Summary Report of the Atlantic Herring Research Track Stock Assessment Peer Review

March 10-14, 2025
Northeast Fisheries Science Center, Woods Hole, Massachusetts

Report prepared by Peer Review Panel Members:
Dr. Conor McManus (Chair), NEFMC SSC
Dr. Henrik Sparholt, independent contractor with CIE
Dr. Florian Berg, independent contractor with CIE
Dr. Gary Melvin, independent contractor with CIE

Introduction

The Northeast Region Coordinating Council (NRCC)[1] has developed an enhanced stock assessment process to improve the quality of assessments. The process involves two tracks of assessment work: 1) a Management Track (MT) that includes routine updates of previously approved assessment methods to support regular management actions (e.g., annual catch limits), and 2) a Research Track (RT) that allows comprehensive research and development of improved assessments on a stock-by-stock or topical basis. The RT assessment process allows for a more thorough review of information available and for the evaluation of different assessment approaches than would be possible in a standard stock assessment process where the results are immediately needed for management advice. RT assessments are followed by a MT assessments; occasionally, there is the ability to improve a RT assessment before or during a subsequent the MT assessment. This Peer Review Panel (hereafter Panel) reviewed the RT Assessment for the Atlantic herring stock in United States waters.

Atlantic herring was first assessed using Virtual Population Analysis (VPA) in the 1960s-1970s. Since 2009, the stock has been assessed using the Age Structured Assessment Program (ASAP) developed by the Northeast Fisheries Science Center (NEFSC). The last benchmark stock assessment was conducted in 2018, with updates occurring every other year since, and the most recent MT assessment was in 2024 using ASAP. The primary goal of the RT assessment for Atlantic herring was to develop an age-structured state space assessment model to allow for estimating process and observation error and considering environmental covariates within the model.

The Atlantic Herring RT Working Group (WG) used the Woods Hole Assessment Model (WHAM) platform to develop a state space model to be used in future MT assessments of the stock. The WG was chaired by Matt Cieri (MEDMR) and included staff from NEFMC, GARFO,

[1] Atlantic States Marine Fisheries Commission (ASFMC), Greater Atlantic Regional Fisheries Office (GARFO), Mid-Atlantic Fishery Management Council (MAFMC), New England Fishery Management Council (NEFMC), and Northeast Fisheries Science Center (NEFSC).

Massachusetts Division of Marine Fisheries (MADMF), University of Massachusetts – Dartmouth School for Marine Science and Technology (SMAST), and Rutgers University. Terms of Reference (TORs) for the WG are provided in Appendix 1.

The Atlantic Herring RT assessment peer review took place in Woods Hole, MA during March 10-14, 2025 using a hybrid meeting platform. The Panel included three independent scientists selected by the Center for Independent Experts (CIE): Henrik Sparholt, Florian Berg, and Gary Melvin. The Panel was chaired by Conor McManus (Member of the NEFMC Scientific and Statistical Committee). The agenda is located in Appendix 2, and attendees in Appendix 3.

The RT WG Assessment Report and supporting working papers were provided to the Panel via the NEFSC data portal (https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php) on February 21, 2025. Additional information requested by panelists was provided during the peer review and uploaded to the data portal. Listed items are in Appendix 4. Prior to the meeting, members of the Panel met with Brian Hooper (NEFSC's Stock Assessment Process Lead) and Kristan Blackhart (Chief, NEFSC Population Dynamics Branch) on March 6, 2025 to review and discuss the meeting agenda, reporting requirements, meeting logistics and the overall process.

In addition to the report and working papers, presentations by the WG members and WG Chair-invited participants were made available on the data portal during the RT Review. WG members were present during the meeting to answer questions from the Panel, and NEFSC staff served as rapporteurs to capture conversations between the Panel and the WG, Panel consensus on various topics, and Panel recommendations. Panelists led drafting individual portions of this report, but the entire Panel edited and reviewed the report such that it represents the consensus views of the Panel. The Panel Chair compiled and edited this Summary Report with assistance from all panelists before submission of a draft report to the WG Chair and WG Lead Analyst. CIE Panelists will submit separate individual reviewer reports to the CIE.

The Panel agreed that all eight TORs were adequately addressed. With specific regards to the transition from ASAP to WHAM, the WG presented a series of WHAM runs testing the sensitivity of the model to various configurations and life history assumptions. The final candidate model recommended by the WG included many of the same data inputs as the previous 2024 MT ASAP configuration (but for select elements outlined in the main WG report); however, the model now includes random effects being applied in the model, most notably on the numbers-at-age (NAA) transitions. After reviewing the various model variants the WG provided, and further investigation of the diagnostics and outputs from the models, the Panel agrees that WHAM is an appropriate tool for the Atlantic herring assessment and offers a scientific improvement over ASAP. Prior to the next MT assessment for Atlantic herring, the Panel provided several recommendations for consideration that strive for model diagnostic improvements, improved biological plausibility, and a deeper understanding of WHAM given its relatively new nature. Of the recommendations provided by the Panel, particular importance was placed on better understanding the impacts of the NAA random effects, including the stability of NAA estimates with additional years of data added, through peel analyses, and their influence or utility in projections. While WHAM reduces retrospective patterns compared to the ASAP for this stock, the uncertainties around the projection methods need to be further investigated, particularly as they relate to projections beyond the bridge year and subsequent year. The Panel commended the WG for the substantial amount of work made in moving to the state space WHAM platform for the Atlantic herring stock and for the collegial nature of the review.

The Panel's evaluation of the WG's responses to the eight TORs is provided below.

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

The Panel concluded that this TOR was adequately addressed.

Significant progress has been made in the identification of ecosystem drivers that can influence Atlantic herring abundance and distribution. The findings in TOR 1 were addressed and further considered in TOR 4 and TOR 7. Uncertainty was documented through the analyses, both within TOR 1 via analyses and a literature review, and through evaluation of TOR 1 elements in TOR 4 (i.e. WHAM sensitivity runs).

The WG provided a thorough review of Atlantic herring life history and ecosystem drivers on the species (as documented in the literature) through an updated Ecosystem and Socioeconomic Profile (ESP). The ESP leveraged essential fish habitat and life history information previously collated and gathered new, more contemporary research. The ESP was instrumental in understanding the processes that influence population dynamics that could be evaluated as environmentally explicit relationships in WHAM. These endeavors led to a conceptual life history model that outlined the various ecosystem influences on the species in this region, with hypotheses for the impact of specific conditions on each life stage (Figure 1). This conceptual model was then used to identify possible environmental datasets for further exploration.

