Independent Peer Review Report on the Georges Bank and Eastern Georges Bank Haddock Research Track Stock Assessment

presented to the
Center of Independent Experts (CIE)

by

Joseph Powers
NOAA retired
Louisiana State University retired

Executive Summary

The Research Track Assessment of Georges Bank (GB) and Eastern Georges Bank (EGB) haddock provided extensive analyses and research results. Of particular importance was the development of the WHAM state-space framework for including various options for random effects and correlation on post-recruitment survival, M, selectivity, survival during age 0 (recruitment). This framework was utilized for the assessments and has been a major step forward. This approach has allowed model structures to be defined that reduce retrospective patterns and incorporate the density-related effects of ultra large year classes that have plagued haddock assessments.

The selected GB model uses autocorrelated age and year variation around survival and selectivity with a fixed M=0.2 (other specifications noted in TOR 4). The EGB WHAM framework fixes M in the early history and then estimates it for recent years (TOR 4). Both models have achieved the desired result of reducing their retrospective patterns to acceptable levels. Research that revisited catch estimations, survey indices and their changes over the years, age-length and growth updates, exploration of possible mechanisms of recruitment processes and spatial dynamics were explored and incorporated into the base models. Thus, the WG has created acceptable model structures to be carried forward for the Management Track Assessment.

The WHAM framework has been extremely useful in defining the current assessment models. But it also has the potential for improvements in stock assessments in general as well as haddock in particular by expanding the WHAM relationships (such as density-dependent M's); by defining rate models and controlled simulations where deviations from underlying relationships are interpretable as ecological effects; and by linkage between single-species assessments to provide trends in "ecosystem effects" that can provide strategic advice for management (climate change). Presumably, this would be an activity of the WHAM WG.

The WG did not fully agree on the best stock identification approach between the two options, entire GB versus EGB alone. In my opinion, the best approach would be to use GB as a base assessment and attempt to link GB and EGB components in the model. The goal is to stabilize the estimation structure such that they will not deviate too much from the common data and information. Then, separate management objectives and constraints of the US and Canadian systems should be explored through management strategy evaluations and management procedures.

Background

Georges Bank haddock (*Melanogrammus aeglefinus*) has previously been assessed by examining both Georges Bank as a whole (GB) and the eastern portion of the distribution Eastern Georges Bank (EGB). The most recent assessment occurred in 2019 for both GB and EGB where VPA's were employed. For GB this was an operational update that provided the determination that the GB stock was not overfished and not undergoing overfishing and formed the basis of the management advice with stock projections through 2022. Nevertheless, a more extensive benchmark assessment of GB haddock had not been done since 2008. Additionally, the 2019 EGB haddock VPA was not accepted as a basis for management advice due to a large retrospective pattern. This situation provided the motivation for a Research Track Assessment through the establishment of a working group (WG) to address assessment research approaches. WG activities extended from September 2020 through February 2022. Their findings, analyses, conclusions and recommendations were documented in a report and presented at a meeting of the Georges Bank and eastern Georges Bank Haddock Research Track Stock Assessment Peer Review Panel that met via WebEx from March 28 through March 31, 2022. The Panel was composed of three scientists selected by the Center for Independent Experts (CIE): Anders Nielsen (Technical University of Denmark), Kevin Stokes (Stokes.Net.NZ Ltd), and Joseph Powers (NOAA retired). Additionally, the Panel consisted of co-chairs Rob Kronlund (Interface Fisheries Consulting, Ltd.) and Richard Merrick (member of the New England Fisheries Management Council Scientific and Statistical Committee). The review focused on 12 Terms of Reference (TORs, listed below). The Panel addressed each of those TORs in turn, based on the WG's documentation and the feedback and opinions of the contributing scientists.

Individual Reviewers' Roles in the Review Activities

The role of each of the Center of Independent Experts (CIE) in this review was to attend and participate in the panel review meeting, conduct an independent peer review in accordance with the requirements specified in this Performance Work Statement (Appendix 2) and TORs, to assist the Peer Review Panel (co)Chairs with contributions to the Peer Reviewer Summary Report and to deliver individual Independent Reviewer Reports to the CIE accordingly explaining whether each research track Term of Reference was or was not completed successfully.

In particular my responsibility as a CIE reviewer is to deliver an independent report addressing the TORs in the work statement. This document represents my report. My specific independent responses to each of the 12 TORs and my overall conclusions follow.

Summary of Findings for Each Term of Reference

The current stage of Georges Bank and Eastern Georges Bank Haddock assessment and the basis of this review is a *Research Track Stock Assessment*. Therefore, considerable effort over many months was extended by the Working Group to address a large number of research issues that had been extant since the last benchmark assessment in 2008 in order to improve the upcoming management track assessment (*circa* July-Sept 2022). To achieve this, the WG was directed to address 12 Terms of Reference. These TORs were typically opened ended with wording like "review and present…" and "review and evaluate…." The open-endedness is as it should be for the research investigation activities of a WG, allowing the WG to explore research options. However, as a Reviewer and non-participant in

the WG, I am left with the conclusion for most TORs that: "yes, the WG did, indeed, review and present…." If I were a WG participant, I might have attempted to steer the group to investigate several other research avenues without knowing *a priori* how useful that would have been. Nevertheless, I am comfortable with the WG's conclusions and the assessment approach they have taken at this stage. With my comments I have attempted to provide guidance about future research: denoting those aspects that are longer term to be included in future research tracks and those that should be addressed in the final implementation of this assessment *circa* July-Sept 2022.

