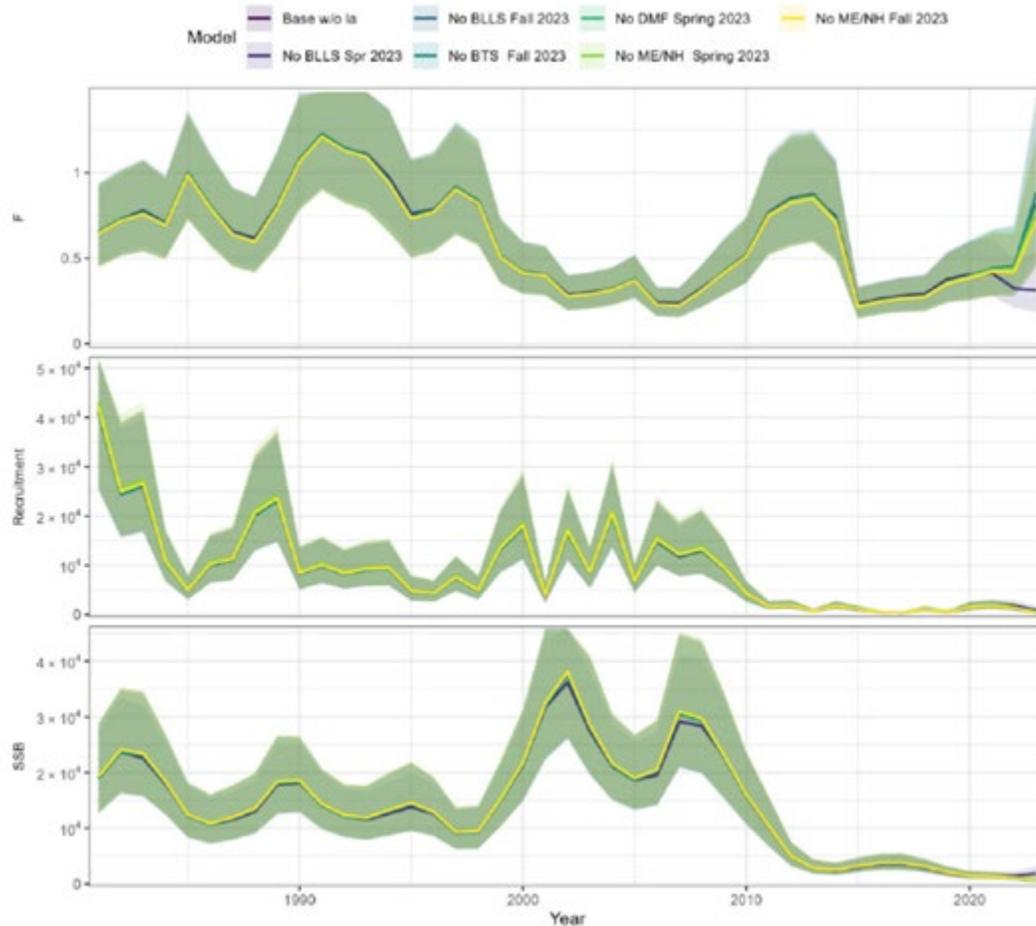

Figure 1. Leave-one-survey-out in 2023 sensitivity runs. Estimates in the terminal and most recent years are highly sensitive to the 2023 spring BLLS data.

The Panel also requested a sensitivity run excluding the BLLS surveys entirely from the assessment (spring and fall BLLS surveys, all years (2014-23)). Removal of the BLLS survey resulted in overall lower F and higher SSB estimates for the recent years. This is not surprising given that the BLLS indicated a more rapid decline in abundance over the past decade than the other surveys (with the possible exception of the IBS). The panel

noted the sensitivity of results to the exclusion of this survey, combined with the above noted issues of divergent quantitative trends as a topic for future research but did not recommend exclusion of this (or any other) survey from the MT assessment.

The Panel considered the assessment results in the most recent years and their extreme sensitivity to the 2023 spring BLLS data at length. There are no reasons intrinsic to the BLLS survey to exclude the 2023 spring data since all stations were sampled, the CV was not high, and no error was found in the calculation of that index. The model was highly sensitive to the 2023 spring BLLS data but did not fit the index well in that year (the index point estimate and CI were far below the model prediction, Fig. 2 below). Exclusion of the 2023 spring BLLS data alleviated the above noted, strong concerns about the SSB and F estimates produced by the base run for the most recent years. The panel further noted that the influence of the 2023 spring BLLS data was likely exacerbated by the sparsity of other index data for the same and recent years, with the spring NEFSC bottom trawl treated as missing in 2023 and the MADMF survey treated as missing in 2022 and 2023 (for details see TOR 2 above). In the light of these considerations, the panel resolved to recommend removing the 2023 spring BLLS index from the 2024 MT assessment. The lead analyst recorded disagreement with this decision. It should be noted here that the decision to exclude the 2023 spring BLLS index was taken to address strong concerns over the model's F and SSB estimates in the most recent years, it does not affect the stock status determination nor the overall outlook for this stock, nor does it prejudice the use of the 2023 spring BLLS index in future assessments of this stock.
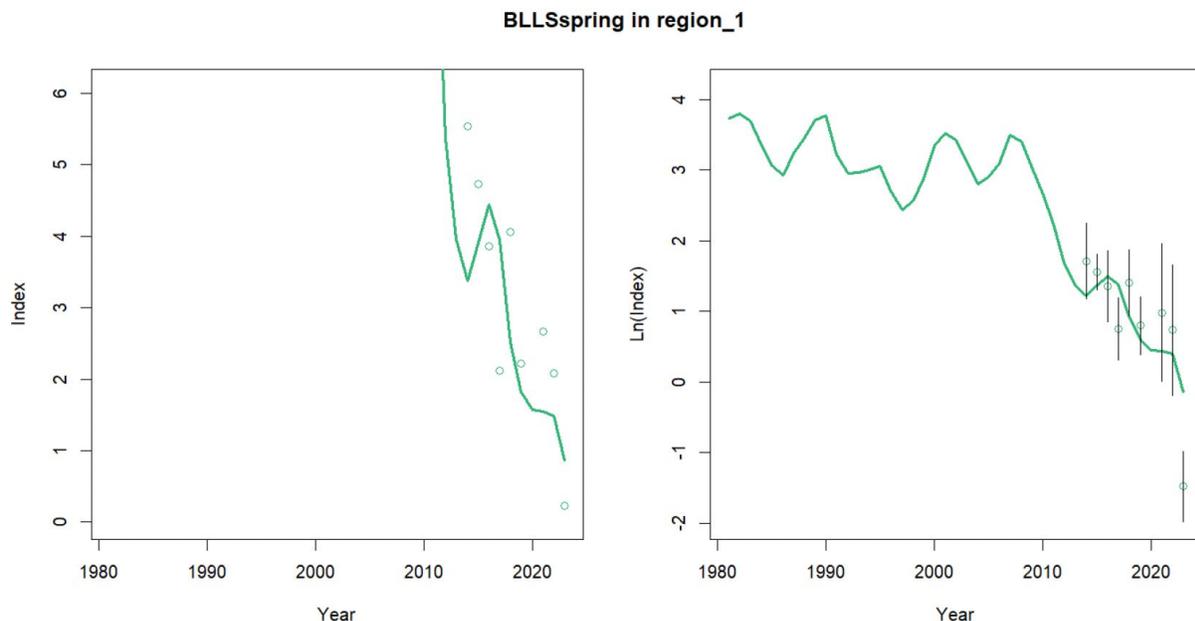

BLLSspring in region_1

Figure 2. Spring BLLS survey estimates (points) and resulting model fit showing the anomalously low 2023 estimate

TOR 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 has been met.***

The Panel was presented with both SPR 40% proxy BRPs and a stochastic projection approach where the population was projected into the future at $F_{40\%}$, and the projected equilibrium biomass was used to calculate SSB_{MSY}. This analysis was originally conducted using the standard projection in WHAM (which they called the "deterministic" projection). However, fishing at F_{MSY} resulted in an SSB_{MSY} around half of the SPR-based estimate. The analyst conducted what was called the "stochastic" projection, which incorporated the NAA random effects throughout the entire projection period. This resulted in an SSB_{MSY} estimate about 66% higher than the SPR-based estimate. The stochastic SSB_{MSY} estimate was suggested as the preferred estimate. The discrepancy in projection estimates from WHAM was concerning to the Panel. In particular, the stochastic estimate seemed unreasonably high. Without clear knowledge of what was driving the different results in the deterministic vs. stochastic projections, the Panel was not comfortable endorsing the stochastic projection estimate of SSB_{MSY}. Instead, the Panel approved the SPR 40% proxy approach to estimate SSB_{MSY}. Weight-, selectivity-, and maturity-at age were fixed at the most recent five year average to calculate $F_{40\%}$, and the resulting spawning biomass per-recruit was then multiplied by the mean recruitment over the assessment period (1981-2023) to get SSB_{MSY}. SSB in the terminal year was 1,847 mt, which is only 3% of the SSB_{MSY} =62,677. Fishing mortality in 2023 was 0.31, exceeding the target F_{MSY} proxy =0.19. Based on the BRPs and estimated current biomass and F,

> **The stock is overfished and overfishing is occurring.**

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

> ***This TOR has been met.***

The Panel approved the bridge year calculation assuming catch in 2024 was equal to the 2023 catch, and recommends using the standard projections in WHAM to calculate the OFL.