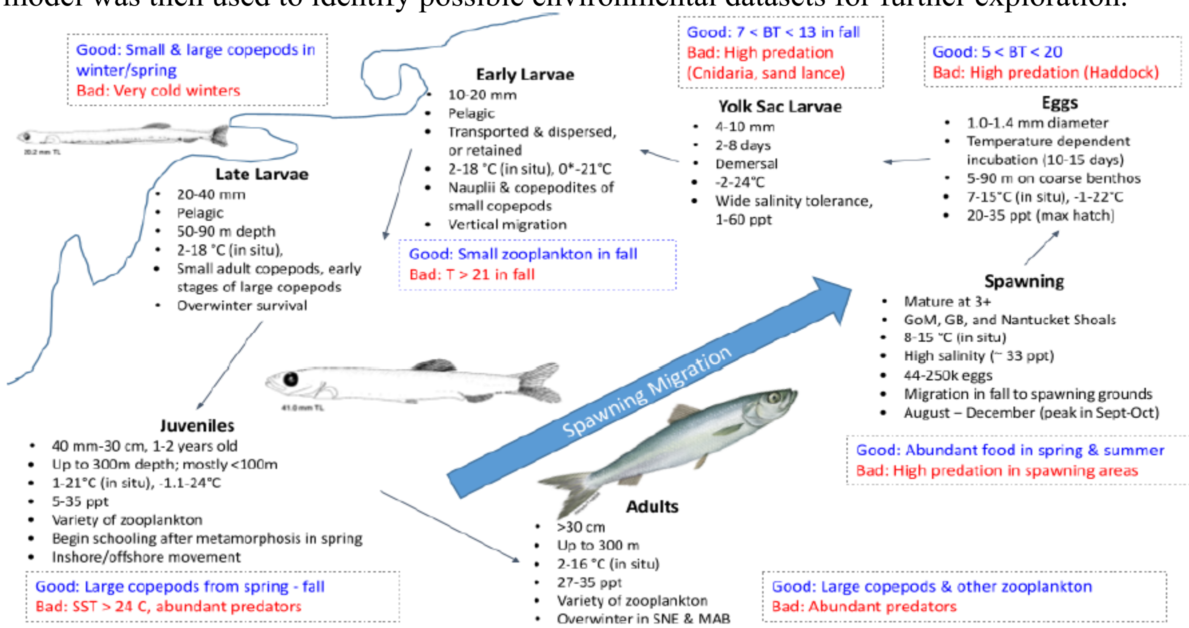

Figure 1. Life history diagram and ecosystem influence hypotheses for Atlantic herring as presented by the WG and in the assessment report.

The WG used boosted regression trees (BRTs) to understand which environmental factors influence Atlantic herring, and their relative importance compared to each other. Several covariates were explored (e.g. copepods, haddock, temperature indices) to understand influences on Atlantic herring (e.g. starvation/food availability, predation pressure, thermal mortality). The response variable for the BRTs was Atlantic herring recruitment estimated from the previous ASAP run to help guide how these covariates could be used in conjunction with recruitment estimates in WHAM.

The Panel commended the WG for their efforts within the ESP, as well as using both hypotheses for impacts on herring recruitment to guide model prioritization runs and ensure biological plausibility. The use of a tool (e.g. BRTs) was also appreciated to determine which variables were most sensible for testing within WHAM.

The Panel recommended several future advancements for this work, including modifications to copepod and haddock predation indices to refine the indices and better capture the interactions between these factors and herring. The Panel also suggested considering competition metrics (e.g. other forage fish abundance) in the future given the abundance and spatiotemporal changes in certain species that have been captured in the literature and reflected in the stakeholders' comments collected via this assessment. The environmental covariates constructed could also be used to understand other processes within the assessment model (e.g. catchability and selectivity, natural mortality, growth) as deemed appropriate in the future.

The Panel provided two substantive comments of caution for the BRTs moving forward. The first concern was with regards to the BRT method. It was unclear from the Panel whether the BRT was the best tool for this prioritization exercise compared to other methods (e.g. principal component analysis, dynamic factor analysis, other regression approaches). Testing WHAM and ASAP recruitment data for the BRTs suggested that variable importance conclusions may be subject to input data with the BRTs (Figure 2), and testing its robustness against other methods would be beneficial in the future. The second concern posed was the use of assessment model output (i.e. Age-1 Atlantic herring estimates) for these analyses. Cautioned has been raised when conducting post-hoc analyses using data output as they often overlook the uncertainty in the assessment products themselves (Brooks and Deroba, 2015; Dickey-Collas et al., 2015). The differences in BRT results from using WHAM and ASAP-predicted recruitment also suggest this concern, in addition to the rather circular nature of these analyses (i.e. use WHAM estimated recruitment to pick an environmental covariate to include in modeling for revised WHAM predicted recruitment). Efforts should be made to analyze Age-1 abundance indices from other sources (e.g. state surveys, seabird diet data) to avoid this circularity.

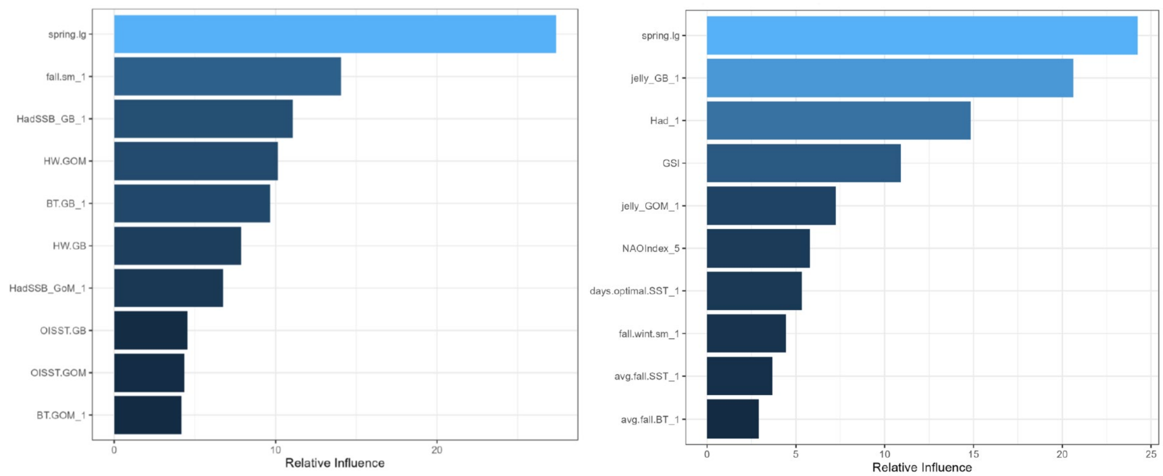

Figure 2. Boosted regression tree results for environmental variables explaining Atlantic herring recruitment estimates produced from ASAP (left) and WHAM (right). Figures were presented during the peer review.

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

The Panel concluded that this TOR was adequately addressed.

The information presented by the WG provided a good and general summary of the landings over time, by region, and gear type given the historical extent of the fishery and the magnitude of catches. They identified potential uncertainty associated with the different reporting mechanisms and how estimates of discards were integrated into the estimates of landings over time. Over the years, the authorities responsible for the data collection have utilized several approaches to account for fishing mortality (catch and discards) from all sources. The implementation of CAMS in 2020 represents an improved methodology in this process and reflects the sequential approach in the enhancement of the catch monitoring/reporting system. While the WG met the TOR regarding fishing effort in the macro-sense, limited information was provided on the spatial and temporal changes or variability of the fishery (fleet activity) which occurred within a broad geographical area. It would have been informative to see how and where fishing effort has spatially decoupled over time in relation to the available surveys and their coverage. The Panel agreed with the decision to use CAMS to estimate catches and discards given its reasonable comparability with past approaches; however, there were a few inconsistencies identified in the reporting with slight differences for the CAMS between the report and the presentation. These differences need to be resolved and corrected in the reporting. The Panel strongly recommends comparative analyses presented from 2020 onward be conducted as far back as possible (e.g. 1996). Doing so will allow for determining if differences between previous methods and CAMS align with smaller differences observed from 2021-2023 or larger differences as observed in 2020 (i.e. up to a 25% difference).