TOR 1. Review existing research efforts, data, and habitat information in the Gulf of Maine and Georges Bank, identify any findings relevant to influences of ecosystem conditions on haddock, and consider those findings, as appropriate, in addressing other TORs. For processes that the working group deems important and promising that are not currently feasible to consider quantitatively, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments.

This TOR was met for both GB and EGB haddock.

Generally, from a stock assessment standpoint, "relevant influences of ecosystem conditions on haddock" relate to the influence of environmental factors on growth, mortality, recruitment and movement/spatial distribution. Variability in environmental variables contribute significantly to the realized rate parameters through annual noise and through trends in the underlying relationships. Typically, in stock assessments some parameters are fixed, others are estimated and then the process and measurement error are combined in MLE. The common exception to this is the process error on the stock-recruitment relationship (sigma-R) which is estimated or assigned in many stock assessments. However, with the development of WHAM the random effects on other biological/ecological variables, such as numbers at age (NAA, i.e., survival), M, selectivity can now be included. I see this as an important step in refining how "ecosystem" effects are incorporated into stock assessments.

As I understand it, there is a research track WG addressing WHAM. I can foresee several activities of such a group: 1) expanding the WHAM relationships (age-specific M's and density-dependent M's (e.g., Powers 2014. ICES Journal of Marine Science, doi.10.1093/icesjms/fst22) and others); 2) defining rate models and controlled simulations where deviations from underlying relationships are interpretable as ecological effects; 3) linkage between single-species assessments to provide trends in "ecosystem effects" that can provide strategic advice for management (climate change). At this stage the exploration by WHAM is just beginning.

The discussion at the review meeting under this TOR focused on habitat models including machine learning techniques to assign habitat scores based on zooplankton blooms using survey data. This research indicates that the habitat where haddock occur has expanded over the years. It is less clear what the underlying factors were that caused this. Was it simply that the three large year-classes resulted in recruits distributing to larger areas? If so, were these areas less suitable (affecting growth and M)? Habitat scores indicate the size of haddock habitat has changed over time, declining from the late 1970s into the 1990s and then increasing over the last two decades and especially in the most recent years of the ultra-large 2013 year class.

The role of climate change and fish density in affecting changes in haddock distribution was explored with GAMs predicting Area of Occupancy of haddock in the GB to have varied substantially over time, with a consistent trend between spring and fall and a clear northward shift in fall and to some extent eastward in the last 10 years.

While these analyses were useful in short term assessment decisions (e.g., movements northwards and to some extent eastward into Strata 29 and 30, perhaps in association with warming bottom temperatures, led to the inclusion of data observed from these strata into the EGB assessment model), the greater benefit is as an "hypothesis-generator". Results can be used to formulate (model) underlying ecological relationships that can be explored by WHAM.

A further caution: a theme of this research track has been issues of changes in growth, possible changes in M and recent high variability in recruitment. Since all three are occurring at the same time, it will be hard to decipher the effects of each. For example, length-specific M's (Lorenzen) can be equivalent to density-dependent M's for common values of von Bertalanffy K and asymptotic M. Thus, this emphasizes the importance of WHAM formulation and testing to arrive at pragmatic simplifications that are estimable and provide usable advice.

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.

This TOR was met for both GB and EGB haddock.

Since the last benchmark was in 2008, it was important that catch estimation procedures be revisited to see if they may have evolved over time. This is especially important in that a major assessment concern has been the occurrence of retrospective patterns, particularly in the EGB. Mis-reported/estimated catches can be an obvious source of that.

Estimation of discards has been done by both the US and Canada through the implementation of observer programs. Currently, Canada observes 50 – 100% of trips and about 37% of the landed catch and the discards thus estimated are small relative to that catch. US discards are relatively small but observer sampling has not been as extensive as it was historically. For the US the size/age distribution of discards are estimated from survey seasonal W-L relationships, as size data are not collected by observers.

The current understanding of the WG on catch uncertainties are reiterated as: 1) minor uncertainty in statistical area landed from the allocation and stock border procedures; 2) unknown uncertainty of over- or underreporting of haddock landings; a recent court case documented a large seafood mislabeling other groundfish as haddock due to having excessive haddock quotas compared to restrictive quotas on the mislabeled fish; 3) minor uncertainty of gutted:whole weight conversion; 4) unknown uncertainty of probabilistic assignment of age from age-length keys for year classes that are adjacent to the exceptional 2003, 2010, and 2013 year classes (age smearing); 5) unknown uncertainty of estimating catch weights from survey length-weight relationship.

Of these, proposals have been aired to more systematically address whole to gilled and gutted relationships, as the original conversion was determined decades ago and product-handling may have evolved. Additionally, I would encourage more within-year and sex-specific W-L and L-fecundity sampling as a means to address fecundity and egg production as well as better catch-weight estimation.

TOR 3. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). Characterize the uncertainty in these sources of data.

This TOR was met for both GB and EGB haddock.