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

The new research recommendations provided to the Panel were to

Continue to monitor and improve projections

Treat index values of zero as zero in WHAM (not missing)

The Panel agrees with these recommendations, and also suggests reviewing the indices used for this stock, particularly in light of the new stock structure. This may include exploring VAST indices to investigate divergent index signals.

Eastern Gulf of Maine Atlantic Cod

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

This TOR has been partially met.

Commercial and recreational discards and landings were updated through 2023 in the model. Updated commercial discards from CAMs included new fisheries, including longline, handline, scallop dredge, surf clam dredge, and lobster pots. These newly added discards dramatically increased the overall estimated catch from 2020 to 2023, almost exclusively from discards in the lobster fishery. The panel was surprised that only four years of data were included, and learned that there is ongoing research at the University of Maine to estimate these discards, but estimates were not available in time for the MT assessment. The Panel was also concerned about the magnitude of these estimates relative to the estimated SSB, which made them seem implausibly high (discussed in more detail in response to ToR 3). There was no information provided about the specifics of these estimates beyond that the CAMS methodology was used.

The presentation for EGOM was on Tuesday June 18th, and on the morning of Thursday June 20th the Panel received a letter from the Commissioner of the Maine Department of Natural Resources expressing concern over the estimated CAMS

discards in the lobster fishery. The Panel felt this letter was extremely helpful in understanding potential issues with the discard estimates. The letter indicated that the discard estimates were coming from federally permitted lobster vessels that have federal permits for other fisheries. This section of the fishery is very small in comparison to the overall fleet, and also very few observer trips are allocated to these vessels (18 observer days were assigned for 2023-24 out of hundreds of thousands of trips in Maine alone).

Ultimately, the Panel **did not accept** the four years of CAMS data in the model because only four years were included and because of concerns about them being biased high. The Panel recommends continued collaboration with the U. Maine researchers to improve the estimation of discards over time in the lobster fishery. It should be noted that the Panel did not reject CAMS estimated discards in the lobster fishery for the other stocks reviewed, but does have similar concerns about their estimation. However, the estimates for the other stocks were negligible compared to other catch sources and did not have an impact on model results.

Weight-at-age of the catch was based on pooled observations over the time period, and were thus time-invariant. While this is a reasonable approach given the available data, the panel has concerns overall about the lack of data in the fishery. In particular, the Panel is concerned about the limited age information in the commercial landings, as there are no age composition data since 2013.

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

This TOR has been successfully met.

The model updated data from six survey indices, the spring and fall NEFSC surveys, the spring and fall Maine / New Hampshire surveys, the Maine / New Hampshire age-0 index, and the Sentinel hook and line survey. There was a formatting issue in the Sentinel survey that was corrected from the RT assessment, and age information for the Sentinel survey was updated with new information provided by the Maine Center for Coastal Fisheries. Similar to the commercial data, the Panel had concerns about limited age information across many of the surveys. In addition, the NEFSC surveys were either not conducted or did not collect any cod (true zeros) since 2020.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review.

> ***This TOR has been partially met.***

As noted in response to ToR 1, the Panel did not accept the estimated lobster discards from 2020-2023. The proposed model which included those four years was presented to the Panel, and the Panel noted that the estimated discards of 79 mt compared to the estimated SSB in that year of 129 mt. The discards are based on an assumed mortality of 16.5%, implying that approximately 479 mt of cod had been collected in lobster traps that year. The Panel also noted the unrealistic behavior of F being minimal in recent years and then spiking with the inclusion of lobster discards from 2020 onward. The minimal F was due to not having estimates of the lobster discards in previous years, and the Panel asked if a sensitivity run had been explored where some assumed level of lobster discards had been hindcasted into years prior to 2020. A sensitivity run had been conducted where the average discards from 2020-2023 (44 mt) was included in earlier years. However, this model resulted in historical F values more than 25x FMSY in some years, being in excess of 10x FMSY more recently (model m2 in Figure 3). The Panel requested additional sensitivity runs with lower assumed catch values, and were provided two additional runs where the assumed catch was 50% and 25% of the average (22 and 11mt; models m4 and m3 in the Figure, respectively). These models also produced very high estimates of historical F.
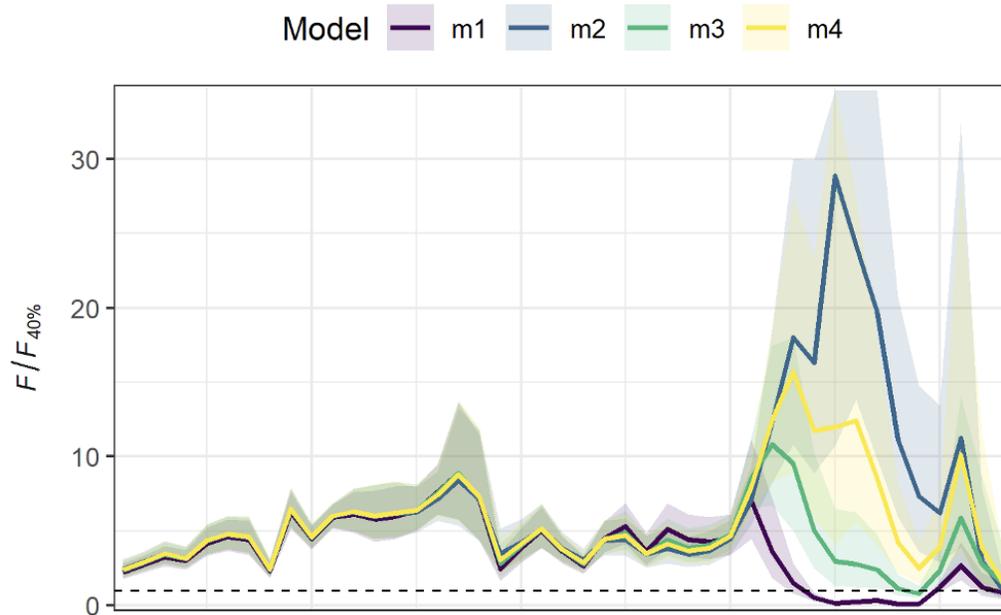
Figure 3. Sensitivity runs comparing the effect of different levels of assumed historical lobster discards on the estimated F to the MT run (m1) that included only four years of lobster discards from 2020-2023. The models assumed historical lobster discards were equal to the average of 2020-2023 (m2), 25% of the average (m3), and 50% of the average (m4).

These sensitivity runs also produced different effects on status in the terminal year, with lower Fs (and no overfishing) for the runs with higher historical catches, and higher Fs (with overfishing) when lower historical catches were used (Table 1).

F estimates	2019	2020	2021	2022	2023
m1	0.03	0.32	0.72	0.33	0.23
m2	1.98	1.70	3.08	0.83	0.28
m3	0.22	0.64	1.60	0.76	0.40
m4	0.67	1.07	2.76	1.07	0.42

Table 1. Estimated F in recent years for the sensitivity runs. The estimated F_{MSY} for this stock is 0.27

m1: MT Model
m2: Hindcasted average lobster fishery discard 2020-2023
m3: Hindcasted 0.25* average lobster fishery discard 2020-2023
m4: Hindcasted 0.5* average lobster fishery discard 2020-2023

These results suggest that this level of removals could not be sustained in prior years, which casts doubt on the magnitude of the average from 2020-2023 also being too high overall. For these reasons, the Panel felt inclusion of only the CAMS estimates from 2020 onward was inappropriate.

There was considerable discussion about what should be done for this ToR, and the Panel weighed two options. The first option was to use the bridge run that had been explored that did not include the new CAMS estimated discards but did update other data streams. The model would of course be missing some important sources of mortality, but the alternative was to reject the model. By rejecting the model, the Panel would also not accept the Plan B option, which was an ASAP-like model with a retrospective adjustment if appropriate. The Panel would not accept the proposed fallback for several reasons. First, that model had the same inputs as the proposed model, including the CAMS discard estimates. Second, the proposed MT model only differed from ASAP in that it estimated lag-1 autocorrelation in recruitment. The Panel felt that such a model would likely be very similar to the original model. Finally, the Panel noted that the justification for this backup approach was not relevant for this stock. The Index-Based Methods Working Group report was the justification for the backup, because they showed that ASAP-like models with retrospective adjustments were very often superior to other data-limited approaches. However, that analysis explored only data-rich cases where assessments were rejected due to large retrospective patterns. The Panel feels that EGOM is data-moderate at best, with considerable uncertainty and gaps in some of the data sources.

If the Panel had rejected the model and the fallback approach, they would not have had clear guidance on what should be done. One option could be the I_{smooth} (formerly called Plan-B smooth) method that has been used for many stocks in the region, but this requires a recent average catch (which would exclude lobster discards), and trends in abundance. Current applications for other stocks use average trends in the spring and / or fall NEFSC survey, but there have been missing years and years with zero catch recently.

Because of these challenges, the Panel proposed using the bridge run, including new CAMS discard estimates for all fisheries except the lobster fishery. While the Panel acknowledged that this model was flawed due to the lack of a known source of mortality, they felt it was more appropriate than including only four years of likely overestimated discards. The Panel felt that this option was best because it kept the door open for the model to be potentially updated with discard estimates from the U. Maine

research in the next MT assessment, as opposed to fully rejecting the model which might require going to another RT assessment.