The number of biological samples (50 fish/sample) from commercial catches has dramatically declined and the temporal coverage has contracted since 2019 due to the reduction in allowed catch and management restrictions. Although the total number of samples is the lowest in the time series (1964-present), the WG concluded that the level of sampling was adequate to describe the characteristics of the mobile fleet. Information on the fixed gear fishery (i.e., weir/shutoff) was provided by the Canadian Department of Fisheries and Oceans and resulted in some slight modifications, but nothing of consequence relative to the assessment. The weight at age for the commercial fishery, which showed a steep decline between 1980 and 1990 for most adult ages (i.e. representing a potential loss in productivity), appears to have leveled off and may even be increasing slightly for some age groups in recent years. This is consistent with observations for adjacent and ocean-wide herring stocks as well as many other fish species. Spring spawning fish were mentioned on several occasions during the meeting. While traditionally, the stock has been dominated by fall spawners and remains so, the continued presence of spring spawners in catches and survey sampling may be a forerunner of things to come. Unfortunately, their presence appears to be more anecdotal with no quantification of their presence. The proportion of spring spawners in the resource and the fishery is not expected to be large or have an impact on the assessment at the current time, but worth considering in the future.

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

The Panel concluded that this TOR was adequately addressed.

The WG provided an overview of the fishery-independent surveys available and used in the Atlantic herring stock assessment. The four surveys currently used in the assessment model included the NEFSC Spring (1968-present) and Fall (1963-present) Bottom Trawl (BT) surveys, the Summer (Shrimp) survey (1980s-2023) and the Acoustic survey (1998 – present) conducted as part of the Fall BT survey. The former three surveys focus on the near-bottom distribution and abundance of herring (Figure 3), while the latter quantifies fish throughout the water column. The Panel suggested that additional explorations of the Acoustic/BT survey data be undertaken to evaluate the catchability of the gear and potentially combining the Fall BT and Acoustic survey data into an age-based acoustic index given they are from the same ship and time.

All surveys have broad geographical coverage and incorporate the general extent of herring distribution at the time of the surveys. The NEFSC BT surveys time series have been split multiple times to account for gear and vessel changes. These surveys have been the subject of several evaluations in past assessments and the current configuration has been consistently applied since the 2018 framework review. The WG felt there was no new information to suggest reevaluating the previously accepted characteristics of these surveys and calibration between the vessels or door changes.

One noted omission in the assessment was the larval survey series that extends back into the 1970s (Richardson et al. 2010). This extensive survey, which has a broad coverage, should be explored further as it may provide additional information on the early life history of herring. It also has the potential to be used as an index of abundance to complement the limited data in the assessment surrounding recruitment (Age-1 herring) or herring in their first year of life. The Summer (Shrimp) survey was paused in 2024 and uncertainty about its future was expressed. The Panel suggested that the WG evaluate the impact of permanently losing a survey on the assessment beyond the traditional leave-one-out approach and hope there is an ability to prioritize its continuation.

For many fish stocks, there is limited data to quantify recruitment and juvenile fish distribution/abundance (i.e., lack of an index of abundance). The WG explored five additional long-term surveys not included in the model: The MENH Spring and Fall BTS, the MADMF Spring and Sall surveys, and the seabird diet composition survey. Particular emphasis was placed on the seabird diet as a potential index of abundance for Age-1 fish. Combined, these surveys cover most nearshore (potentially broader) juvenile habitat and contain quantitative information on the abundance and distribution of juvenile fish. Although these surveys were dismissed from inclusion in the current assessment for various reasons, it was noted that most of these surveys followed trends in abundance consistent with this assessment. The Panel highly recommended exploration of these surveys as indices of abundance, and better leveraging the Age-1 composition data from the state surveys and the seabird diet. Comparing the state surveys and the seabird diet might also allow for confirming the accuracy of the diet data tracking inshore recruitment dynamics. The Panel also suggested that a more quantitative model-based approach (e.g. VAST, sdmTMB) might be useful to integrate surveys with spatial footprints smaller than the stock area into a single index for the assessment.

A limiting factor in the abundance indices currently used in the herring assessment is absence or poor representation of small fish (Age-1) even though the sampling gear, in theory, is capable of capturing this size fish, when present. Selectivity for each survey clearly shows how the various ages are treated in the model and the limits imposed on the younger ages. It should be noted that acoustic technology will detect and quantify all age groups of herring if they are present. A knife-edge full selection for ages 3-8 may not be appropriate. Better understanding and quantification of the recruitment dynamics for this species is needed. Enhanced age composition data collection in recent years (i.e., 25 years) have improved the reliability of all indices.

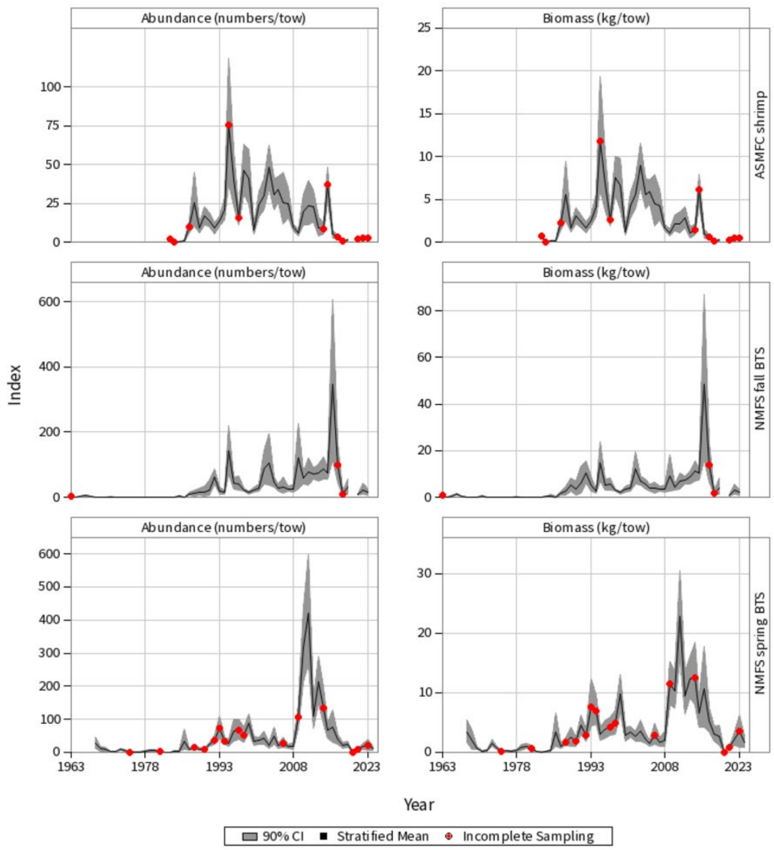

Figure 3. Stratified mean indices from NMFS trawl surveys (number per tow) as provided in the assessment report.

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

The Panel concluded that this TOR was adequately addressed.

An underlying goal of the Atlantic herring RT assessment was to move from ASAP used in previous Atlantic herring assessments to the WHAM platform. Although the WHAM platform was designed to utilize state space model formulations, there are many options where the model can be simplified such that it can use fixed effects similar to ASAP.