There is a long history of resource surveys in Georges Bank by both the US (fall and spring bottom trawl surveys since the late 1960's) and Canada (since 1987). These surveys have formed the basis for the assessment of many stocks of which haddock is one. For the years that the 3 surveys overlap (1987-present), there is strong agreement in trend, showing a steady increase from the 1990s in both biomass and abundance, peaking in the mid-2010s, and some decline to 2019. However, the Canadian series does not survey the GB, only sampling in the higher density strata on eastern Georges Bank.

An important event in the survey history was the replacement of the survey vessel *RV Albatross* with a new vessel *RV Bigelow* in 2009. Calibration studies for the two vessels were conducted and several subsequent analyses were conducted for multiple species. The calibration factor so derived was for adjusting numbers per tow in the survey. For biomass per tow for haddock there had to be an adjustment for the variation in size frequencies induced by the very large year-classes. The model implemented uses the seasonal weight length parameters. This approach was acceptable and provided a useful calibration such that the US fall and spring surveys could be considered for the complete time series.

Despite having a calibration, it was proposed that the series could be parsed into pre-2011 (Albatross) and post-2011 (Bigelow) with separate catchabilities (q's) to be estimated. However, the post- series only has about 8 years and given the concerns about retrospective patterns this could result in instabilities. Also, the trends are similar, as are the cohort progressions in the surveys. Thus, the idea was rejected at this stage. I agree with this decision.

As noted, the Canadian survey started in 1987 and focuses on EGB. Additionally, there are individual years where a subsequent vessel was used with no conversion factor. Despite this the trends in the surveys are quite similar. So, in summary, the surveys are well documented and will be an integral part of the assessment.

Remaining uncertainties include: variance weightings of the surveys do not encompass the calibration factor variances; there are instances of a single very large tow influencing the annual estimate substantially; there is likely additional uncertainty in the calibration factors for small fish samples and no 2020 surveys were conducted (COVID).

The surveys also provided basic data which indicated variability in maturity, weight/length at age and possibly mortality all of which were examined through the assessment modeling.

TOR 4. 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 model and evaluate the strength and direction of any retrospective pattern(s) in both the current and the previously accepted model. Enumerate possible sources of the retrospective patterns and characterize plausibility, if possible.

This TOR was met for both GB and EGB haddock.

As noted by this TOR, an over-arching theme of this assessment is the source and solution to retrospective patterns and the density-dependent implications of huge differences in year class strength on growth, mortality and selectivity. These considerations argue for more sophisticated modeling that includes random effects (process error) on parameters. But prior benchmark modeling started (2008) with VPAs and then moved to ASAP. Additionally, several data changes (calibration, elimination of some survey strata, etc.) were incorporated. Thus, transitions needed to be built from VPA to ASAP and ASAP to WHAM, i.e., demonstrating that ASAP could be structured as the VPA and then data changes evaluated as to whether they were a result of the model structure or of the data changes. Similarly, the transition of ASAP to WHAM was evaluated by structuring WHAM as ASAP and reevaluating data changes. Then a number of WHAM options were explored to come to a final model structure for implementation in the final assessment to be conducted later this year.

In the case of GB, this process and the documentation of the results were very extensive and extremely organized, far exceeding most stock assessments. The analyst is to be commended for the thorough model exploration and the easily followed arguments on why model structures were accepted or rejected.

The final WHAM model chosen was using random effects autocorrelation with year age (2DAR1) for numbers at age (NAA) and fleet selectivity. Variance and correlation were estimated separately for recruitment versus ages 1+. Other model constructions were: logistic fleet selectivity with random effects in second block (1963-2019, the years with survey data); indices with age-specific selectivity; and M fixed at 0.2 for all ages, all years. The differences in the model outcomes compared to previous models are in the following figure (figure 1).

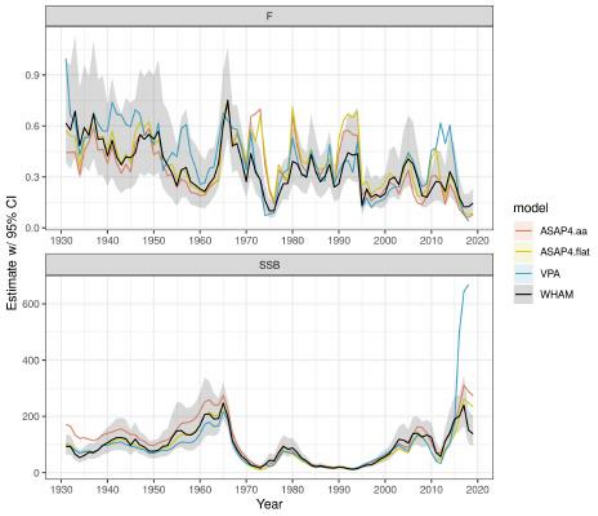

Figure 1: Differences in model outcomes

A retrospective pattern still exists but was considerably less than the VPA and ASAP models (biomass overestimated, F's underestimated). The analysis indicated that Mohn's rho for the final WHAM was within confidence intervals and that adjustment was not recommended. It was noted at the review meeting that one exercise using a conditional simulation test showed that the model appeared to have a substantial bias (of about 30-40%) on both SSB and *F* estimated time series compared to the true. This is unlikely to be a valid result or else the model comparisons on other analyses would have shown similar effects. Thus, there is perhaps some coding/translation error in this graph which should be resolved before the management track assessment.