The Panel also had concern about the combined selectivity assumption for the single fleet, especially if recent catches are heavily dominated by lobster discards. Future models should explore a separate selectivity block for the years where catches are dominated by lobster discards.

TOR 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 has been successfully met.***

The BRPs were calculated using the SPR 40% proxy approach. Weight- and maturity-at-age were fixed at the most recent five year average, while selectivity used the terminal year to calculate $F_{40\%}$, and the resulting spawning biomass per-recruit was then multiplied by the mean recruitment to get SSB_{MSY}. SSB in the terminal year was 267 mt, which is only 12% of the SSB_{MSY} = 2,184 mt. Fishing mortality in 2023 was 0.006, below the target F_{msy} proxy =0.27. Based on the BRPs and estimated current biomass and F estimates,

The stock is overfished, but overfishing is not occurring.

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

> ***This TOR has been met.***

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

The most important research priority that was iterated by the members of the 2023 RT and re-iterated by the RT peer reviewers, was the necessity to estimate and include dead discards of cod that were captured in EGOM lobster pots.

The discard data from CAMS that was used in this management track directly includes these dead discards 2020-2023.
  - Spatial coverage of data for this region (especially inshore) is low.

- There may be a mismatch between cod length/age distributions in lobster pots and cod landed commercially.
- Discard mortality rate in CAMS is from peer reviewed literature, but actions on the water may vary from study conditions (i.e. actual discard mortality is probably higher).

There is still a strong need to estimate these dead discards prior to 2020 and verify estimates used 2020-2023.

An increase in sampling is needed to appropriately characterize the age composition of the stock in the recent time series.

Appropriate coverage of the new EGOM stock for biological sampling of cod is a high priority.

Overall the Panel agrees with these Research recommendations. Although we did not agree with the use of the CAMS estimates, continued research into these estimates and potential sources of bias is warranted. The highest priority is to continue to collaborate with researchers estimating lobster discards over time, and careful consideration of the selectivity assumption for these discards.

Georges Bank Atlantic Cod

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

> *This ToR has been successfully met.*

The model includes catch and discards from the U.S. and Canadian fleets, combined into a single fleet. These data were updated from the RT assessment, although there were some issues and changes in updating the data. First, insufficient port sampling in the U.S. resulted in no catch-at-age information in 2022 and 2023 in the U.S. fleet. Canadian age composition was used in these years. In addition, the Canadian method for calculating landings-at-age was revised, and was updated in this assessment. Estimated discards from CAMS were updated, although there was some concern about instability in the CAMS estimate based on the date the data were accessed. The discard estimates also included discards in the lobster fishery which were not included in the RT assessment, and these estimates were less than 1% of the total removals 2022 and 2023. Calculation of weight-at-age in the catch was modified to account for differences in Canadian and U.S. methods and units, and filled gaps in 2022 and 2023 using Canadian weight-at-age.

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

> ***This ToR has been successfully met.***

The model included three survey indices of abundance: the DFO spring survey, and NEFSC spring and fall surveys. There was a year of missing data for the DFO survey 2022 due to a vessel change and lack of calibration, and for the NEFSC spring survey in 2023 due to an interruption of the survey. The RT assessment used uncalibrated Bigelow estimates in the NEFSC surveys, and this error was corrected in the current MT assessment. The RT review panel recommended splitting the NEFSC surveys into the Albatross and Bigelow vessels, and to explore the VAST model. Neither approaches were considered in the MT because the correctly calibrated data resolved much of the diagnostic issues identified in the RT assessment.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review.

> ***This ToR has been successfully met.***

A number of bridge runs were explored, and most continued to have a strong retrospective pattern. The inclusion of 2DAR1 random effects for number-at-age greatly improved the fit (AIC) and reduced the retrospective pattern. However, the model would not converge until a change was made to the selectivity in the NEFSC fall survey that was deemed "not biologically realistic" by the analyst because it resulted in full selectivity at ages 1-3, and declining selectivity thereafter. The Panel had concerns about this unrealistic result that allows for model convergence, but that this was not sufficient grounds for rejecting the model. The Panel recommends that the cause of this

selectivity pattern be explored further in future assessments. In addition, the Panel had concerns about the impact of the combined Albatross and Bigelow surveys. The Panel recommends future assessments explore splitting the surveys and estimating separate catchabilities and selectivities.

TOR 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 has been successfully met.*

The BRPs were calculated using the SPR 40% proxy approach. Weight- and maturity-at age were fixed at the most recent five year average, and selectivity in the terminal year was used to calculate $F_{40\%}$, and the resulting spawning biomass per-recruit was then multiplied by the mean recruitment to get SSB_{MSY}. SSB in the terminal year was 2,668 mt, which is 32% of the SSB_{msy} = 8,290 mt. Fishing mortality in 2023 was 0.13, below the target F_{msy} proxy = 0.23.

Based on the BRPs and estimated current biomass and F estimates, the stock is overfished, and overfishing is not occurring.

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

> *This ToR has been successfully met.*

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

The recommendations coming out of the assessment were:

Develop model with separate Bigelow and Albatross indices

Develop VAST indices to investigate conflicting spring index signals

Expand effort to obtain biological samples

Participate in broader discussion to define prevailing conditions for reference points and projections

Explore stochastic projections

Evaluate possible time variation in DFO survey index that could contribute to poorer fit

Investigate mechanism for extreme age truncation in NEFSC fall index

The Panel supports all of these recommendations. The exploration of the projections should consider both the deterministic and stochastic projections in WHAM, and possible reasons for divergence if they occur (as occurred for WGOM cod). In addition, there should be an exploration into why the model does not converge without the unusual NEFSC fall selectivity. Splitting the NEFSC survey is a top research priority, and might help address this issue.

Southern New England Atlantic Cod

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

> *This ToR has been successfully met.*

The model includes a time series of catch and discards from commercial and recreational fishery for 1981-2023. This stock is spread across a large number of statistical areas, covering the area from Cape Cod in the north to Outer Banks in the south. However, commercial landings are coming mostly from very few northern areas with trawl being a dominating gear followed by the handline. Commercial catch and discards were based on the Area Allocation method (prior to 2021) and more recently on CAMS methodology (beginning in 2021). Historically total landings were derived from the weigh-out reports from commercial seafood dealers and considered a census of total landings. The time series of Atlantic cod recreational and commercial catch were updated by the Working Group for the Research Track Assessment in 2023 where the spatial allocation of catch was revised to align with the new spatial footprint of Atlantic cod stock areas. The time series was further updated through 2023 for the Management Track Assessment.

Recreational harvest and discards are available from the NOAA MRIP survey starting in 1981. The survey intercepts a relatively large number of trips which results in the CV or the proportional standard errors (PSE) estimated in the range of 20-30%. Recreational landings of SNE cod dominated for nearly the entire time series making over 90% in recent years. Furthermore, the assessment responded to the research track recommendation and developed recreational harvest and discard estimates for wave 1 from VTR data for charter fleet covering 1996-2023. There are studies available that estimate discards mortality. Overall, total removals by commercial and recreational fleet appeared to be well estimated.

The assessment uses age structure information from commercial landings and NMFS spring survey. Commercial age sampling is missing for many years, especially prior to 2004 and between 2014 and 2018. This is compensated to some degree by the availability of age samples from NEFSC spring survey. There are no age samples for the recreational fishery, but there are many length measurements. The assessment team developed a state space model to convert lengths into ages, thus resolving the problem. Weight at age and maturity at age are assigned to multiyear averages, as annual collections do not always have a sufficient sample size. This is a potential source of uncertainty, as data from other stocks indicate significant changes in weights at age (a strong declining trend for legal size older fish) and a trend towards early maturation.

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

This ToR has been successfully met.

The model included NEFSC spring survey and recreational LPUE which was updated for the Management Track Assessment to the CPUE by including the trips with zero catch. While the NEFSC trawl index has large and consistent spatiotemporal coverage (except for the 2023), the survey does not encounter cod frequently in recent years and covers mostly younger ages. Conversion of recreational length measurements into age with state space model provides consistent characterization of recreational catch age structure through time. The assessment increasingly relies on the recreational survey and a development of recreational CPUE was a significant improvement. However, the Panel notes that the recreational catch estimates and the CPUE index are not independent.

TOR 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 backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review.

> ***This ToR has been successfully met.***

Key concerns from the RT assessment were the exclusion of 0s in the recreational index of abundance, and in the absence of Wave-1 estimates in the recreational catch. Both of these issues were successfully addressed in the MT assessment. Overall the model fits were reasonable and the retrospective bias was low.
Similar to WGOM cod, the Panel had concerns over the scale of estimates, with total catch in weight close to the spawning biomass. This pattern was apparent in many years going back to at least 2014 (based on Table 1 in the MT summary report), which commonly resulted in F values above 1. Relatively high uncertainty in the recreational catch may be playing a role in this pattern in some years, as the catch is dominated by recreational landings and discards.
No back up assessment alternative was needed. However, if one was needed, the Panel may not have approved of an ASAP-like model. This model was very similar to ASAP, except it included time-varying catchability in the recreational CPUE. The Panel recommends proposing a Plan-C fallback option for this stock as well.