The basic data used in the assessment have gone unchanged since the 2018 benchmark stock assessment (NEFSC 2018). It contains two fishing fleets: a mobile fleet and fixed fleet. The mobile fleet selectivity is age-based with selectivity fixed at 1.0 for ages 7 and 8, but estimated for all other ages. The fixed fleet selectivity is age-based and is fixed at 1.0 for Age-2, but estimated for all other ages (see TOR 2 for more details). The surveys included in the assessment model are the NEFSC Spring Bottom Trawl survey, the NEFSC Fall Bottom Trawl survey, the NEFSC Summer survey, and the Acoustic survey (see TOR 3 for more details).

A bridging analysis was constructed to define a WHAM model that mimics the most recent ASAP-based assessment (Figure 4). Run0 is “WHAM as ASAP” (i.e., no random effects, annual recruitments are fixed effects), and although the model did not converge, it replicated ASAP well.

Figure 4. Spawning stock biomass (left) and recruitment (right) using ASAP (ASAP 2022) and WHAM with no random effects used (Run0). Figures are from the assessment report.

After bridging to WHAM, the WG explored several model configurations using eight model evaluation criteria including convergence, residuals, retrospective patterns, AIC, and self-test. Firstly, the WG tried to solve the convergence issue which was achieved by making recruitment an ar1_y random effect process, but even with convergence the model still had some selectivity parameters on bounds and patterned OSA residuals. None of the proposed model configurations resulted in a model without convergence issues or patterned OSA residuals. Next, the WG tried beginning the model in 1987 when age composition observations first become available from the bottom trawl survey (only age composition data from the fishery is available prior to 1987). This truncation resolved most issues of unidentifiable parameters, convergence, and residual patterns. However, this truncation raised concern about the model’s ability to possibly estimate a stock-recruit curve, though the WG further highlights that previous attempts to estimate a stock-recruit curve have failed. Consequently, the WG agreed to begin the model in 1987. The Panel agrees with the conclusion of the WG, but recommends the WG investigate any possibilities to include data prior to 1987 in the future, especially when estimating reference points.

The WG continued with the WHAM model development using the truncated time series. First, the likelihood used to fit the age composition data was reevaluated using dirchlet-multinomial, dirichlet-pool0, logistic-normal-miss0, logistic-normal-ar1-miss0, or logistic-normal-pool0. The WG agreed that the logistic-normal-ar1-miss0 option provided the best residual diagnostics, and used this in all subsequent model configurations. The Panel supported this conclusion. Next, a more thorough evaluation of including various combinations of process uncertainty and correlation structures was conducted, e.g. testing iid, ar1_a, ar1_y, 2dar1 structures of abundance transitions (i.e. NAA), fishery selectivity, natural mortality (M), or recruitment. Over 200 model configurations were tested by the WG, and they used a suite of diagnostics to evaluate the different configurations. They investigated potential causes for model configurations that failed the diagnostics and provided solutions to account for those issues. The final model configurations used (1) the time- and age-invariant M=0.35, (2) decoupled NAA and recruitment processes, (3) iid distribution for the recruitment process, and (4) ar1_a NAA random effects, and) estimated initial numbers-at-age (results compared to previous MT ASAP run in Figure 5). The Panel supported the final selected model configuration. However, the Panel suggested that the WG evaluate the possibility to estimate recruitment at Age-2, as the current recruitment estimates are solely based on computational estimations (Age-1) without any observations. Furthermore, the effects of the NAA random effects need to be explored, especially during the projections. In addition, the Panel recommends investigating more biologically plausible configurations (e.g., implementing an age-based natural mortality). The WG presented one model configuration with scaled age-dependent mortality rates, but from a modeling point of view, the estimates did not change and the less complex model was selected at this time. The Panel suggests that the MT allow for some flexibility for age-dependent natural mortality investigations.

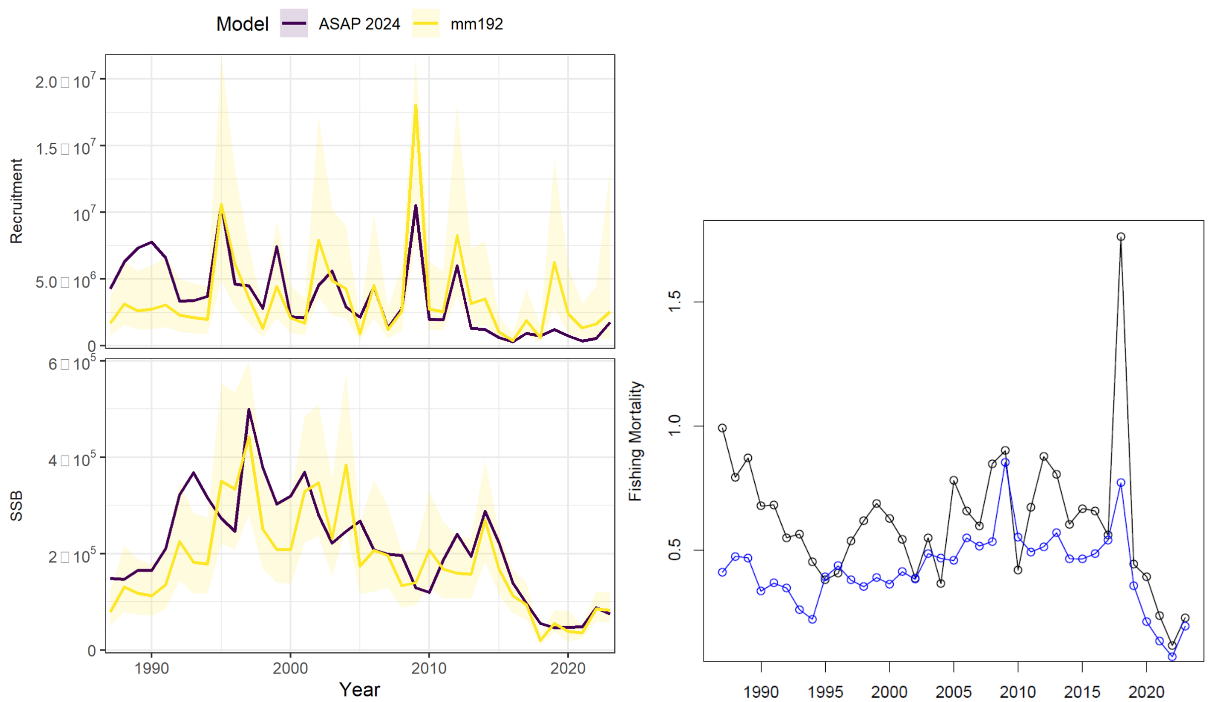

Figure 5. Results compared for the 2024 MT assessment (ASAP) and 2025 RT assessment (WHAM, mm192). Comparisons are provided for spawning stock biomass (SSB) and recruitment (left) and fishing mortality (right). For fishing mortality, the WHAM estimates are provided in black, and the ASAP results are presented in blue.

The Panel wants to emphasize the effort of the WG to implement an Age-1 recruitment index based on seabird diet or the utilization of other environmental indices as covariates in the final WHAM model. In all cases, there was a positive signal in the data, but the model did not improve enough to justify the inclusion of these extra parameters. Some variables used also produced counterintuitive relationships and thus need further exploration. The Panel recommends to continue work on this topic.

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

The Panel concluded that this TOR was adequately addressed.