Development of a EGB model began with exploration of a bridge from the earlier VPA model (which had failed due to a strong retrospective bias) to a WHAM framework model. The selected base model was similar to the GB model using similar explorations. The major difference was that M was fixed at 0.2 for years preceding 2010 but was estimated as an age-invariant fixed effect parameter for the last 10 years (2010-2019). This choice was made because (a) analyses by the WG suggested a hypothesis that M has increased significantly in the past decade in the EGB unit, and (b) model performance relative to survey index trends improved substantially relative to alternative formulations.

Implementation of this WHAM-based EGB assessment model produced results with good model consistency (e.g., small retrospective patterns), and generally good diagnostics (generally good behavior of model residuals and good simulation self-testing performance). Model sensitivities were explored and generally supported the choice of model. There were issues of asynchrony in the model predictions of peak years for the survey indices (NFMS fall and DFO spring), low model estimates of selectivity for younger age classes, and residual errors in survey age composition that should be investigated. However, these issues do not preclude consideration of the model for management application.

An overall impression of both GB and EGB models under this TOR is that a balance is trying to be achieved between the dampening of retrospective patterns and the inclusion of density-dependent parameters presumably caused by the extraordinary year classes in recent years. In the GB case, the chosen modeling structure focused on random effects on survival, recognizing that this was a pragmatic statistical approach and that survival estimates can be affected by any number of factors (misreported catches, M, growth K etc). Thus, all those possible effects are collapsed into a survival estimate. This

achieved the desired result of reducing the retrospective pattern. At this stage of WHAM modeling, I support this approach until such time as factors can be isolated using WHAM (through the WHAM WG).

Conversely, the EGB has assigned changes to M. While I do not disagree with the outcomes of this model structure. I think more exploration of how M operates in the haddock ecosystem is needed (again WHAM WG).

An underlying theme of several TORs is density-dependence of parameters in response to the extraordinary year classes. But the experiments and analyses focus on time periods when density was higher and growth rates were lower and M was (probably) higher. So, the focus was on blocks of time where rates were constant. This approach suffices for estimating SSB and F in the past, but it raises the question of when does density-dependence change back. This arises in projections and BRPs (TORs 5, 6). The problem is that a density-dependent model of the parameters has not been proposed in this assessment. One such model is the Beverton-Holt model using post-recruitment parameters for a cohort (dN/dt=-(Minf+bN)N). In this case Minf might be fixed at 0.2. This can be similar to a Lorenzen M relationship (already implemented in WHAM) but rescales for cohort size (Powers 2014. ICES Journal of Marine Science, doi.10.1093/icesjms/fst22; also note that the catch equation under this M relationship is not the same as the Baranov equation). At any rate, I am comfortable with the current modeling approaches going forward for the management track assessment. But the activity of a WHAM WG will be extremely important.

Essentially, the GB and EGB models are two separate hypotheses about stock structure. Given TOR 12, I would recommend that the GB model be the focus of further evaluation in the future, but that effort is needed to link the two and to harmonize the management advice accordingly. If separate US and Canada management regimes are desired, these can be accommodated with an overall GB model, linking key spatial parameters and achieving equivalent FSPR criteria as desired. However, the modeling to do this will require additional effort that I expect will be unavailable before the ~July-Sept management track assessment.

TOR 5. Update or redefine status determination criteria (Status Determination Criteria, point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and Maximum Sustainable Yield [MSY]) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for Biological Reference Points (BRPs).

This TOR has been met for GB haddock. Note that this TOR was not a requirement for EGB haddock.

The proposal is to continue to use $F_{40\%SPR}$ as the proxy for F_{MSY}. For the current status and short term projections, Bmsy=B40%SPR is to be based on the full time series for recruitment, with 5-year averages for selectivity, weight-at-age, and SSB. While a B40%SPR BRP is conceptually acceptable, the actual value can change considerably with shifts in M, growth and selectivity especially with large year classes. Thus, the actual values of BRPs should be revisited every 2nd year as part of Level 2 management track assessments.

The assessment model essentially assumes a steepness of 1, or more correctly that the time series of SSB has never declined to a point that expected recruitment is significantly less than R0. It would be interesting to examine the SSB-R-FSPR history to see how it lines up with F40%SPR. Perhaps, some

support might be obtained for selecting this BRP. Also, because F40%SPR =FMSY aligns with a specific steepness, it would be interesting to see steepness estimated during the management track.

TOR 6. Define the methodology for performing short-term projections of catch and biomass under alternative harvest scenarios, including the assumptions of fishery selectivity, weights at age, maturity, and recruitment.

This TOR has been met for both GB and EGB haddock, though this TOR was generally most relevant to the GB assessment process.

WHAM is proposed for the projections of the numbers-at-age consistently with the assumed (stochastic) process model. Two-year averaging was deemed appropriate for numbers-at-age and weight-at-age, with 5-year averaging for selectivity. Again, because of the poor performance of projections beyond 3 years, the expectation is that these averages will be updated every 2nd year as part of management track assessments. Note, however, that those projections are typically made for four years including the bridge year. I supported the approach for GB projections and supported the use of WHAM to perform the projections.

For the EGB projections, the analyst converted the EGB base model implemented in WHAM to an operating model to develop closed-loop simulations. The primary purpose of this step was to allow investigation of reference points (TOR 11 for EGB only), projection of various management options (TOR 6), and evaluation of "Plan B" options (TOR 8). For TOR 6 two operating models were configured for investigating short-term projections. Both operating models assume the estimation of M from 2010 to 2019 ("step up" of the M_{est} base model). However, for simulated future data one of the models reverts to the historical M=0.2 ("low M") while the other maintains the recent estimated M for all projection years ("high M", MLE of 0.473).