TOR 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 has been successfully met.***

The BRPs were calculated using the SPR 40% proxy approach. Weight-, selectivity-, and maturity-at age were fixed at the most recent five year average to calculate F40%,

and the resulting spawning biomass per-recruit was then multiplied by the mean recruitment to get SSB_{MSY}.

Based on WHAM assessment, SSB has been declining since the start of the time series and was low over the last two decades. SSB in the terminal year was 339 mt, which is only 3% of the SSB_{msy} =11,359. Fishing mortality in 2023 was 0.975, exceeding the target F_{msy} proxy =0.12. Based on the BRPs and estimated current biomass and F estimates, the stock is overfished, and overfishing is occurring.

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

This ToR has been successfully met.

Short term projections were conducted in the WHAM model projections were conducted from 2025-2027. For projections, the terminal year catch was used as the interim estimate for 2024 treated as a bridge year. However, owing to the high uncertainty in recreational catch estimates over time, the Panel had concern that carrying forward the last years catch into the bridge run could be problematic if that estimate were biased high. As a result, the Panel the most recent three-year average F be used in the bridge year for the projections, instead of calculating F based on an assumed catch. Three years was selected to account for some of the variability in F that also results from uncertainty in the annual recreational catch estimates.

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

The 2023 Research track assessment Panel noted the lack of recreational removals estimates for Wave 1 as well as the fact that the recreational LPUE index did not include trips with zero catch for cod. The assessment lead addressed both deficiencies by developing Wave 1 recreational harvest estimates from VTR database and by recalculating recreational index as catch per unit of effort, CPUE including trips targeting cod but not catching it.

Additional research recommendations included:

Improved biological sampling

Further work to develop the rec CPUE time series

Fitting the model directly to length observations from the recreational fleet. This potentially could be accomplished using the growth branch of WHAM

Discard mortality study from the recreational fishery (Proposal in the works)

The Panel agrees with these recommendations, and adds that exploration of splitting the NEFSC spring survey into the Albatross and Bigelow surveys is warranted.

Appendix A. Summary of Assessment Oversight Panel Meetings for June 2024 Management Track Stock Assessments

The NRCC Assessment Oversight Panel (AOP) met on February 28, 2024 to review the Management Track Assessment plans for Butterfish, Atlantic Herring, Atlantic Surfclam, Black Sea Bass, and the four newly designated stocks of Atlantic Cod: Eastern Gulf of Maine, Western Gulf of Maine, Georges Bank, and Southern New England. A second meeting followed on April 4, 2024 to review the Management Track Assessment Plans for Golden Tilefish following the review of the Research Track Assessment of the stock in March; the AOP also revisited discussions on assessment review level for Atlantic Surfclam and Georges Bank Cod at this time based on requests from the assessment leads, who had noted shifts in assessment plans potentially necessitating a shift in the plans presented during the earlier discussions. After convening the second meeting, the AOP was advised by the WGOM Atlantic Cod assessment lead that additional projection analyses for the stock (described below) would be included in the assessment and may necessitate additional review; the AOP discussed this issue asynchronously via email. One assessment was recommended for Level 1 Reviews (Direct Delivery); this assessment will undergo an internal review before being delivered to the appropriate management body. The assessments for stocks/species recommended for Level 2 and 3 peer reviews will be reviewed during meetings scheduled for June 18-21 and June 24-28, 2024.

Assessment Oversight Panel Members

Kristan Blackhart (Chair), Northeast Fisheries Science Center, Woods Hole, Massachusetts

Mike Celestino, representing the Atlantic States Marine Fisheries Commission, New Jersey Department of Environmental Protection

Lisa Kerr, Ph.D., Chair of the NEFMC Scientific and Statistical Committee, Gulf of Maine Research Institute

Paul Rago, Ph.D., Chair of the MAFMC Scientific and Statistical Committee, NOAA Fisheries (retired)

Meeting Details
This meeting was guided by the NRCC-approved stock assessment guidance documents. Standard background documents were provided to the Panel in advance of the meeting:
- An updated prospectus for each stock

- An overview summary of all the salient data and model information for each stock
- The NRCC Guidance memo on Management Track Assessments

Additional documents submitted by stakeholders and partners were provided to the AOP where available. Prior to the meeting, each assessment lead prepared a proposal for their Management Track Assessment. The proposal reflected the Research Track or most recent assessment results, the peer review panel Summary Report results and any initial investigations conducted for the Management Track Assessment.

At the meeting, each assessment lead gave a presentation on the data to be used, model specifications (if applicable), evaluation of model performance, the process for updating the Biological Reference Points, the basis for catch projections, and an alternate assessment approach if their analytical assessment was rejected by the peer review panel.

Major Recommendations for Review of Individual Stocks

In general, the AOP approved the plans presented, but recommended several points of emphasis to the recommended review levels as summarized below. AOP guidelines can be found in the [stock assessment process document](#).

Stock	Assessment Lead	Review Level	Rationale and Comments
Butterfish	Charles Adams	Level 2	**Rationale:** No new sources of information; no anticipated changes to the assessment model; recruitment stanza used for projections will be updated to start in 2010 instead of 2011 (as in 2021 RT/2022 MT) due to updated condition analysis - change in recruitment stanza technically is a Level 3 review, but extra information was presented indicating this change propagated through with minor changes to assessment results.
Atlantic Herring	Jonathan Deroba	Level 1	**Rationale:** No new sources of information; no changes to model; no changes to projection methods; RT review scheduled for 2025. Concerns about overly optimistic projections and impacts on catch advice; also missing survey data, potential changes to change point analysis, retro adjustment - if the cumulative impacts of changes to projections on catch advice appear large as assessment develops, lead to flag this as requiring an elevated review.

Stock	Assessment Lead	Review Level	Rationale and Comments
Atlantic Surfclam	Daniel Hennen	~~Level 1~~ Level 2	**Rationale:** Originally**,** based on no new sources of information, no changes to model, no changes to projection methodology, and backup approach previously vetted, led to Level 1 review. Subsequent identification of significant trends in VB K parameters that the lead wanted to include in the modeling framework resulted in Level 2 review.
Black Sea Bass	Emily Liljestrand	Level 2	**Rationale:** First MT following RT in 2023; uses RT accepted model; presented sensitivity analyses required by RT peer review focused on VAST index development for guidance on proceeding with VAST vs. individual indices. AOP agrees with the assessment plan for use of VAST index but feels that the additional sensitivity analyses requested by the RT review panel require additional peer review via the MT process.
Eastern Gulf of Maine Cod	Cameron Hodgdon	Level 3	**Rationale:** Newly defined stock, first MT following RT in 2023; MT plans to incorporate cod bycatch estimates (data not available yet, so magnitude still unknown) from EGOM lobster fishery into total catch; uses same projection methodology from RT; backup as recommended from RT (simplified WHAM w/ fewer RE). Lobster bycatch represents an unvetted, new data resource that deserves flexibility for the review panel to investigate.
Western Gulf of Maine Cod	Charles Perretti	~~Level 2~~ Level 3	**Rationale:** Newly defined stock, first MT following RT in 2023; no new data; no changes to model; no changes to projection methodology; backup as recommended from RT (simplified WHAM w/ fewer RE). Suggests Level 2 because it is a new stock. Considering projections explorations if time permits; if any significant improvements are found, would request elevated review. Level 2 is appropriate because assessment will be supporting status determination for the first time. In April, assessment lead notified AOP of plans to explore stochastic projection approaches, which could result in changes to BRPs so review level elevated to Level 3.

Stock	Assessment Lead	Review Level	Rationale and Comments
Georges Bank Cod	Amanda Hart	~~Level 2~~ Level 3	**Rationale:** Originally based on stock area redefined, first MT following RT in 2023; no new data; model changes include correcting calibration of NEFSC survey indices, adopt correction in equilibrium initiation numbers-at-age, and correct weight-at-age input; no changes to projection methodology; and backup as recommended from RT (simplified WHAM w/ only FE), this assessment was assigned a Level 2 review. An additional request for analysis on spatial allocation of biomass estimated for the U.S. Georges Bank stock area to the eastern Georges Bank transboundary area required elevation to Level 3.
Southern New England Cod	Alex Hansell	Level 3	**Rationale:** Stock area redefined, first MT following RT in 2023; no new data; model changes include new recreational CPUE index (previous LPUE index did not include zero observations) - will impact model fits, but unknown what overall impacts are until index is available and incorporated - addresses a major recommendation from RT review panel; projection methodology unchanged from RT; backup as recommended in RT.
Golden Tilefish	Paul Nitschke	Level 3	**Rationale:** RT review recommended additional model exploration (i.e. did not identify preferred configuration or agree with WG base model). Challenges remain on the sensitivity of adding random effects to the data poor WHAM model for tilefish, as well as balancing model diagnostics with biological realism. Disagreement between review panel and WG on plan B approaches. Panel suggested using a suite of models to indicate uncertainty around projections, but did not provide advice on how to accomplish this.

Individual Stock Discussion Summaries

Butterfish
(AOP Lead: Mike Celestino)
Recommendation: Level 2 (Expedited Review)

The current assessment model for Butterfish is the 2021 RT accepted WHAM model, as updated during the 2022 MT. The proposed work plan for the 2024 MT assessment is to update all data through 2023, using the current WHAM model configuration with no changes. Biological reference points will be updated using the 2021 RT-approved methodology, though the recruitment stanza for short term projections and reference points will be updated. The alternative backup assessment plan is a loess smooth of the spring and fall NEFSC (Bigelow) and NEAMAP fishery-independent surveys.

