Center for Independent Experts (CIE) Butterfish and Shortfin Squid Research Track Stock Assessments Independent Peer Review, March 7-11, 2022

Individual Reviewer Report

Robin Thomson, CSIRO, Australia

1 Executive Summary

The work presented for both the Butterfish and *Illex* Review were of a high scientific standard. Both stocks present considerable challenges for stock assessment. The Butterfish and *Illex* Working Groups have not yet found solutions to all of those challenges, however considerable progress has been made.

For Butterfish, all Terms of Reference (ToRs) were met, with only one (TOR 4) considered to have been partly rather than adequately met. The proposed assessment model and method for making short-term projections are accepted. The research recommendations made by the Butterfish Working Group are appropriate, and several additional recommendations that were proposed by the Review Panel are repeated in this Individual Review report. The Review Panel were in unanimous agreement on all points regarding Butterfish.

The *Illex* stock presents considerable challenges for stock assessment and conventional management approaches, such that not all ToRs could be fully met. Lifespan is short for this species (approximately 7 months), and cohort-specific growth rates and abundance can be highly variable due to high plasticity in abundance, growth and maturity resulting from strong environmental influence. Measurements made from one cohort might not be a good reflection of the attributes of subsequent cohorts, and the short-life span mean that the usual time lag between gathering data, performing quantitative assessments, setting catch limits, and the fishing season to which those catch limits apply, can be as long as 4 *Illex* lifespans. Considerable work of high quality has been performed but, due to the nature of the stock, models can provide only broad guidance regarding the plausible range of scenarios (e.g., stock abundance and fishing pressure relative to stock size) that might lead to the observed data. The Review Panel agreed that the general depletion model (GDM) in its current form cannot be used to provide management guidance due primarily to convergence issues that suggest over-parameterization. However, the Panel differed on whether the GDM was likely to provide useful estimates given more work and finer-scale data. My view is that collecting data to support the model using a daily time step will increase the available data but will also increase the number of model parameters to estimate (migrations into and out of the population) and is therefore unlikely to solve the convergence issues. However, the variants of the GDM that did converge can be added to the suite of models presented by Rago and used as an informal ensemble to indicate that the envelope of possibilities mainly encompass situations in which overfishing is not occurring. I recommend work towards the development an MSE approach, using a Harvest Control Rule based on one or more empirical (in-season) measurements that are likely to be indicative of stock abundance.

2 Background

This report presents my (Robin Thomson's) individual response to the Terms of Reference of the 2022 Research Track stock assessments for Butterfish (*Peprilus triacanthus*) and Northern shortfin squid (*Illex illecebrosus*). A separate Summary Report of the views of the Review Panel was previously provided. Similarity between my Individual report and that of the Review Panel is inevitable, given that I was an active participant in the drafting of the Panel Report and provided the wording for some sections. The Review Panel were in agreement on all but one point, which is discussed below. For expedience, I have copied paragraphs of the Panel Report that present procedural information such as the list of presenters and participants of the video meetings. All of the opinions expressed in this report are my own. Appendix A gives a bibliography of materials provided for the review; Appendix B is a copy of the Performance Work Statement that was provided to reviewers (and that has Appendices of its own).

The Research Track Review meetings for Butterfish and *Illex* squid took place via WebEx video meetings during the week of 7-11 March 2022 EST (8-12 March 2022 EAST).
The Panel was composed of three scientists selected by the Center for Independent Experts (CIE): Yong Chen (SUNY Stonybrook), Robin Cook (University of Strathclyde) and Robin Thomson (CSIRO). The Panel was chaired by Mike Wilberg, as a member of the Mid-Atlantic Fishery Management Council's Scientific and Statistical Committee. The Panel was assisted by Michele Traver (NEFSC's Stock Assessment Process Lead) and Russ Brown (Chief, NEFSC Population Dynamics Branch).

Prior to the meetings, on 23 Feb 2022 EST, assessment documents were made available to the Panel through a NEFSC website (https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php). On 3 March 2022 (EST) I met briefly with Michele Traver and Russ Brown to review and discuss the meeting agenda, Panel reporting requirements, meeting logistics and the overall process. Due to incompatible time zones a separate meeting was held with the remaining members of the Review Panel, Yong Chen and Robin Cook.