Ultimately, projections and risk evaluation based on the "low M" scenario were conditionally identified for consideration, while acknowledging the possibility of a future scenario with higher M. This conclusion was reached given the higher M produced estimates of projected SSB lower than any historical SSB, i.e., stock levels outside of historical precedent. But if decreasing M is a function of density dependence, it is not clear how long it would take for the density dependence effect to dissipate. The analyst suggested that the final selection of an M projection scenario for the upcoming TRAC assessment could be based on inspection of two years of additional biological and survey data, a sort of "Plan C". I can support this approach in principle.

It is interesting that projections for both GB and EGB are suggesting reversion to "base" M estimates. I think this is a natural consequence when density-dependence is not modeled explicitly (see TOR 4).

TOR 7. Review, evaluate and report on the status of the Stock Assessment Review Committee (SARC) and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

This TOR was met for both GB and EGB haddock.

The WG through a presentation at the meeting and through their assessment report document provided an extensive list of previous research recommendations emanating from the SARC and WGs. Largely, those recommendations were the driving force for the TORs for this research track assessment. I have made comments on the research related to each TOR.

From an assessment standpoint I would emphasize two aspects for implementation: 1) the exploration of WHAM model structures and estimation techniques for evaluating density-dependence, climate shifts, etc; and 2) the linkage of GB-EGB assessments through, perhaps, MSEs and MPs.

TOR 8. Develop a "Plan B" for use if the accepted assessment model fails in the future.

This TOR was met for the GB assessment which includes the EGB component.

The GB "Plan B" suggestion was to use an index-based method in which the annual dynamics of the NMFS spring and fall surveys were Loess-smoothed and then the annual change of that smoothing was developed into a multiplier which adjusts a new quota based on the multiplier of previous quotas.

For EGB haddock, at the Review meeting, the analyst proposed a single biomass index developed by averaging the NMFS spring, NMFS fall and DFO spring biomass surveys from the EGB area. This index could be used to inform three index-based management procedures: Plan B smooth, a constant index ratio approach, and an alternative set of age-structured indices. Ultimately, the constant index ratio approach appeared to provide outcomes somewhat consistent with SSB of the $F_{40\%SPR}$ level based on closed-loop simulation evaluation. However, it is unclear how much the information base required of these methods depends on key data components that were a part of the "failed" assessment in the first place.

Apparently, a "Plan B" option is now required when going forward to a full assessment. But frankly, I believe this TOR is poorly worded and extremely vague in its intent. Given the amount of data that is available for GB and EGB haddock and the analyses that were presented thus far, I cannot foresee a circumstance in which analysts would forego an application of some form of the existing models and data and argue that the "Plan B" above is somehow better. While an index method might be an adequate method to define a harvest control rule (after MSE testing) for periods between research track or benchmark assessments, it would be hard to accept such a method for scaling a TAC decision given the haddock sources of data that are available.

In the future for research track assessments, some guidance is needed to define this TOR. What does "…accepted assessment model fails..." mean? Accepted by whom? Failed at what? Models don't fail, analyses fail to provide estimates of "X" with required precision and accuracy of "Y" and "Z". What are the X, Y and Z? So, I agree that this TOR was met, but I don't see it as a useful stand-alone exercise.

TOR 9. Review and present any research related to recruitment processes (e.g., spawning and larval transport, and retention), and potential hypotheses for large recruitment events.

This TOR was met for both GB and EGB haddock.

The WG addressed this TOR through several studies relating survival ratios (Age1 recruitment/SSB) to temperature and the magnitude and timing of chlorophyll a blooms, primarily from the work of Friedland (see figure 2 below).

Friedland 2021 dx.doi.org/10.1139/cjfas-2020-0453

Fig. 1. Scatter between survivor ratio and bloom magnitude for estimates using the base set of STARS parameters (*a*) and with mean of the ensemble of estimates using the ranges of parameters (*b*). The error bars in panel (*b*) are 95% confidence intervals around the mean bloom magnitude.

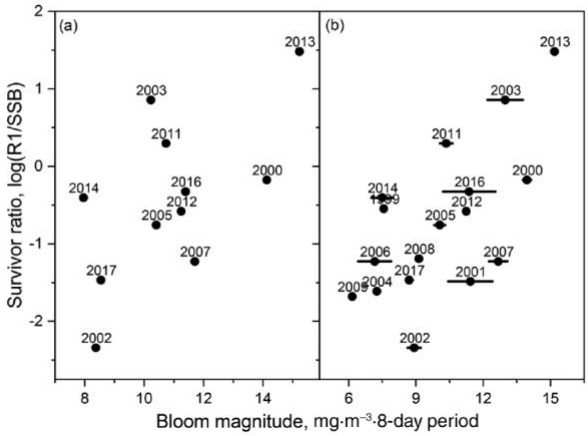

Figure 2: Scatter between survivor ratio and bloom magnitude

The argument that is being made is that blooms increase the opportunities for adult food and thus better recruitment. After discussion at the meeting, it was clear that the research was effectively addressing "quality" of the spawners, i.e. either the weight of an individual spawner increased and/or the number of eggs produced per individual increased (fecundity). Recall that stock-recruitment models are essentially mortality functions of the survival of eggs to recruitment (Brooks, E. N., and Powers, J. E. 2007. ICES Journal of Marine Science, 64: 413–424). Also, recall that the usual assumption is that the number of eggs produced is proportional to SSB (as was done in this assessment). This assumption implies that an individual female's egg production scales to approximately L^3. But these results would suggest how one measures SSB might make a difference and that fecundity/egg viability measures might be important. Food for thought.