The AOP and analyst discussed at length the change point analysis that resulted in selection of a recruitment stanza that differed from that selected by the RT analysis. The RT analysis showed a significant change point between 2010 and 2011, while, after addition of two more years of condition data, the updated analysis showed a significant change point between 2009 and 2010 (and thus the start year for the recruitment stanza would now be 2010).

The analyst proposed a Level 2 review to accommodate the proposed changes. According to the 'Description of New England and Mid-Atlantic Region Stock Assessment Process' document, a change in recruitment stanza is associated with a Level 3 review. The AOP discussed this at length and supported the Level 2 recommendation for a number of reasons. The AOP was reassured by the fact that the analyst had done some additional work suggesting the impact of this stanza change was very small (e.g., 1% change in B_{msy} proxy). The AOP described the intent of Level 3 reviews associated with changes in recruitment stanzas likely being related to exploring and understanding impacts when an assessment moves from using a full recruitment time series to a reduced time series, and not the more modest update in the present situation. Additionally, the AOP thought the time allocated to a Level 2 review (1-2 hours) was sufficient for a panel to adequately review the proposed work.

The AOP did raise questions about how best, broadly, to update change point analyses. The AOP noted that in the present case, the change in recruitment stanza was modest and not impactful to reference points, but acknowledged that there could be scenarios where more meaningful changes are seen, and having a set of best practices would be advantageous in dealing with non-trivial cases. The AOP revisited this discussion at the

end of the meeting on 26 February 2024. The AOP discussed that development of best practices for change point analyses as it relates to impacts on reference points could be considered in association with the future RT (2027) scheduled to address projection methodologies.

For the reasons described above, the AOP recommended a Level 2 review for Butterfish.

Atlantic Herring
(AOP Lead: Lisa Kerr)
Recommendation: Level 1 (Direct Delivery)

The 2022 management track assessment for Atlantic herring indicated that the stock is overfished (i.e., 2021 SSB was 21% of the SSBMSY proxy = 185,750 mt) and overfishing is not occurring (i.e., 2021 F was 31% of the FMSY proxy = 0.5). The stock assessment used the Age Structured Assessment Program (ASAP) which was used in previous assessments; however, the methods used to derive biological reference points (BRPs) and conduct short-term projections were changed as part of the management track assessment. Retrospective patterns remain an issue for this assessment and a retrospective adjustment was necessary. The continued poor recruitment of the stock also remains a concern and the SSC has expressed concern that the short-term projections have been consistently overly optimistic for this stock.

The analyst proposed a Level 1 review (direct delivery) for the 2024 management track assessment for Atlantic herring as there are no new sources of information being added to the assessment and no changes anticipated to the model. The management track assessment plan will focus on updating all survey and catch time series through 2023. The spring 2023 NEFSC bottom trawl survey will be missing for this stock but is not expected to be a problem in the context of this analytical model as there will be three surveys available for that year (fall NEFSC bottom trawl survey, acoustic survey, and the shrimp survey). There is also extensive work underway in the Atlantic herring Research Track Stock assessment which is scheduled to be completed in 2026 (catch advice for 2027). This management track assessment will inform catch advice setting for 2025-2027 (with the last year likely being replaced by next assessment).

The AOP supported a Level 1 review for this stock conditional on no additional concerns being identified by the analyst during the update process with the intention being to flag issues early and to elevate to a higher level of review if necessary.

Atlantic Surfclam
(AOP Lead: Mike Celestino)
Recommendation: Level 1 (Direct Delivery)

The current assessment model for Atlantic Surfclam is the 2016 SAW 61 accepted Stock Synthesis model (and a subsequent Level 3 MT reviewed in 2020). The original (February 28th) proposed work plan for the 2024 MT assessment was to update survey and commercial data through 2023 as well as upgrade to the latest version of Stock Synthesis. The alternative assessment plan, should one be necessary, is a swept area biomass estimate from the surfclam survey, adjusted with the median catchability coefficient from depletion studies. At that time, the analyst recommended a Level 1 review.

The AOP commented that B/Bmsy is well above the threshold and that F/Fmsy is well below the threshold, and this has simplified the Mid-Atlantic SSC's task of specifying catch advice. In terms of work for potential inclusion as part of subsequent management or research track assessments, the analyst noted that he could revisit the methods used to derive the trend-based reference points once or if the fishery has a measurable effect on the stock, which at present, it does not. The analyst also did not anticipate any modeling challenges or changes as a function of upgrading the modeling software.

The analyst noted that harvest and discard estimates will come from CAMS and noted that the years that are available for comparison with the commercial fisheries dealer database (CFDERS) are similar.

Since the time of the February AOP meeting the analyst compiled data and completed some preliminary analyses and model runs. This work showed significant trends in Von Bertalanffy (VB) growth curve K parameters for both modeled surfclam regions. Significant trends in VB L_{inf} are already included and modeled in the base model. Additionally, sensitivity runs from the 2020 assessment indicated that inclusion of the additional growth parameter (K) had no notable impact from the final base model output, but did improve some diagnostics. The analyst and AOP agreed at the April 4th 2024 AOP meeting that this proposed change (modeling time varying K) met the criteria for level 2 review [i.e., "Adjustment of method for estimating biological information (growth, maturation, sex ration, changes to length-weight relationships, etc.) when based on methods developed with sufficient peer review or justification for its use"].

In light of the new information, the AOP revised their original recommendation from a Level 1 review for this assessment at their February 28th meeting, to a Level 2 review at their April 4th meeting.

Black Sea Bass
 (AOP Lead: Paul Rago)
Recommendation: Level 2 (Expedited Review)

The MT assessment for Black Sea Bass will be the first assessment following the RT in December 2023. Major innovations introduced in the RT included:

Transition from a statistical catch at age model (ASAP) to a stage-space model with random effects (WHAM). The new model has two spatial components (North and South) with movement between components.

Use of an autoregressive spatial model (VAST) to synthesize age-specific indices from 10 separate surveys into a single set of indices at age.

Introduction of an improved recreational catch-per-angler index.

Incorporation of bottom temperature with random effects as a covariate for recruitment.

Use of random effects for recruitment, numbers at age, and fisheries and index selectivities in the Northern component.
The RT review panel recommended three specific sensitivity analyses to be conducted prior to the MT. The most significant issue related to the use of the VAST model outputs as measures of relative abundance. The stock assessment lead investigated the sensitivity of the model results to deletion of individual indices. These results were reviewed by the AOP to help guide staff work prior to the MT in June 2024. Analyses presented to the AOP suggested that the model was robust to deletion of single indices, except for the MADMF survey in the northern stock area. A potential cause of this effect is that the model was weighting the MA index, which covers a small fraction of the northern stock area, equally with the NEFSC survey that samples a much larger area. Residual patterns for other surveys had undesirable properties, suggesting that the synthesis afforded by the VAST model was more realistic.

Additional questions by the AOP clarified that VAST indices should be robust to missing observations (e.g., missing 2020 and 2023 NEFSC BTS) because it utilizes spatial and temporal patterns explicitly in a hierarchical model.

An important change from the previous assessment model results is the biomass trend in recent years. The ASAP model suggested that the declining abundance of the 2011 and 2015 year classes would lead to rapid stock decline through 2025. The new model, using data through 2021, suggests continued high abundance but no evidence of decline in recent years. Moreover, the population's center of gravity appears to be moving northward. Confirmation of these trends await the MT update that will include survey and catch data through 2023. The Black Sea Bass assessment will be the first assessment in the Northeast to update an assessment based on application of the VAST model. The behavior of the overall model with additional data is unknown, but major changes in perceptions of historical relative abundance are not expected.

The AOP expressed appreciation that a detailed supplemental report describing the responses to the RT peer review recommendations would be prepared as part of the MT. The AOP recommended continued use of VAST for the MT and did not recommend reversion to tuning WHAM with multiple independent indices.

The AOP approved the use of a simplified WHAM model as the backup assessment plan. This approach retains the information from multiple data sources and age data that would otherwise not be considered in an index approach.

Technically, the assessment model to be used in the MT corresponds to the peer-reviewed model in the RT and would qualify as a Level 1 review. After discussion, the consensus opinion of the AOP was that an additional level of review would reduce the uncertainty of the first-time implementation of an assessment that differs quantitatively and qualitatively from its predecessor.

In light of these considerations and the review of necessary follow-up tasks requested by the RT Peer Review Panel, the AOP recommended a Level 2 review.

Eastern Gulf of Maine Atlantic Cod
(AOP Lead: Paul Rago)

Recommendation: Level 3 (Enhanced Review)

The Eastern Gulf of Maine Cod (EGOM) stock is a newly defined stock that previously was part of the Gulf of Maine cod stock. Revised stock definitions were approved in August 2023 at the RT Peer Review. The stock was assessed with a state-space model (WHAM) using data from 1981 to 2021. Per the RT Terms of Reference, stock status was not determined. Estimated biological reference points were based on a MSY proxy based on F40% MSP; the resulting fishing mortality threshold ($F_{40\%}$) estimate is 0.26 and spawning stock biomass target ($SSB_{F40\%}$) is 2,274 mt.