The meeting opened on 12:00 EST, Monday March 7, with welcoming remarks and comments on the agenda by Russ Brown, Michele Traver, and Panel Chair Mike Wilberg. The first two days of the meeting focused on presentations and discussion of the 10 ToRs for the Butterfish 2022 research track assessment, and the second two days focused on the 11 ToRs for *Illex*.

All reviewers attended all video meetings and actively participated by questioning presenters and participating in subsequent discussions. Towards the end of the second day the Review Panel met to discuss their findings in an open meeting that was attended by most of that day's participants. During that meeting the reviewers addressed some questions to Presenters to clarify certain points. Towards the end of the fourth day the Review Panel met in a closed session, along with Michele Traver and Russ Brown who provided guidance only on aspects of the review process. The review panel meetings were cordial and co-operative with agreement on most points.

Documentation was prepared by the Butterfish and *Illex* Working Groups (see Appendix A), and presentations were made by Charles Adams, Andrew Jones, Jason Didden, Tori Kentner, Eric Robillard, Laurel Smith, and Rob Vincent for Butterfish and Lisa Hendrickson, Brooke Lowman, Jessica Jones, Sarah Salois, Paul Rago, John Manderson, and Anna Mercer for *Illex*. Members of the Working Groups and public also provided valuable discussion. Jason Boucher, Tony Wood, Russ Brown, Ben Levy, Brian Linton, Toni Chute, Laurel Smith, and Abigail Tyrell (all from the NEFSC) acted as rapporteurs throughout the meeting.

3 Response to the Terms of Reference for Butterfish (*Peprilus triacanthus*)

3.1. 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 adequately addressed.

Catches of Butterfish are thought to have peaked during the 1970s when foreign fleets were operating in the area. Discards are thought to have been non-negligible but accurate data are not available for this period for either discards or landings. After the phasing out of foreign catches (after 1987) records of landed catches are thought to be accurate. Estimates of discard rates are available from observer data and these are used to calculate a total discard tonnage for the year through the application of a method that scales observed fishing discards to all fishing operations by estimating the Butterfish discard percentage of the total landed catch for all species and applying that discard to all landings for the year. This is a reasonable method to us and is commonly used in the region. However, further work to investigate factors that influence the discard rates of Butterfish (e.g., spatial zone, fleet or metier, market prices of Butterfish and other species caught with Butterfish) could be used to further refine the estimates. A model-based estimate of Butterfish discards might better account for such variables. If information on the number of fishing shots is available, then an alternative method is to scale the observed discard fraction per shot in the fishery to the total number of shots (even those that did not result in landings of Butterfish). This method avoids the influence of fluctuations in the abundance or market for other species.

An attempt was made to apply assessment models to data for the foreign fishing period but this was not successful. The decision was made to confine the assessment to the period where catches and discards are more precisely known. This was a reasonable decision, particularly given an apparent environmentally-driven shift in productivity in the region which renders the earlier period of less relevance to current management.

Yong Chen noted a shift in landings taken from Statistical Area (SA) 537 off the RI coast to SA 526 off MA. Investigation of the cause of this shift and whether it results from changes in the distribution of the stock or of the fishery is recommended.

The assessment models that were presented all require information on the catches-at-age in both the fishery and the surveys. This information was derived by constructing Age-Length-Keys (ALKs) and multiplying those by the length-at-age information collected from port samples of length and age data. Gaps in both the length frequencies and the ALKs were filled using methods commonly used in the region. The methods that were used are reasonable, however it would be greatly preferable to fit the assessment model to the length and age-at-length data (Lee et al. 2019). Gap filling can obscure information on the relative sizes of cohorts and, because the ‘filling’ occurs outside of the assessment model, uncertainty due to the missing data is not propagated into the assessment. Estimates of uncertainty from the assessment are therefore too small. Fitting to length data would require estimation within the model of growth rates, which might be environmentally driven, which might increase model complexity too much. Nevertheless, I encourage exploration of the possible utility of such a model.

Data presented to the Reviewers indicated that discarded fish tend to be smaller than retained fish. It is therefore desirable to give the model information on the age (or preferable size) of the discarded fraction separately from that of the landed fraction. Moreover, targeted fishing of Butterfish was stopped between 2002 and 2012 – the pattern of size-based discarding during that period might differ and, if it does, this should be included in the model.

3.2. Present the survey data available (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.), and describe the basis for inclusion or exclusion of those data in the assessment. Characterize the uncertainty in these sources of data.

This ToR was adequately addressed.