The R1/SSB survival ratio measures two different factors: 1) the production of eggs, and 2) the survival of eggs to recruitment. This research is addressing the former. But it is unclear how much the three exceptional year-classes relied on one or both of these. This is another reminder that understanding recruitment processes will remain elusive. Also, note that the sigma-R recruitment deviations in the GB assessment was estimated to be 1.6. This compares to many groundfish-type assessments where the sigma-R's that are used are 0.4-0.6. Clearly the 1.6 was driven by the three exceptional recent year-classes, but it is an open question whether 1.6 is a characteristic of the haddock population-ecosystem long-term or if that system has changed in recent times.

TOR 10. Review and present any research related to density-dependent growth.

This TOR was met for both GB and EGB haddock.

Although explicit modeling of density dependent growth was not implemented into the assessment, indirectly it was accounted for by the size-age data throughout the WHAM assessment. Multiple

analyses and data sets indicate that growth rates are affected by very large year classes. With haddock this is especially important because much of the biological parameter dynamics (M, K) are likely driven by the ultra-strong year classes.

So, I believe that the research has been extensive in documenting that density changes are associated with growth changes and that these are important in the use of size data for assessment, determination of BRPs, maturity and selectivity. However, as noted in TOR 4, density dependent rates should be explicitly explored in WHAM to determine the best way to project future rates such as M and K.

TOR 11. For eastern Georges Bank, provide advice to TMGC on appropriate reference points.

This TOR was very vague as to the expected outcome. It's hard to say what is appropriate without knowing what TMGCs objectives are. After some back and forth at the review I surmised that an appropriate BRP is the same as for the GB, i.e. F40%SPR. The original *F*ref = 0.26 which was derived in 2002 was defined as the F40%SPR at that time but was well below the values of $F_{40\%SPR}$ investigated in the retrospective analysis. Therefore, it is proposed to adopt a $F_{40\%\,SPR}$ (current mean value of 0.488) updated every four years and calculated as the mean over the last 10 years.

The closed-loop simulation approach should lead to management procedures (including control rules). Each management procedure implies trade-offs in performance that could be considered by the TMGC with respect to objectives related to reference points and status determination (whereas the GB analysis under TOR6 would provide BRPs and SDCs). I believe that the MSE/MP approach would be especially relevant for TMGC management within a larger GB assessment context.

TOR 12. Review data related to stock structure of haddock on Georges Bank (including eastern Georges Bank management area) and implications for assessments conducted on the whole bank and on subareas of the bank.

This TOR has been met for GB and EGB.

The classic assumptions of stock identification for a stock assessment are reproductive isolation, demographic independence, homogeneous vital rates, no immigration/emigration. Neither a GB nor an EGB management unit met all stock assumptions perfectly; each had pros/cons with respect to meeting the assumptions. Clearly, the spatial distribution is dynamic and is likely driven by strong year-classes, but I agree with the WG conclusion that "the current distribution and connectivity of haddock across the Bank suggest that haddock on Georges Bank (eastern Georges Bank and western Georges Bank) is a single stock." But future climate drivers may alter this view.

Collectively the Review Panel suggested that stock structure and assessments should focus on a biological stock, and stock management should be dealt with somewhat separately but within the context of the biological stock. I (as a Panel member) agree with that statement, but I also have a pragmatic view of what a biological stock is. Thus, I support the overall GB assessment. However, there is a need to move toward linked GB-EGB components and common F-based reference points to be evaluated through MSEs/MPs. The goal is to use the overall GB assessment to define FMSY related catches and then MPs to establish acceptable TAC allocations for the TMGC users.

Conclusions

The Research Track Assessment of GB and EGB haddock provided extensive analyses and research results. Of particular importance was the development of the WHAM state-space framework for including various options for random effects and correlation on post-recruitment survival, M, selectivity, and survival during age 0 (recruitment). This framework was utilized for the assessments and has been a major step forward. This approach has allowed model structures to be defined that reduce retrospective patterns and incorporate the density-related effects of ultra large year classes that have plagued haddock assessments.

The selected GB model uses autocorrelated age and year variation around survival and selectivity with a fixed M=0.2 (other specifications noted in TOR 4). The EGB WHAM framework fixes M in the early history and then estimates it for recent years (TOR 4). Both models have achieved the desired result of reducing their retrospective patterns to acceptable levels. Research that revisited catch estimations, survey indices and their changes over the years, age-length and growth updates, exploration of possible mechanisms of recruitment processes and spatial dynamics were explored and incorporated into the base models. Thus, the WG has created acceptable model structures to be carried forward for the Management Track Assessment.

The WHAM framework has been extremely useful in defining the current assessment models. But it also has the potential for improvements in stock assessments in general as well as haddock in particular by expanding the WHAM relationships (such as density-dependent M's); by defining rate models and controlled simulations where deviations from underlying relationships are interpretable as ecological effects; and by linkage between single-species assessments to provide trends in "ecosystem effects" that can provide strategic advice for management (climate change). Presumably this would be an activity of the WHAM WG.