The RT Peer Review recommended incorporation of cod bycatch into the estimate of total catch. This represents a major change in the model configuration evaluated by the Peer Review. Population size is expected to increase because the scale of a population in most assessments is largely determined by the magnitude of total catch. The magnitude of such changes for EGOM cod are unknown. Additional requests by the RT Peer Review to modify the stock assessment model, particularly those related to consideration of survey indices were considered a lower priority and cannot be addressed in this MT. These requests will be handled incrementally in future MT assessments.

The AOP appreciated the inclusion of lobster discard data, but noted the difficulties of hindcasting historical data based on only recent samples. About 20 years of observer and harvester reported data are available. Total discards will be assumed to be proportional to lobster landings. Details on the hindcasting method have not been developed to estimate discards prior to 2000. It was noted that substantial differences in the lobster fishery have occurred in the past decade. Such changes may require modification of the assumptions used to hindcast historical discard data. Ideally, the estimation will generate age-specific discard estimates in the lobster fishery. Assumptions about the discard survival rates will also be a concern.

The effects of introducing a revised catch time series on the behavior of the assessment model are unknown. Alternative decisions regarding weighting of model parameters and survey indices (generalized tweaking) may be required. This MT will be the first assessment to assign status determinations (i.e., overfishing/overfished) to this newly defined stock.

The AOP expressed concern about the number of biological samples collected by Port Agents, particularly at finer spatial scales. The partitioning of historical data from the former Gulf of Maine stock into the new Eastern and Western GOM stocks reduces the number of samples historically. These concerns have been addressed in the RT Peer

Review. However, recent reductions in numbers of port samples in the Northeast could pose problems for support of many stock assessments, particularly those with smaller spatial domains.

As a backup assessment, the assessment lead proposed the use of a simplified WHAM model with few or no random effects. This approach was considered sound by the AOP because it retained the information content from multiple surveys and the age composition of the indices and catch.

In view of the significant changes proposed, the AOP recommended a Level 3 review.

Western Gulf of Maine Atlantic Cod
(AOP Lead: Kristan Blackhart)
Recommendation: Level 2 (Expedited Review)

The Western Gulf of Maine (WGOM) Cod stock represents a new stock definition, although there is a good deal of overlap with the previously defined Gulf of Maine stock. This is the first MT assessment for this stock following the Atlantic Cod RT assessment in 2023. The assessment methods adopted for this stock in the RT used the WHAM configured with two fishery fleets (commercial landings/discards, recreational landings/discards), three fishery selectivity blocks, and a lifetime estimate of natural mortality (M=0.2) based on a suite of life history based estimators developed during the RT. The proposed assessment approach for the June MT uses the current WHAM configuration with no changes and all fishery and survey data updated through 2023. Commercial data updates for 2022 and 2023 will be pulled from CAMS. As recommended by the RT review panel, the backup method uses a simplified WHAM model with fewer random effects. At the February AOP meeting, the assessment lead noted that if time permits, additional explorations into the projection methodologies may be explored for this assessment as suggested by the RT review recommendations. It was unknown at that time whether those explorations would be incorporated into the June MT or lead to any changes in the projection methods or biological reference points. Other explorations suggested by the RT review panel for this stock (i.e., decoupling process error across ages, examining selectivity configurations, modeling recruitment) will likely be deferred until later MTs due to time constraints.

After the April AOP meeting, the assessment lead submitted a request to the AOP to increase the review level based on his work on the assessment projection

methodologies, as noted above. WGOM Cod uses the standard projection and reference point methodology found in WHAM. However, there are two issues associated with the standard method for the stock: 1) long-term projections at F_{MSY} do not result in SSB reaching SSB_{MSY}; and 2) projections do not include the effect of new process errors in the projection period. Since the February AOP meeting, the assessment lead has spent time exploring a potential solution to both of these issues which is to include process error in the projections (i.e., perform stochastic projections), and use those long-term stochastic projections to estimate SSB_{MSY}. This work will be presented alongside the standard approach at the Management Track, and could lead to changes in biological reference points for the stock.

At the time of initial discussion in February, no changes to the model or data inputs were planned and the lead analyst suggested a Level 2 review because this is a newly defined stock. The AOP agreed with that suggestion and noted this assessment will support a first-time status determination, justifying a Level 2 review. After the April request from the assessment lead, the AOP agreed that elevating this to a Level 3 review is appropriate.

Georges Bank Atlantic Cod
(AOP Lead: Lisa Kerr)
Recommendation: Level 2 (Expedited Review)

A research track stock assessment for Georges Bank cod was concluded in 2023. This assessment process resulted in a change in the spatial scale of management for cod and a shift from application from an empirical approach (i.e., Ismooth method) to an analytical model, the Woods Hole Assessment model, for the Georges Bank cod stock. This is the first management track assessment that will apply the approved methods coming out of the research track assessment process.

Between the conclusion of the research track and the management track errors were identified that will be addressed in the management track assessment. It was found that the survey data used in the research track assessments for GB Atlantic Cod did not incorporate the Bigelow:Albatross calibration. In addition, an issue with the specification of equilibrium initial numbers at age and the weight-at-age input were also identified as needing correction. These corrections were made and a rerun of the model conducted to evaluate the impact. The model return indicated that the corrections resulted in the anticipated adjustment of results and improved model diagnostics.

During the February 28, 2024 meeting of the AOP, the analyst proposed a Level 2 review (expedited review) for the 2024 management track assessment for Georges Bank Atlantic cod. There are multiple changes being made in this management track: 1) calibration NEFSC survey indices, 2) correction to equilibrium initial NAA, 3) correction weight-at-age input. However, work to date indicates that the cumulative impact of these changes improved diagnostics and resolved areas of concern from the Research Track. There are no new sources of information being added to the assessment and no additional changes anticipated to the model. The management track assessment plan will focus on updating all survey and catch time series through 2023. The NEFSC 2023 spring survey is missing due to sampling interruption (could result in only 1 spring survey in 2023) and the DFO spring survey is missing in 2022. The missing surveys will be treated as missing in the model and is not expected to be a significant problem in the context of this analytical model as there will be at least one spring survey available for each year.

The AOP reconvened on April 4, 2024 and discussed new information needs regarding the Georges Bank cod stock assessment. Due to recent changes in the TRAC process there is now a need for spatial allocation of the biomass estimated for the full Georges Bank stock area to the eastern Georges Bank transboundary area. The analyst plans to deliver this information and suggested elevation of the review from a Level 2 to a Level 3 due to the current uncertainty around the approach that will be taken.

The AOP supports a Level 3 review for this stock due to the multiple changes being made to the stock assessment, as well as the request for analysis on spatial allocation of biomass estimated for the U.S. Georges Bank stock area to the eastern Georges Bank transboundary area.

Southern New England Atlantic Cod
(AOP Lead: Kristan Blackhart)
Recommendation: Level 3 (Enhanced Review)

As with the other cod stocks, Southern New England Cod was redefined in the last RT and this represents the first MT for this stock using the new stock definition. The current model from the 2023 RT uses WHAM based on data inputs from 1981-2021. Natural mortality is age based and catch-at-age data is available for commercial and recreational landings/discards. The RT assessment utilized two abundance indices: NEFSC spring and a recreational LPUE index from 1996-2021. The planned MT assessment does not include any new data inputs beyond updating fishery and survey

data through 2023, but does change the WHAM configuration by addressing a major RT review recommendation and developing a new recreational CPUE index that includes zero observations to replace the existing LPUE index. Development of this CPUE index is pending based on ongoing exploration of various methodologies, so the overall impact of the inclusion of this new index on the assessment is hard to predict. The lead analyst suggested a

Level 3 review, in line with the NRCC process document requirements for an Enhanced Review when assessments include a new or alternate interpretation of an index. The AOP concurs with a Level 3 review.

Golden Tilefish
(AOP Lead: Kristan Blackhart)
Recommendation: Level 3 (Enhanced Review)

The previous assessment model for Golden Tilefish used ASAP and was last updated in the 2021 MT. A RT for the stock was completed in March 2024 exploring the use of WHAM, using the same data as the 2021 MT assessment (terminal year 2020). The RT working group found that the WHAM model results for this data limited stock were sensitive to the inclusion of random effects. Models without random effects estimate large domes at the end of the time series, producing results similar to the 2021 MT ASAP assessment with reasonable diagnostics. Adding additional random effects to the model provides relative improvements to the diagnostics while also estimating a flattening of the selectivity curve, reducing the cryptic biomass in the 10+ age group through estimation of larger reductions in biomass since the development of the directed longline fishery in the 1970s and relatively lower rebuilding of the stock after management was implemented.

The RT also developed several new time series of data for Golden Tilefish that were not included in the RT WHAM model, but will be added for the MT in addition to updating all data series to 2023. Additionally, the RT review panel suggested additional exploration of WHAM model configurations to optimize model performance, which will be done as necessary as new data is added to the base configuration. Because of the uncertainty present in this assessment, the AOP discussed the assessment lead providing information on the implications of alternative models to provide additional context for the SSC in producing cautionary management advice.

Because of the extensive review required for this assessment following the RT, the lead analyst suggested a Level 3 review.

The AOP concurs with a Level 3 review.