The Working Group investigated the large number of surveys that were available from federal and state organizations. The federal surveys had best overlap with Butterfish distribution and provided age information. The state surveys provided only length data, so the data were used for only those fish young enough to have age assigned from length (i.e., age 0). A joint young-of-the year (YOY) index was calculated from survey data for which at least 50% of the catch was Butterfish. This was an appropriate use of the available information, given the assessment model’s need for catch-at-age data. As noted above, it is advisable to fit assessment models to length composition and conditional age-at-length data where possible. If this can be done for Butterfish, then state survey data for all lengths could possibly be used.

A seasonal shift in Butterfish distribution is evident, resulting in a more widespread distribution over the shelf during the Spring. For that reason, Albatross spring surveys should be included only for a sensitivity analysis as the model would misinterpret the greater catches on the shelf during Spring as increased abundance whereas it is likely the result of increased availability at that time.

A significant challenge to assessment of the Butterfish stock comes from its apparent differential distribution by age and season and that this distribution seems to have changed over time. Further investigation of the environmental factors that influence that distribution might help to better disentangle signals of abundance from those of availability. Combining data from the available state and federal surveys might help to identify influential factors and how these impact Butterfish distribution.

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

This ToR was adequately addressed.

Both statistical catch at age and state space models were applied to data for Butterfish. The time series was confined to only the period of domestic fishing when adequate catch and discard information are available. The models used an AR(1) process to describe recruitment rather than a stock recruit relationship. During the model period, fishing seems not to have greatly impacted the stock so that the absolute size (scale) of the stock is poorly estimated. To achieve convergence a catchability (q) parameter for one of the surveys had to be specified. This choice will largely drive the model estimate of stock abundance and should therefore be explored using sensitivity tests. The work on predator consumption (discussed below) raises questions about whether the scale of the Butterfish stock in the model is higher than it is in reality. Because of the lack of impact of catches on the stock, a stock recruit relationship would not be estimable because recruitment strength has not been observed over an adequate range of abundance. If one were to be used, at least one of its parameters would therefore have to be pre-specified rather than estimated. Given the environmentally driven changes in distribution and recruitment strength observed within the region, an autoregressive time series approach such as that presented by the Working Group seems a sensible choice (although questions regarding its use for forecasting are discussed below).

The models that were applied all used an annual time step, which seems long for a species that is represented mostly at ages 0 to 3 in the catches. The use of a model start date of 1 January in contrast to mid-year spawning creates a disconcerting mismatch in which fish are present in the model as zero year olds for only 6 months. A shorter model time step (of at least 6 monthly, and perhaps 3 or 4 monthly) might better reflect the dynamics. Alternatively, a ‘model year’ that begins at the time of spawning could be trialed.

The WHAM model estimated the selectivity-at-age for each age group separately, not constrained by the use of a functional form for selectivity-at-age. Such a formulation could allow the model to ‘hide’ an incorrect natural mortality rate. Sensitivity tests that force selectivity to follow a pre-specified functional form (in particular, a logistic form) could help to reveal mismatches, especially if M were estimated for alternative choices of selectivity function.

The decision was made to present the Review Panel with only the accepted model and a relatively small subset of alternative models that were considered. Presumably, a larger number of sensitivity trials were conducted than were seen by the Panel. It would have been useful to see a wider range of tests of the final WHAM model as a way to better understand the compatibility of the model and data with alternative states of nature and hypotheses. Such tests include:

- a plausible range of alternative values for the fixed value of catchability (q) for the Albatross surveys,
- exclusion of the Albatross survey,
- alternative (and age dependent) values for natural mortality (M),
- earlier start years for the model,
- the use of functional forms for selectivity-at-age (especially logistic).

3.4. Update or redefine status determination criteria (SDC point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This ToR was partially addressed. The WG considered several potential candidate reference points and recommended $F_{50\%}$ and $B_{50\%}$, for fishing mortality and biomass, respectively. The resulting fishing mortality rate was concerningly very high, corresponding with 99.9% mortality for fully selected ages. Such a high value would be of concern for any stock. The reference point used in the past was 2/3 M, which seems more appropriate because it provides a lower and therefore more precautionary fishing rate. However, it must be noted that the uncertainty in the true size of the stock and the value of M mean that any increase in catch for this stock should be implemented carefully and its impact on the stock closely monitored.