The WG did not fully agree on the best stock identification approach between the two options, entire GB versus EGB alone. In my opinion, the best approach would be to use GB as a base assessment and attempt to link GB and EGB components in the model. The goal is to stabilize the estimation structure such that they will not deviate too much from the common data and information. Then separate management objectives and constraints of the US and Canadian systems should be explored through management strategy evaluations and management procedures.

My final opinion has to do with the Research Track Assessment process in general. Research tracks avoid the "taint" of management by not including the most recent years of catch and by not addressing the details of management. However, all assessments are conducted to provide management advice and the demarcation between what is assessment research and what is management is not so sacrosanct. In my opinion, research track assessments would benefit from TORs that incorporate how management is or is likely to be structured. This would provide some guidance for determination of BRPs, projection approaches and tradeoffs between acceptable accuracy and precision of parameters leading to key management quantities. In the case of haddock in particular, exploration of assessment/management structures leading to compatible TGMC management would have been useful. It is unclear to me that this will be achieved for the next management track assessment.

Appendix 1: Bibliography of materials provided for review

Brooks, E.N, M.L. Traver, S.J. Sutherland, L. Van Eeckhaute, and L. Col. 2008. In. Northeast Fisheries Science Center. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii. http://www.nefsc.noaa.gov/publications/crd/crd0815/

Friedland, K.D., J.A. Langan, S.I. Large, R.L. Selden, J.S. Link, R.A Watson, and J.S. Collie. 2020. Changes in higher trophic level productivity, diversity and niche space in a rapidly warming continental shelf ecosystem. Science of the Total Environment 704: 135270.

Hordyk, A.R., Q. Huynh, and T.R. Carruthers. 2022. *OpenMSE R Package* (1.0.0) [Computer software]. Blue Matter Science. https://cran.r-project.org/package=openMSE.

Legault, C.M. and V.R. Restrepo. 1998. A flexible forward age-structured assessment program. ICCAT. Col. Vol. Sci. Pap. 49: 246–253.

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

NEFSC (Northeast Fisheries Science Center). 2019. Georges Bank haddock 2019 assessment update report. Unpubl. Rpt. 10 pp. nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php

NEFSC (Northeast Fisheries Science Center). In Review. Final Report of the Haddock Research Track Assessment Working Group. Unpubl. Rpt. 65 pp. https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php

Stock, B.C., and T. J. Miller. 2021. The Woods Hole Assessment Model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fisheries Research 240, 105967. https://doi.org/10.1016/j.fishres.2021.105967

TRAC (Transboundary Resources Assessment Committee). 2019. Eastern Georges Bank haddock [5Zjm; 551,552,561,562]. TRAC Status Report 2019/01.

TRAC (Transboundary Resources Assessment Committee). 2021. Eastern Georges Bank haddock [5Zjm; 551,552,561,562]. TRAC Status Report 2021/01.

Appendix 2: Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Eastern Georges Bank and Georges Bank Haddock
Research Track Peer Review

Background

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

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

Scope

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

The purpose of this meeting will be to provide an external peer review of Eastern Georges Bank and Georges Bank and haddock stocks. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: **Appendix 1**: TORs for the research track, which are the responsibility of the analysts; **Appendix 2:** a draft meeting agenda; **Appendix 3:** Individual

[1] https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2005/m05-03.pdf

Independent Review Report Requirements; and **Appendix 4:** Peer Reviewer Summary Report Requirements.

Requirements

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

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

Tasks for Reviewers

- Review the background materials and reports prior to the review meeting
    - Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the Peer Review Panel (co)Chair with contributions to the Peer Reviewer Summary Report
- Deliver individual Independent Reviewer Reports to the Government according to the specified milestone dates
- This report should explain whether each research track Term of Reference was or was not completed successfully during the peer review meeting, using the criteria specified below in the "Tasks for Peer Review Panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

- During the peer review meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the Peer Review Panel chair shall identify or facilitate agreement among the reviewers for each research track TOR.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for Peer Review Panel chair and reviewers combined:

Review the Report of Haddock Research Track Working Group.

1) The Peer Review Panel (co)Chair, with the assistance from the reviewers, will write the Peer Reviewer Summary Report. Each reviewer and the (co)chair will discuss whether they hold similar views on each research track Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the peer review meeting. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such opinions. Reviewers are not required to reach a consensus.

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

Place of Performance

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

Period of Performance

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

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

Schedule	Milestones and Deliverables
Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
March 28-31, 2022	Panel review meeting
Approximately 2 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

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

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

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

Restricted or Limited Use of Data

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

NMFS Project Contact

Michele Traver, NEFSC Assessment Process Lead
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov

Appendix 1 to PWS. Haddock Research Track Terms of Reference

1. Review existing research efforts, data, and habitat information in the Gulf of Maine and Georges Bank, identify any findings relevant to influences of ecosystem conditions on haddock, and consider those findings, as appropriate, in addressing other TORs. For processes that the working group deems important and promising that are not currently feasible to consider quantitatively, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments.

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 being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). Characterize the uncertainty in these sources of data.