The principles of previous assessments have been maintained or slightly improved since previous assessments, and the parameter setting and updates of input data were deemed appropriate. As in previous assessments, a stock-recruit relationship was not established because even very low stock sizes are as likely to produce average recruitment as large stock sizes. The WG believed that stock-recruit data from WHAM also did not suggest an identifiable relationship (Figure 6). Therefore, a proxy, F40% (based on SPR calculations) for Fmsy was used by the WG. Assuming continued average recruitment might be problematic at low stock sizes, but given the current management strategy of reducing F when SSB is below half B40% for the major fishing fleet (the mobile-fleet), this results in a low fishing pressure at low stock size.

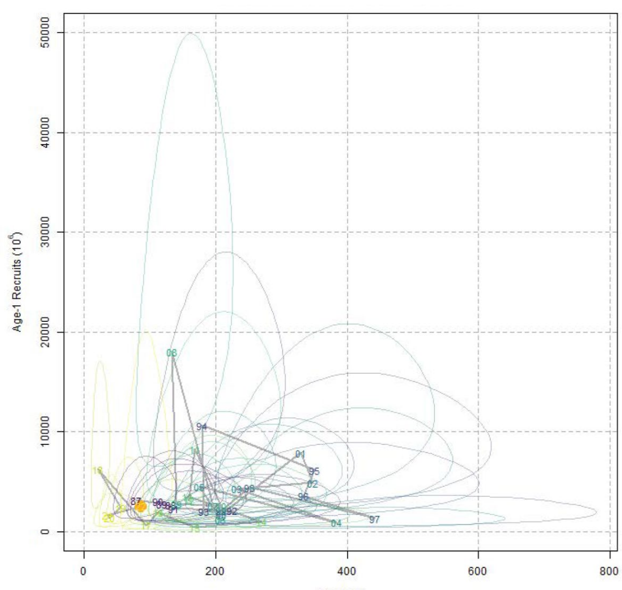

Figure 6. Stock-recruit plot from WHAM, as provided in the TOR 4-6 working paper.

In spite of the difficulties in establishing a suitable stock-recruit relationship in most fish stocks, there has been a trend internationally over the past decades to do so. The International Council for the Exploration of the Sea (ICES) outlines various types of stock-recruit patterns and suggestions on models and software to apply and estimate stock-recruit model parameters (ICES 2003). One advantage of forcing a stock-recruit curve on the often very unclear stock-recruit pattern is forcing recruitment to zero when spawning stock is zero in the assessment. Doing so will also allow for analytical reference points. As such, the Panel encourages a deeper investigating into prospective stock-recruit relationships for Atlantic herring in the future.

F40% is calculated differently in ASAP versus WHAM. In ASAP, the mobile-fleet selectivity is considered to find F40%, while the fixed-fleet is entirely ignored (so F40% will be higher in ASAP because the fixed-fleet ignored). In contrast, WHAM estimates a global F40% that when distributed among the two fleets produces SPR40%, implicitly accounting of both fleets. WHAM estimates a global F40% and then the fleet-specific F40%s are derived based on relative selectivity. B40% is only slightly considers the fixed fleet in ASAP, but it is more accurate in the WHAM approach.

The very large difference in selectivity between the two fleets (Figure 7), makes interpretation of the F40% complicated when F is defined as F on Age-7, as done in the current assessment. Normally, F is supposed to reflect fishing pressure on the dominating age groups in the fishery. For these two fleets, the dominating age groups are very different. To improve the situation, the Panel suggests to investigate new F metrics such as the use of average F for Ages 2-6.

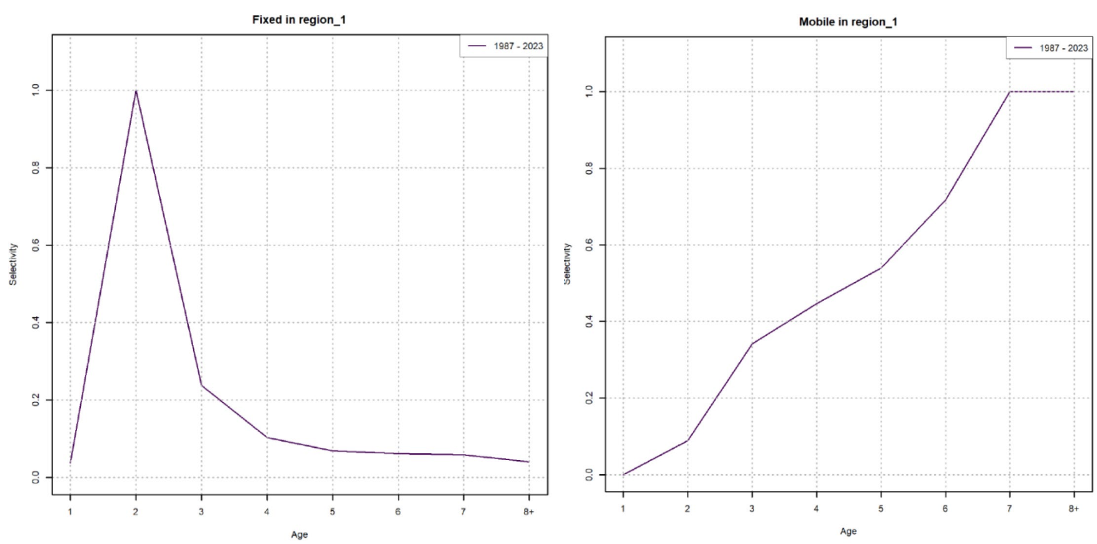

Figure 7. Selectivity patterns by age for the fixed (left) and mobile (right) fleets, as provided in the TOR 4-6 working paper.

Internationally, there has been a growing focus in recent years on density-dependence (DD) in growth, maturity and natural mortality in long-term projects including in reference point calculations. DD in recruitment (to be more precise in survival from egg to recruitment) is well taken care of by using stock-recruit model or as in this assessment, by assuming recruitment to be constant and independent of spawning stock. Density-dependence on recruitment is the main compensation mechanism for fishing in fish stocks generally, but recent science has shown that DD in growth, maturity and natural mortality can also contribute substantially in fish stock dynamics, and missing it can result in an overestimation of stock size in reference points. This might be explored in future assessments of this stock as well.

While using F40% and B40% may be considered antiquated, they have worked reasonably well for many years when managing fish stocks. Newer science involving multispecies interactions, cannibalism, DD, and meta-analysis are available and may be worth considering. For instance there are now proxy Fmsy estimates available based on life history parameters like age at 50% maturity and von Bertalanffy growth parameters K and Linf. Some of these new approaches could be considered. Or at minimum, comparing the results of these to F40% or B40% may be insightful. Alternative reference points (e.g. at 30% or 50%) were not presented, so it difficult to determine how sensitive results are to the reference point specification. It is intuitive regarding the directionality of reference point change impacts when considering alternative percentages, but the magnitude of the impact could not be assessed by the Panel. While the Panel understands and ultimately supports the continued use of 40% based on life history information, further comparisons to other reference points would be useful.

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

The Panel concluded that this TOR was adequately addressed.

The assumptions and input data to the projections were considered adequate by the Panel. However, the projection results seemed highly uncertain (e.g. Figure 8). As such, the Panel was not sure how useful projection estimates beyond the bridge year + 1 are for providing management advice. The Panel was of the opinion that the median value projected showing a quite steep increase, was a balanced result of the data and analysis available.