The uncertainty in the estimated reference points ought to include consideration of the uncertainty in populations size (currently fixed by the chosen catchability value for the NEFSC fall trawl) and in the value of M.

The use of an AR(1) process instead of a stock recruit relationship, although seemingly appropriate for this stock, does lead to questions regarding how to implement future projections and how to calculate reference points. The WG's choice to project forward using the average recruitment over the past 5 years seems reasonable, as does the choice to use a recent average to calculate the reference points. However, if changing environmental conditions result in a continuous downward trend in recruitment then this choice could result in the calculation of reference points that also tend downwards. Careful consideration, particularly of the threshold reference point, is recommended. The use of 'dynamic reference points' for a species whose productivity is declining could allow the stock to be fished to a low level relative to past abundance (albeit not relative to a declining reference point). This is of particular concern for a prey species such as Butterfish. It might be necessary to consider the needs of predators, at least in fixing a threshold reference point, in terms of absolute rather than relative abundance.

Yong Chen was concerned that the justification of a "Schaefer production function" for B50% is not warranted because a production function is not used to estimate the reference point (i.e., a Schaefer model is not being used). Although I agree, I am not greatly concerned by this. The choice of 50% is relatively arbitrary. Of more concern to me is that 50% might not be a sufficiently conservative choice for a forage fish. Butterfish are a prey species so that reducing their abundance to a low level might have consequences for several other species. Reference points above 50% of unfished levels have been adopted for other prey species for that reason (e.g., 75% for Antarctic krill).

3.5. Make a recommended stock status determination (overfishing and overfished) based on new modeling approaches developed for this peer review.

This ToR was adequately addressed.

The consensus of the evidence suggests that the stock is not overfished or experiencing overfishing. This conclusion is likely to be robust to the major sources of uncertainty including those expressed above about reference point estimation. More precise determination of stock status (i.e., current biomass as a percentage of unfished biomass) is not determinable with acceptable accuracy given the uncertainty regarding stock size resulting from the apparent lack of impact of the fishery during the modelled period.

3.6. 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, and maturity.

This ToR was adequately addressed.

Short term (three year) projections were performed using the '17-NAA5' and '04-NAA2' WHAM models. Weight- and maturity-at-age for the projection were set equal to average over the most recent 5 years. The AR(1) process used by the 17-NAA5 model was projected into the future with the autocorrelation dampening over time. The '04-NAA2' model treats recruitments as fixed effects drawn from a distribution with mean and variance calculated from a subset of model years. The 17-NAA5 WHAM model is the recommended model. The projections used future (fully selected) F values of 0, F_{2019} and the relatively high F_{proxy} value.

The methods used to perform the short-term future projections seem appropriate, and given the short-lived nature of the stock it is appropriate to perform projections of only 3 years.

3.7. 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, as well as the most recent management track assessment report. Identify new research recommendations.

This ToR was adequately addressed.

The research recommendations are appropriate. In particular, the recommendation to continue working on incorporating environmental covariates is encouraged.

I support the Review Panel recommendation to expanding the research recommendation (2) "Explore the possibility of spawning south of Cape Hatteras, NC and potential contribution to the northern stock" to include a general study of stock structure and distribution that includes the different life history stages (e.g., revisit larval survey data and how well they match the older life stages).

Below, I repeat (and support) the new research recommendations suggested by the Panel:

1. Conduct a new evaluation of survey catchability. The current value of q is based on an analysis of habitat distribution to estimate availability to the survey. A q~0.2 implies that 80% of the stock is not within the survey area, which seems potentially problematic given that Butterfish are widely caught throughout the survey that covers most of their range.

2. Consider alternative ways of calculating total discards. Current estimates raise samples based on a ratio estimator that uses total fish catch as the denominator. Raising using number of trips or shots (or other effort measures) are possible alternatives. Applying a time series smoother to the ratio estimator may be able to exploit information across years to improve estimates.

3. Investigate whether environmental variables or time varying catchability can be applied to the spring Albatross/Bigelow survey so that it can be included in the assessment. Time varying catchability should be estimable within an assessment.

4. Consider an age- and length-structured model that allows increased use of the state survey data (by including all the length data). This may help to avoid the need for gap-filling. The derived data that are used to fill gaps will give a false sense of precision and the approach is likely to over-smooth estimates of recruitment.

5. Consider alternative (area, or habitat, weighted) averaging for the aggregated state survey YOY index. The Conn model used by the assessment team assumes a common signal across multiple areas and cannot, therefore, take into account spatial effects that might be important.