4. 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 model, and evaluate the strength and direction of any retrospective pattern(s) in both the current and the previously accepted model. Enumerate possible sources of the retrospective patterns and characterize plausibility, if possible.

5. Update or redefine status determination criteria (SDC point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs.

6. Define the methodology for performing short-term projections of catch and biomass under alternative harvest scenarios, including the assumptions of fishery selectivity, weights at age, maturity, and recruitment.

7. Review, evaluate and report on the status of the Stock Assessment Review Committee (SARC) and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

8. Develop a “Plan B” for use if the accepted assessment model fails in the future.

9. Review and present any research related to recruitment processes (e.g., spawning and larval transport, and retention), and potential hypotheses for large recruitment events.

10. Review and present any research related to density-dependent growth.

11. For Eastern Georges Bank, provide advice to TMGC on appropriate reference points.

12. Review data related to stock structure of haddock on Georges Bank (including Eastern Georges Bank management area) and implications for assessments conducted on the whole bank and on subareas of the bank.

Research Track TORs:

General Clarification of Terms that may be used in the Research Track Terms of Reference

Guidance to Peer Review Panels about "Number of Models to include in the Peer Reviewer Report":

In general, for any TOR in which one or more models are explored by the Working Group, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the Working Group and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Research Track Working Group:

Anyone participating in peer review meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Appendix 2 to PWS. Draft Review Meeting Agenda

{Final Meeting agenda to be provided at time of award}

Eastern Georges Bank and Georges Bank Haddock
Research Track Assessment Peer Review Meeting

April 4 – April 7, 2022

WebEx link: TBD Phone: TBD

DRAFT AGENDA* (v. 1/6/2022)

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

Monday, April 4, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:30 a.m.	Welcome/Logistics Introductions/Agenda/ Conduct of Meeting	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief TBD Panel Chair	
9:30 a.m. - 10:30 a.m.	TOR #2	Liz Brooks, Monica Finley	Catch data
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2 cont.	Liz Brooks, Monica Finley	Catch data
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #3	Liz Brooks, Monica Finley	Survey data
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #3 cont.	Liz Brooks, Monica Finley	Survey data
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	

5 p.m.	Adjourn		

Tuesday, April 5, 2022

Time	**Topic**	**Presenter(s)**	**Notes**
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead TBD, Panel Chair	
9:15 a.m. - 10:30 a.m.	TORs #1 and #9	Kevin Friedland, Liz Brooks, Scott Large	Ecosystem and Recruitment Processes
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TORs #1 and #9 cont.	Kevin Friedland, Liz Brooks, Scott Large	Ecosystem and Recruitment Processes
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TORs #10 and #12	Liz Brooks, Steve Cadrin, Yanjun Wang	Density Dependent Growth and Stock Structure
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TORs #10 and #12 cont.	Liz Brooks, Steve Cadrin, Yanjun Wang	Density Dependent Growth and Stock Structure
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Wednesday, April 6, 2022

Time	**Topic**	**Presenter(s)**	**Notes**
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead	

		TBD, Panel Chair	
9:15 a.m. - 10:30 a.m.	TOR #4	Liz Brooks, Tom Carruthers	Mortality, Recruitment and Biomass Estimates
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #4 cont.	Liz Brooks, Tom Carruthers	Mortality, Recruitment and Biomass Estimates
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TORs #5, #6, and #11	Liz Brooks, Tom Carruthers	BRPs, Projections and EGB Reference Points
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TORs #5, #6, and #11 cont.	Liz Brooks, Tom Carruthers	BRPs, Projections and EGB Reference Points
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Thursday, April 6, 2022

Time	**Topic**	**Presenter(s)**	**Notes**
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead TBD, Panel Chair	
9:15 a.m. - 10:30 a.m.	TOR #8	Liz Brooks, Tom Carruthers	Alternative Assessment Approach
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #7	Brian Linton	Research Recommendations

11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 2:30 p.m.	Follow-ups/Key Points	Review Panel	
2:30 p.m. - 5 p.m.	Report Writing	Review Panel	
5 p.m.	Adjourn		

Friday, April 7, 2022

Time	**Topic**	**Presenter(s)**	**Notes**
9 a.m. - 5 p.m.	Report Writing	Review Panel	

Appendix 3 to PWS. Individual Independent Peer Reviewer Report Requirements

1. The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the Peer Reviewer Summary Report.

    a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the Peer Reviewer Summary Report that they believe might require further clarification.

d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

    Appendix 1: Bibliography of materials provided for review
    Appendix 2: A copy of this Performance Work Statement
    Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Appendix 4 to PWS. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the Research Track Peer Review Panel chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the peer review meeting. Following the introduction, for each assessment /research topic reviewed, the report should address whether or not each Term of Reference of the Research Track Working Group was completed successfully. For each Term of Reference, the Peer Reviewer Summary Report should state why that Term of Reference was or was not completed successfully.

To make this determination, the peer review panel chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and peer review panel chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

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

3. The report shall also include the bibliography of all materials provided during the peer review meeting, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.

The report shall also include as a separate appendix the assessment Terms of Reference used for the peer review meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Allen (Rob) Kronlund (CoChair), Interface Fisheries Consulting, Ltd.
Richard Merrick (CoChair), NOAA retired
Anders Nielsen, Technical University of Denmark
Joseph Powers, NOAA retired
Kevin Stokes, Stokes.Net.NZ Ltd