The Panel believed that the SSB estimates are likely overestimated, given that the projections suggest that SSB can be rebuilt to historic (or time-series high) levels over the next three to four years (Figure 8). Density-dependent effects on growth, maturity and natural mortality would also likely be realized in this case and prevent these levels. Conversely, the lower confidence limit is probably an underestimate as the weight-at-age input data for the projections ignores a slight upward trend in weight-at-age in the most recent couple of years.

The rapid increase in the stock under the projections did not seem to be driven by recruitment assumptions, since these recruitment levels are small compared to SSB in 2025 and 2026 and would not enter SSB before 2027 in significant numbers. However, the Panel did recommend further development in the WHAM code to allow for better sensitivity exercises on the recruitment projections (comparison in recruitment average values versus the recruitment parameter are provided in Figure 9). The random effects were discussed in length. They are ignored in the projections, but even with an ar_y process, the NAA random effects would quickly trend to 0. The negative random effects in recent years in the assessment was a concern of concern of the Panel, but the model did not support a correlation between years.

It was suggested to try to run the projections in a retrospective fashion. The increasing trend of SSB from 2018 to 2023 was theorized to influence projections. Changing to an earlier terminal year and re-running the projections would allow for understanding how optimistic projections are and whether the large uncertainties are linked to the method or the recent years' data.

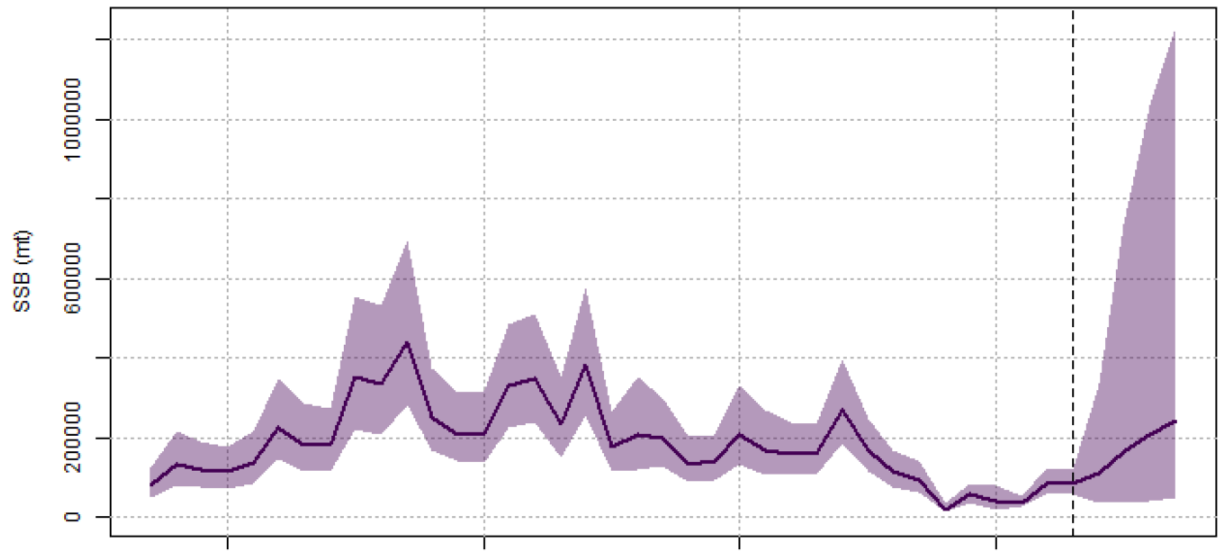

Figure 8. Estimated spawning stock biomass (SSB) from WHAM and projections. The vertical dashed line represents the demarcation between model estimates and projection estimates after 2023. Figure was provided in the TOR 4-6 working paper.

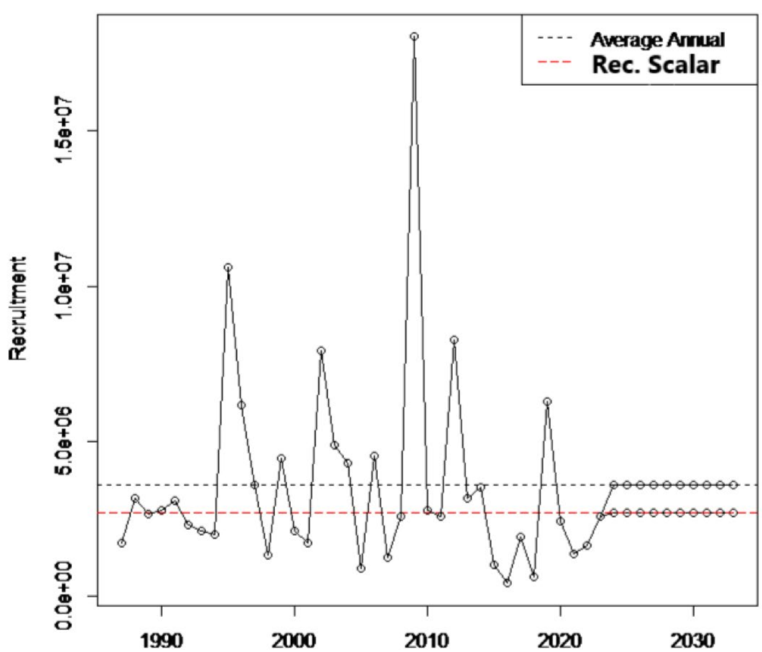

Figure 9. Recruitment time series from WHAM and recruitment projection values between recent time series averages (black) and recruitment estimates from the scalar parameter (red). Figure was provided during the review.

During the review, it was noted that a Projections RT WG would be started soon to evaluate how to improve short term projection methods. The Panel strongly encouraged Atlantic herring being a case study within this RT.

TOR 7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

The Panel concluded that this TOR was adequately addressed.

The WG reviewed proposed recommendations from previous assessments, review panels, and SSC discussions. Some elements previously listed and noted from this exercise were not fully discussed or responded to by the WG (e.g. utility of the study fleet data, further expansion of acoustic surveys and research). The Panel asked for the WG to respond to these verbally during the Review. The Panel noted that the research recommendations posed by the WG were not prioritized, making it challenging for the Panel to respond to recommendations on priorities with WG feedback. As such, during the Review, the Panel asked the WG to note the highest priority research fields from their list in the reports. These recommendations included three themes with sub-elements:

<u>Survey Data</u>

- Continue enhancing survey methodologies, such as incorporating age-length data, to improve reliability and consistency.
- Apply VAST methodologies to combined State and Federal Surveys.
- Evaluate the inclusion of innovative recruitment indices, like seabird diet data, for stock assessments.

<u>Assessment Methods</u>

- Explore alternative age-specific selectivity models and the use of random effects selectivity parameters.
- Investigate the use of larval indices as a potential index of recruitment.
- Refine projection models by integrating ecosystem indicators such as zooplankton abundance and thermal habitat suitability.

<u>Projection Methods</u>

- Integrate environmental datasets, including temperature and prey availability indices, to refine projection assumptions.
- Address overly optimistic projections by exploring methods that better account for recruitment variability and environmental drivers.
- Evaluate how and if NAA random effects should be treated during the projection period.
- Consider refining how weight- and maturity-at-age are specified in the projection period (i.e., perhaps a recent 5-year average is not best, especially with trending traits).