Meeting Conclusions

The AOP met on February 28 and April 4, 2024 to review the stock assessment plans for 9 stocks scheduled for the June 2024 Management Track cycle. Additional discussion occurred asynchronously following the April meeting. The panel concluded that a Level 1 review (Direct Delivery) was warranted for Atlantic Herring; Level 2 reviews (Expedited Review) for Atlantic Surfclam, Butterfish, and Black Sea Bass; and Level 3 review (Enhanced Review) for Eastern Gulf of Maine Cod, Georges Bank Cod, Golden Tilefish, Southern New England Cod, and Western Gulf of Maine Cod. The Level 2 and 3 reviews will occur during the June 2024 Management Track Peer Review scheduled for June 18-21 and June 24-28, 2024. Any additional changes in the required review level would be triggered by a Northeast Fisheries Science Center request to increase the review level for a given stock. The AOP could concur to increase the review level via email or request to reconvene the AOP panel to have further discussions with the stock assessment lead. Any need to reconvene the panel would be a publicly announced meeting and any subsequent changes to the review level would be publicized to assessment partners and stakeholders.

Appendix B. Assessment Oversight Panel Meeting participants (names only, no call-in numbers).

Kristan Blackhart, AOP Chair (NEFSC)
Paul Rago, AOP (MAFMC)
Mike Celestino, AOP (ASMFC)
Lisa Kerr, AOP (NEFMC)
Michele Traver - NEFSC
Alan Bianchi - North Carolina Division of Marine Fisheries
Alex Dunn - NEFSC
Alex Hansell - NEFSC
Amanda Hart - NEFSC
Andrew Jones - NEFSC
Angela Forristall - NEFMC staff
Anna Mercer - NEFSC
Anthony Wood - NEFSC
Ashley Asci - GARFO
Brad Schondelmeier - MADMF
Brandon Muffley - MAFMC staff
Brian Linton - NEFSC
Cameron Hodgdon - NEFSC
Carrie Nordeen - GARFO
Cate O'Keefe - NEFMC Executive Director
Charles Adams - NEFSC
Charles Perretti - NEFSC
Chengxue li - NEFSC
Chris Kellogg - NEFMC staff
Chris Legault - NEFSC
Conor Davis - New Jersey Department of Environmental Protection
Dan Hennen - NEFSC
Dave McElroy - NEFSC
Emilie Franke - ASMFC staff
Emily Bodell - NEFMC staff
Emily Keiley - GARFO
Emily Liljestrand - NEFSC
Gareth Lawson - Conservation Law Foundation
Jacqueline Odell - Northeast Fisheries Coalition
Jamie Cournane - NEFMC staff
Jason Boucher - NEFSC
Jeff Kaelin - Lund's Fisheries

Jessica Blaylock - NEFSC
Jessica Coakley - MAFMC staff
John Pappalardo - Cape Cod Commercial Fisherman's Alliance
John Wiedenmann - Rutgers University
Jon Deroba - NEFSC
Joseph Meyers - ASMFC staff
Julia Beaty - MAFMC staff
Julie Nieland - NEFSC
Kai Lorenzen - University of South Florida
Kathy Sosebee - NEFSC
Kiersten Curti - NEFSC
Kiley Dancy - MAFMC staff
Larry Alade - NEFSC
Libby Etrie - Conservation Law Foundation
Liz Sullivan - GARFO
Mark Grant - GARFO
Mary Sabo - MAFMC staff
Melanie Griffin - MADMF
Olaf Jensen - University of Wisconsin (Madison)
Paul Nitschke - NEFSC
Rebecca Peters - Maine Department of Marine Resources
Rick Bellavance - NEFMC Council Member
Robin Frede - NEFMC staff
Sam Truesdell - NEFSC
Samantha Tolken - GARFO
Sefatia Romeo Theken - Deputy Commissioner for MA Fisheries and Game
Spencer Talmage - GARFO
Steve Cadrin - SMAST
Susan Wigley - NEFSC
Tara Dolan - MADMF
Thomas Alspach - Sea Watch International
Tracey Bauer - North Carolina Division of Marine Fisheries
Will Poston - American Saltwater Guides Association

Key:
ASMFC - Atlantic States Marine Fisheries Council
GARFO - Greater Atlantic Regional Fisheries Office
MADMF - Massachusetts Division of Marine Fisheries
MAFMC - Mid-Atlantic Fisheries Management Council

NEFMC - New England Fisheries Management Council
NEFSC - Northeast Fisheries Science Center
SMAST - University of Massachusetts School of Marine Science and Technology
Appendix 2. Acronyms Used in This Report
AOP - Assessment Oversight Panel
ASAP - Age Structured Assessment Program
CAMS - Catch Accounting and Monitoring System
CFDERS -
CPUE - catch per unit of effort
EGOM - Eastern Gulf of Maine
FE - fixed effects
LPUE - landings per unit of effort
MAFMC - Mid-Atlantic Fishery Management Council
MSP - maximum spawning potential
MT - Management Track
NEAMAP - NorthEast Area Monitoring and Assessment Program
NEFMC - New England Fishery Management Council
NEFSC - Northeast Fisheries Science Center
NRCC - Northeast Region Coordinating Council
RE - random effects
RT - Research Track
SAW - Stock Assessment Workshop
SSC - Scientific and Statistical Committee
VAST - Vector Autoregressive Spatio-Temporal model
WHAM - Woods Hole Assessment Model

Appendix C. Management Track Stock Assessment Terms of Reference

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

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

3. Estimate annual fishing mortality, recruitment, and stock biomass (both total and spawning stock) as possible (depending on the assessment method) for the time series using the approved assessment method and estimate their uncertainty. Include retrospective analyses if possible (both historical and within-model) to allow a comparison with previous assessment results and projections, and to examine model fit.

    a. Include bridge runs to sequentially document each change from the previously accepted model to the updated model proposed for this peer review.

    b. Prepare a backup assessment approach that would serve as an alternative for providing scientific advice to management if the analytical assessment were to not pass review.

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

5. Conduct short-term stock projections when appropriate.

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

[*] Major changes from the previous stock assessment require pre-approval by the Assessment Oversight Panel.