The Panel supported this list, as well as their recommendations put forth in the other TORs. The Panel did iterate multiple research recommendations again to ensure they were pursued in the future. The first was to continue investigating recruitment either as an index of abundance or redefining it in WHAM. Using the inshore surveys or the seabird diet data show promise in constructing a recruitment index for Atlantic herring. Additionally, the WHAM could start at Age-2 as opposed to Age-1 given if the consensus continues that there is no informative data for Age-1 abundance. The Panel reemphasized the need to better understand the impacts of NAA random effects on the WHAM model, particularly their influence on projections. Further development of the environmental covariates for future WHAM testing could be fruitful, as well as additional evaluation of natural mortality for a species that can have high natural mortality rates at younger age classes.

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

The Panel concluded that this TOR was adequately addressed.

The WG felt that if WHAM is rejected (WHAM configurations with random effects or those with fixed effects and thus replicating ASAP), the issues causing this rejection would not be resolvable within an alternative age-based assessment platform. Thus, the WG concluded that an index or survey trend-based backup would be appropriate. The WG chose the existing I-smooth approach (Legault et al. 2022) as an appropriate backup plan method, which has been used for several stocks in the northeast U.S. region. The approach integrated all fisheries-independent survey indices used in the proposed WHAM model into the I-smooth approach. The WG noted that weighting the indices used in I-smooth based on some confidence criteria (e.g. by survey area coverage, likelihood values from previous ASAP or WHAM MT/RT assessments) could be evaluated as needed. However, the Panel noted that this was not a priority. The Panel concluded that the proposed backup approach would be appropriate if the WHAM model was rejected.

References

Brooks, E.N., and Deroba, J.J. 2015. When “data” are not data: the pitfalls of post hoc analyses that use stock assessment model output. Canadian Journal of Fisheries and Aquatic Sciences. 72(4): 634-641.

Dickey-Collas, M., Hintzen N. T., Nash, R.D.M., Scho ̈n, P-J., and Payne, M. R. Quirky patterns in time-series of estimates of recruitment could be artefacts. – ICES Journal of Marine Science, 72: 111–116.

ICES 2003. Report of the Study Group on Precautionary Reference Points For Advice on Fishery Management. ICES CM documents. 2003/ACFM:15.

Legault, C.M. et al. 2022. Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. CJFAS 80:27-42. http://dx.doi.org/10.1139/cjfas-2022-0045.

NEFSC. 2018. 65th Northeast Regional Stock Assessment Workshop (65th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 18-11; 659 p.

Richardson, D. E., Hare, J. A., Overholtz, W. J., and Johnson, D. L. 2010. Development of long-term larval indices for Atlantic herring (Clupea harengus) on the northeast US continental shelf. ICES Journal of Marine Science, 67: 617–627.

Appendix 1 - Terms of Reference for the Atlantic Herring Research Track Stock Assessment

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

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

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

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

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

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

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

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

Appendix 2 - Atlantic Herring Research Track Assessment Peer Review Meeting Draft Agenda

March 10 - 14, 2025

Clark Conference Room, NEFSC Aquarium Building, Woods Hole, MA

Google Meet joining info: https://meet.google.com/her-pcjt-aog

Or dial: (US) +1 316-536-0378 PIN: 521 394 633#

DRAFT AGENDA* (v. 2/12/2025)

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

Monday, March 10, 2025

Time	Topic	Presenter(s)	Notes
9:00 a.m. - 9:15 a.m.	Welcome, Introductions & Agenda Logistics	Conor McManus Brian Hooper	
9:15 a.m. - 9:30 a.m.	Research Track Process Overview	Kristan Blackhart	
9:30 a.m. - 10:00 a.m.	Overview of Herring RT	Matthew Cieri	
10:00 a.m. - 10:30 a.m.	Management History and Current Practices	Jamie Cournane	

10:30 a.m. - 10:45 a.m.	BREAK		
10:45 a.m. - 12:00 p.m.	TOR 1 - Ecosystem and Socioeconomic Profile & Supporting Analyses	Adelle Molina	Ecosystem
12:00 p.m. - 1:00 p.m.	LUNCH		
1:00 p.m. - 2:00 p.m.	TOR 1 - Copepods & VAST	Sarah Gaichas	Ecosystem
2:00 p.m. - 3:00 p.m.	TOR 1 - Haddock Predation	Micah Dean	Ecosystem
3:00 p.m. - 3:15 p.m.	BREAK		
3:15 p.m. - 4:00 p.m.	TOR 1 - Stakeholder Observations	Jamie Cournane	Ecosystem
4:00 p.m. - 4:15 p.m.	Public Comment	Public	
4:15 p.m. - 5:00 p.m.	Discussion/ Summary	Review Panel	Conclusions, Recommendations, & Final Wrap-up for TOR 1
5:00 p.m.	ADJOURN		

Tuesday, March 11, 2025

Time	Topic	Presenter(s)	Notes
9:00 a.m. - 9:05 a.m.	Welcome/Logistics/Agenda	Conor McManus Brian Hooper	
9:05 a.m. - 10:30 a.m.	TOR 2	Ashley Asci Matthew Cieri	Catch
10:30 a.m. - 10:45 a.m.	BREAK		
10:45 a.m. - 12:00 a.m.	TOR 3	Matthew Cieri	Surveys
12:00 p.m. - 1:00 p.m.	LUNCH		
1:00 p.m. - 3:00 p.m.	TOR 4	Jonathan Deroba	Model
3:00 p.m. - 3:15 p.m.	BREAK		
3:15 p.m. - 3:45 p.m.	TOR 4, continued	Jonathan Deroba	Model
3:45 p.m. - 4:00 p.m.	Public Comment	Public	

4:00 p.m. - 5:00 p.m.	Discussion/ Summary	Review Panel	Conclusions, Recommendations, & Final Wrap-up for TORs 2, 3, & 4
5:00 p.m.	ADJOURN		

Wednesday, March 12, 2025

Time	**Topic**	**Presenter(s)**	**Notes**
9:00 a.m. - 9:05 a.m.	Welcome/Logistics/Agenda	Conor McManus Brian Hooper	
9:05 a.m. - 10:00 a.m.	Review Homework	Conor McManus	If needed
10:00 a.m. - 10:30 a.m.	TOR 5	Jonathan Deroba	BRPs
10:30 a.m. - 10:45 a.m.	BREAK		
10:45 a.m. - 12:00 p.m.	TOR 5, continued	Jonathan Deroba	BRPs
12:00 p.m. - 1:00 p.m.	LUNCH		

1:00 p.m. - 3:00 p.m.	TOR 6	Jonathan Deroba	Projections
3:00 p.m. - 3:15 p.m.	BREAK		
3:15 p.m. - 3:45 p.m.	TOR 8	Jonathan Deroba	Backup Approach
3:45 p.m. - 4:00 p.m.	Public Comment	Public	
4:00 p.m. - 5:00 p.m.	Discussion/ Summary	Review Panel	Conclusions, Recommendations, & Final Wrap-up for TORs 4, 5, 6, & 8
5:00 p.m.	ADJOURN		

Thursday, March 13, 2025

Time	**Topic**	**Presenter(s)**	**Notes**
9:00 a.m. - 9:05 a.m.	Welcome/Logistics/Agenda	Conor McManus Brian Hooper	
9:05 a.m. - 10:00 a.m.	Review Homework	Conor McManus	If needed
10:00 a.m. - 10:30 a.m.	TOR 7	Jamie Cournane	Research Recommendations

10:30 a.m. - 10:45 a.m.	BREAK		
10:45 a.m. - 12:00 p.m.	TOR 7, continued	Jamie Cournane	Research Recommendations
12:00 p.m. - 1:00 p.m.	LUNCH		
1:00 p.m. - 3:00 p.m.	Revisit any remaining issues	Conor McManus	If needed
3:00 p.m. - 3:15 p.m.	BREAK		
3:15 p.m. - 3:30 p.m.	Public Comment	Public	
3:30 p.m. - 5:00 p.m.	Discussion/ Summary	Review Panel	Conclusions, Recommendations, & Final Wrap-up for TOR 7, any remaining issues
5:00 p.m.	ADJOURN		

Friday, March 14, 2025

Time	Topic	Presenter(s)	Notes

9:00 a.m. - 5:00 p.m.	Closed panel writing session	Review Panel	*If necessary, open session may be reconvened to revisit any remaining issues
5:00 p.m.	ADJOURN		

Appendix 3 - Atlantic Herring Research Track Assessment Peer Review Attendees (March 10-13)

CIE - Center for Independent Experts
GARFO - Greater Atlantic Regional Fisheries Office
MAFMC - Mid Atlantic Fisheries Management Council
NEFMC - New England Fishery Management Council
NEFSC - Northeast Fisheries Science Center
SSC - Science and Statistical Committee
~~~~~~

Peer Review Panel

First	Last	Affiliation
Conor	McManus	NEFMC SSC (Chair)
Henrik	Sparholt	CIE
Lorian	Berg	CIE
Gary	Melvin	CIE

Working Group

First	Last	Affiliation
Ashley	Asci	GARFO
Matt	Cieri	Maine DMR (Chair)
Jamie	Cournane	NEFMC
Kiersten	Curti	NEFSC
Micah	Dean	MA DMF
Jonathan	Deroba	NEFSC (Assessment Lead)
~~~~~~

Sarah	Gaichas	NEFSC
Gavin	Fay	SMAST
John	Wiedenmann	Rutgers

Attendees

First	**Last**	**Affiliation**
Emilie	Franke	ASMFC
Richard	Klyver	Blue Green Future
Gerry	O Neill	Cape Seafoods Inc
Gareth	Lawson	Conservation Law Foundation
Kurt	Gottschall	CT DEEP
Kevin	Job	CT DEEP
M	Smith	Public
Tim	Barrett	Fisheries and Oceans Canada
Nathan	Hebert	Fisheries and Oceans Canada
Kaleigh	Hill	GARFO
Carrie	Nordeen	GARFO
Ashley	Trudeau	GARFO
Jonathan	Watson	GARFO
Raffaela	Terzo	Gloucester Fishermen's Wives Association
Jeff	Kaelin	Lunds Fisheries
Sefatia	Romeo Theken	MA Department of Fish and Game (FWE)
Melanie	Griffin	MA DMF
Kelly	Whitmore	MA DMF
Jason	Didden	MAFMC
Brandon	Muffley	MAFMC
Megan	Ware	Maine Department of Marine Resources
Addie	Binstock	ME DMR
Andrew	Applegate	NEFMC
Emily	Bodell	NEFMC
Rachel	Feeney	NEFMC
Cate	OKeefe	NEFMC

Charles	Adams	NEFSC
Larry	Alade	NEFSC
Christina	Asante	NEFSC
Brandon	Beltz	NEFSC
Kristan	Blackhart	NEFSC
Jason	Boucher	NEFSC
Russell	Brown	NEFSC
Maxwell	Grezlik	NEFSC
Amanda	Hart	NEFSC
Daniel	Hennen	NEFSC
Cameron	Hodgdon	NEFSC
Brian	Hooper	NEFSC
Christopher	Legault	NEFSC
Chengxue	li	NEFSC
Emily	Liljestrand	NEFSC
Brian	Linton	NEFSC
Tim	Miller	NEFSC
Adelle	Molina	NEFSC
Paul	Nitschke	NEFSC
Stephanie	Owen	NEFSC
Charles	Peretti	NEFSC
Burton	Shank	NEFSC
Laura	Solinger	NEFSC
Katherine	Sosebee	NEFSC
Michele	Traver	NEFSC
Samuel	Truesdell	NEFSC
Abigail	Tyrell	NEFSC
Sara	Weeks	NEFSC
Anthony	Wood	NEFSC
Renee	Zobel	NH Fish & Game
Cheri	Patterson	NH Fish & Game

Joseph	Warren	NOAA Affiliate
Mary Beth	Tooely	O'Hara Corporation
Justin	Didden	Public
Bill	Lucey	Save the Sound
Alexander	Jensen	SEFSC
Maria Cristina	Perez	SMAST
Jaclyn	Higgins	Theodore Roosevelt Conservation Partnership
Andrew	Jacobs	WTGHA

Appendix 4 - Materials provided or referenced during the Atlantic Herring Research Track Stock Assessment Peer Review meeting

Working papers and presentations are available on a NEFSC website (https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php) by selecting the species and year of assessment.

Working Papers and Background Documentation:

Information Type	File Name
All Files	2025_HER_UNIT_1.zip
All Files	2025_HER_UNIT_2.zip
Assessment Report	2025_Herring_RT_assessment report_final_508.pdf
Figures	Figures_Homework and other_Deroba_revised.pdf
Background	Acoustic-WP.pdf
Background	Boosted Regression Tree-WP.pdf
Background	CAMS-WP.pdf
Background	Copepod Indices-WP.pdf
Background	ESP-WP.pdf
Background	Fecundity-WP.pdf
Background	Haddock_predation-WP.pdf
Background	Management-WP.pdf
Background	RecCovariates-WP.pdf
Background	Seabird_Diet-derived_Recruitment_Index-WP.pdf
Background	Spawn_timing-WP.pdf
Background	Stakeholder-WP.pdf
Background	TOR_2-WP.pdf
Background	TOR_3-WP.pdf
Background	TOR_4-5-6-WP.pdf
Background	TOR_7- WP.pdf
Background	TOR_8-WP.pdf
Background	Threshold-WP.pdf
Background	ToR_1-WP.pdf

Background	_Working_paper_list.pdf
Presentations	Presentation_Management History and Current Practices_Cournane.pdf
Presentations	Presentation_Overview_TORs_2-3_Cieri.pdf
Presentations	Presentation_TOR 1_Copepods and VAST_Gaichas.pdf
Presentations	Presentation_TOR 1_Ecosystem and Socioeconomic Profile_Molina.pdf
Presentations	Presentation_TOR 1_Stakeholder Engagement_Cournane.pdf
Presentations	Presentation_TOR 1_haddock predation_Dean.pdf
Presentations	Presentation_TOR 2_Asci.pdf
Presentations	Presentation_TOR 4-6_Deroba.pdf
Presentations	Presentation_TOR 4_Zooplankton indices as covariates in WHAM_Gaichas.pdf
Presentations	Presentation_TOR 7_Cournane.pdf
Presentations	Presentation_TOR 8_Deroba.pdf