Appendix D. June 2024 Management Track Peer Review Meeting Attendees

ASMFC - Atlantic States Marine Fisheries Commission
CLF - Conservation Law Foundation
DFO - Fisheries and Oceans Canada
GARFO - Greater Atlantic Regional Fisheries Office
GMRI - Gulf of Maine Research Institute
MA DMF - Massachusetts Division of Marine Fisheries
MAFMC - Mid-Atlantic Fisheries Management Council
MD DNR - Maryland Department of Natural Resources
ME DMR - Maine Department of Marine Resources
NEFMC - New England Fisheries Management Council
NEFSC - Northeast Fisheries Science Center
NC DMF - North Carolina Division of Marine Fisheries
NJ DEP - New Jersey Department of Environmental Protection
NYSDEC - New York State Department of Environmental Conservation
SMAST - University of Massachusetts School of Marine Science and Technology
TNC - The Nature Conservancy
UMASS - University of Massachusetts
~~~~~~

Week 1
John Wiedenmann - Chair
Kai Lorenzen - Panel
Alexei Sharov - Panel

Kristan Blackhart - NEFSC
Michele Traver - NEFSC

Al Catone - Captain
Alain d'Entremont - Scotia Harvest President
Alex Hansell - NEFSC
Alex Dunn - NEFSC
Amanda Hart - NEFSC
Andrew Jones - NEFSC
Andrew Allyn - GMRI
Angelia Miller - SMAST
Anjali Bhardwaj - GARFO
Anooshka Sethi - Colby College (Student)
Ashley Asci - GARFO
Aubrey Ellertson - Chatham Fish Pier
Bart DiFiore - GMRI

Benjamin Galuardi - GARFO
Benjamin Levy - NEFSC
Ben Martens - Maine Coast Fishermen's Association
Beth Casoni - MA Lobstermen's Association Executive Director
Brian Linton - NEFSC
Cameron Hodgdon - NEFSC
Carolyn Iwicki - GARFO
Cate O'Keefe - NEFMC Executive Director
Charles Adams - NEFSC
Charles Perretti - NEFSC
Chris Legault - NEFSC
Corrin Flora - NC DMF
Dan Hennen - NEFSC
Daniel Caless - GARFO
Daniel Hocking - GARFO
Daniel Salerno - NEFMC Member
Dave McElroy - NEFSC
Emily Bodell - NEFMC Staff
Emily Liljestrand - NEFSC
Gareth Lawson - CLF
Gavin Fay - SMAST
Geoffrey Smith - TNC
Hank Soule -
Irene Andrushanko - Canada DFO
J. Michael Lanning - GARFO
Jacqueline ODell - Northeast Seafood Coalition
Jamie Behan - GMRI
Jamie Cournane - NEFMC Staff
Janice Plante - NEFMC Staff
Jason Boucher - NEFSC
Jennifer Couture - NEFMC Staff
Jerry Leeman - New England Fishermen's Stewardship Association
John Pappalardo - Cape Cod Commercial Fishermen's Alliance
Jon Deroba - NEFSC
Kailee Berge - U. Maine
Karen Greene - Office of Science and Technology
Kathy Cooper-MacDonald - Canada DFO
Kiersten Curti - NEFSC
Kristopher Winiarski - GARFO
Larry Alade - NEFSC

Laura Smith - GARFO
Libby Etrie - CLF
Lindsey Nelson - NEFSC
Lisa Kerr - U. Maine
Liz Sullivan - GARFO
Margaret Conroy - ASMFC
Mark Grant - GARFO
Mary Hudson - Maine Coast Fishermen's Association
Megan Ware - ME DMR
Melanie Barrett - Canada DFO
Melanie Griffin - MADMF
Meredith Mendelson - ME DMR
Mike Celestino - NJ DEP
Mike Simpkins - NEFSC
Nicole Morgan - GARFO
Paul Nitschke - NEFSC
Rebecca Peters - ME DMR
Rick Bellavance - NEFMC Member
Robin Frede - NEFMC Staff
Robyn Linner - U. Maine
Roger Brothers - GMRI
Roksanna Keyvan - Wake Forest University (Student)
Sara Turner - GARFO
Sarah Robinson - Gloucester Fisheries Project
Sefatia Romeo Theken - MA Department of Fish and Game
Spencer Talmage - GARFO
Steve Cadrin - SMAST
Susan Wigley - NEFSC
Tara Dolan - MADMF
Thomas Heimann - GARFO
Tim Miller - NEFSC
Togue Brawn - Captain
Tony Wood - NEFSC
Vito Giacalone - Northeast Seafood Coalition
Will Helt - TNC

Week 2
Paul Rago - Chair
Ed Camp - Panel
Yan Jiao - Panel

Katie Drew - Panel

Kristan Blackhart - NEFSC
Michele Traver - NEFSC

Alan Bianchi - NC DMF
Alex Dunn - NEFSC
Alex Hansell - NEFSC
Alexei Sharov - MD DNR
Amanda Hart - NEFSC
Andrew Jones - NEFSC
Benjamin Levy - NEFSC
Ben Wasserman - Delaware Division of Fish and Wildlife
Brandon Muffley - MAFMC Staff
Brian Linton - NEFSC
Cameron Hodgdon - NEFSC
Carolyn Iwicki - GARFO
Charles Adams - NEFSC
Charles Perretti - NEFSC
Chris Legault - NEFSC
Dan Hennen - NEFSC
Daphne Munroe - Rutgers University
Douglas Potts - GARFO
Elise Koob - MADMF
Emily Keiley - GARFO
Emily Liljestrand - NEFSC
Gavin Fay - SMAST
Greg DiDomenico - Lund's Fisheries
Jason Boucher - NEFSC
Jason Didden - MAFMC Staff
Jeff Kipp - ASMFC
Jessica Coakley - MAFMC Staff
Joe Myers - Sea Watch International
Jon Deroba - NEFSC
John Maniscalco - NYSDEC
Jose Montanez - MAFMC Staff
Julia Beaty - MAFMC
Kathy Sosebee - NEFSC
Kiersten Curti - NEFSC
Laura Quaglia - University of Texas (Student)

Laurel Smith - NEFSC
Laurie Nolan - Industry Representative
Lindsey Nelson - NEFSC
Liz Brooks - NEFSC
Margaret Conroy - ASMFC
Mike Celestino - NJ DEP
MIke Simpkins - NEFSC
Paul Nitschke - NEFSC
Peter Clarke - NJ DEP
Sara Turner - GARFO
Steve Cadrin - SMAST
Susan Wigley - NEFSC
Tim Miller - NEFSC
Toni Chute - NEFSC
Tony Wood - NEFSC
Tracey Bauer - ASMFC
Wendy Gabriel - NEFSC (retired)

Appendix E. Realized Agenda for June 2024 Management Track Peer Review

June Management Track Peer Review Meeting
June 18-25, 2024

Google Meet joining info: meet.google.com/cvj-xzxh-vuj
Or dial: (US) +1 505-596-1588 PIN: 594 430 759#

AGENDA (v. 6/14/2024)

*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.

Tuesday, June 18, 2024

Time	Subject	Presenter
9:30 a.m. - 9:45 a.m.	Welcome/Logistics/Conduct of Meeting	Michele Traver, Kristan Blackhart, John Wiedenmann, Chair
9:45 a.m. - 11:15 a.m.	WGOM Cod	Charles Perretti
11:15 a.m. - 11:30 a.m.	Break	
11:30 a.m. - 12:30 p.m.	WGOM Cod cont. Discussion/Questions	Charles Perretti Panel
12:30 p.m. - 1:00 p.m.	Public Comment	Public
1:00 p.m. - 2:00 p.m.	Lunch	
2:00 p.m. - 3:30 p.m.	EGOM Cod	Cameron Hodgdon
3:30 p.m. - 3:45 p.m.	Break	
3:45 p.m. - 4:45 p.m.	EGOM Cod cont. Discussion/Questions	Cameron Hodgdon Panel
4:45 p.m. - 5:00 p.m.	Daily Wrap Up Summary/Discussion	Panel
5:00 p.m. - 5:30 p.m.	Public Comment	Public

Time	Subject	Presenter
5:30 p.m.	Adjourn	

Thursday, June 20, 2024

Time	Subject	Presenter
9:30 a.m. - 9:35 a.m.	Welcome/Logistics	Michele Traver, John Wiedenmann, Chair
9:35 a.m. - 11:00 a.m.	SNE Cod	Alex Hansell
11:00 a.m. - 11:15 a.m.	Break	
11:15 a.m. - 12:15 p.m.	SNE Cod cont. Discussion/Questions	Alex Hansell Panel
12:15 p.m. - 12:45 p.m.	Public Comment	Public
12:45 p.m. - 1:45 p.m.	Lunch	
1:45 p.m. - 3:15 p.m.	GB Cod	Amanda Hart
3:15 p.m. - 3:30 p.m.	Break	
3:30 p.m. - 4:30 p.m.	GB Cod cont. Discussion/Questions	Amanda Hart Panel
4:30 p.m. - 4:45 p.m.	Daily Wrap Up Summary/Discussion	Panel
4:45 p.m. - 5:15 p.m.	Public Comment	Public
5:15 p.m.	Adjourn	

Friday, June 21, 2024

Time	Subject	Presenter
9:30 a.m. - 9:35 a.m.	Welcome/Logistics	Michele Traver, John Wiedenmann, Chair
9:35 a.m. - 11:00 a.m.	Meeting Wrap Up/Key	Panel

Time	Subject	Presenter
	Points	
11:00 a.m. - 4:00 p.m.	Report Writing	Panel
4:00 p.m.	Adjourn	

Monday, June 24, 2024

Time	Subject	Presenter
1:00 p.m. - 1:15 p.m.	Welcome/Logistics/Conduct of Meeting	Michele Traver, Kristan Blackhart, Paul Rago, Chair
1:15 p.m. - 3:15 p.m.	Black Sea Bass	Emily Liljestrand
3:15 p.m. - 3:30 p.m.	Break	
3:30 p.m. - 4:30 p.m.	Black Sea Bass cont. Discussion/Questions	Emily Liljestrand Panel
4:30 p.m. - 4:45 p.m.	Daily Wrap Up Summary/Discussion	Panel
4:45 p.m. - 5:00 p.m.	Public Comment	Public
5:00 p.m.	Adjourn	

Tuesday, June 25, 2024

Time	Subject	Presenter
8:30 a.m. - 8:35 a.m.	Welcome/Logistics/Conduct of Meeting	Michele Traver, Paul Rago, Chair
8:35 a.m. - 10:30 a.m.	Golden Tilefish	Paul Nitschke
10:30 a.m. - 10:45 a.m.	Break	
10:45 a.m. - 11:45 a.m.	Golden Tilefish cont. Discussion/Questions	Paul Nitschke Panel

Time	Subject	Presenter
11:45 a.m. - 12:00 p.m.	Public Comment	Public
12:00 p.m. - 1:00 p.m.	Lunch	
1:00 p.m. - 2:30 p.m.	Butterfish Discussion/Questions	Charles Adams Panel
2:30 p.m. - 2:45 p.m.	Break	
2:45 p.m. - 4:15 p.m.	Atlantic Surf Clam Discussion/Questions	Dan Hennen Panel
4:15 p.m. - 4:30 p.m.	Daily Wrap Up Summary/Discussion	Panel
4:30 p.m. - 5:00 p.m.	Public Comment	Public
5:00 p.m.	Adjourn	

Monday, July 8, 2024

Time	Subject	Presenter
12:30pm - 12:40pm	Welcome, Logistics, and Meeting Conduct	Michele Traver, Kristan Blackhart, John Wiedenmann (Chair)
12:40pm - 12:50pm	Confirm GB Cod Assessment	Amanda Hart, assessment lead Panel
12:50pm - 1:10pm	Confirm SNE Cod Assessment Projections	Alex Hansell, assessment lead Panel
1:10pm - 1:40pm	Confirm WGOM Cod Assessment Results & Diagnostics	Charles Perretti, assessment lead Panel
1:40pm - 2:20pm	Confirm EGOM Cod Assessment Results & Diagnostics	Cameron Hodgdon, assessment lead Panel
2:20pm - 2:30pm	Meeting Wrap Up/Key Points	Panel
2:30pm	Adjourn	

Appendix F. References

Deroba, J and A. Schueller. 2013. Performance of Stock Assessment with Misspecified Age- and Time-Varying Natural Mortality. Fisheries Research 146:27–40.

Lorenzen, K. (1996), The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49: 627-642. https://doi.org/10.1111/j.1095-8649.1996.tb00060.x

Turner, S.C. 1986. Population dynamics of, and impact of fishing on tilefish, Lopholatilus chamaeleonticeps, in the Middle Atlantic-Southern New England region during the 1970's and early 1980's. New Brunswick, N.J.: Rutgers University. Ph.D. dissertation.