6. Consider implementing a wider range of assessments/data processing to understand effects of decisions and provide a basis for ensemble modeling. Recent work at the SEDAR68 assessment of Atlantic scamp grouper implemented methods for ensemble modeling while diagnostics developed by Carvalho et al. (2021) have been used to weight models from an ensemble in order to obtain estimates of uncertainty for quantities of interest.

7. Develop a wider range of diagnostics for state-space models (e.g., plots of the random effects predictions). Include MCMC methods to estimate posterior distributions of critical parameters and quantities of interest, e.g., $F_{50\%}$ and $B_{50\%}$.

8. Consider alternative model selection criteria that are more appropriate for mixed-effects models. AIC was used to inform model selection, but this may not be appropriate where random walk models reduce the number of effective parameters. DIC and WAIC may prove more appropriate in these circumstances.

9. Consider developing an age- or size-dependent M. The current value used is a mean value over all ages/sizes but it is highly likely that M is greatest on the youngest fish. Mis-specification of M by size may lead to biased estimates of selectivity and hence BRPs. One common approach is to scale the Lorenzen weight-based Ms to the overall mean derived from meta-analyses.

10. Consider using stomach contents data to inform time-varying M. Data were presented at the review meeting on consumption of Butterfish by marine mammals, birds and some fish. These data may offer an insight into temporal effects on M.

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

This ToR was adequately addressed.

The 'Plan B smoother' was applied. The ASAP and WHAM models are superior to the Plan B model and the preferred WHAM model 17-NAA5 is an appropriate choice. The "Plan B" model is not needed, but would be an adequate replacement if, for some reason, none of the ASAP or WHAM models could be used.

3.9. Additional Terms of Reference

3.9.1. Describe life history characteristics and the stock's spatial distribution, including any changes over time. Describe ecosystem and other factors that may influence the stock's productivity and recruitment. Consider any strong influences and, if possible, integrate the results into the stock assessment.

This ToR was adequately addressed given the time available. However, further work (see Research Recommendations; ToR 7) is recommended to better understand these influences and to integrate them into an assessment model.

3.9.2. Evaluate consumptive removals of Butterfish by its predators, including (if possible) marine mammals, seabirds, tunas, swordfish and sharks. If possible, integrate results into the stock assessment.

This ToR was adequately addressed. Additional work on tunas (e.g., bluefin), swordfish and sharks is possible although might not be a high priority.

The estimated (sometimes assumed) natural mortality rate for Butterfish seems very high. It was interesting that attempts to account for the resulting tonnage of fish that appear to die each year fell well short in terms of estimated tonnage consumed by known predators. This suggests that the stock size in the model (largely driven by the choice of q parameters for one of the surveys) might be too large.

Butterfish are likely to be high in oils (hence the name) suggesting that they should be a 'high value' prey item, yet they seem under-represented in consumption studies. Some of those studies (used, in part, to list the likely predators) are based on inspection of stomach contents and might be subject to rapid deterioration of Butterfish. More modern approaches such as DNA and isotope detection might provide differing results.

4 Northern shortfin squid (*Illex illecebrosus*)

4.1. Estimate catches from all sources, including landings and discards, and characterize their uncertainty.

This ToR was adequately addressed.

Like Butterfish, the size of catches and discards by foreign fleets earlier in the time series is not known with sufficient precision to support quantitative approaches prior to the beginning of a purely domestic fishery. Estimates of catch and discard are thought to be sufficiently accurate from 1997-2019 (the terminal year of the assessment) due to mandatory reporting requirements and increased fishery observer coverage. Given the short lifespan of *Illex* (approximately 7 months) this represents many generations. Discards for this period are adequately estimated and relatively low compared to the retained U.S. fishery catch (~7%). Canadian catches (assumed zero discards due to jig gear) and US fishery catches were obtained. Newfoundland recreational catches are unknown, because the fishery is unmanaged, but are likely small relative to the US fisheries.

4.2. Evaluate indices used in the assessment, including annual abundance and biomass indices based on research survey data and standardized industry CPUE data. Characterize the uncertainty of the abundance and biomass index estimates. Explore the relationship between fishing effort and economic factors (e.g., global market price) in order to determine whether the addition of an economic factor will improve the fit of the CPUE standardization model.

This ToR was adequately addressed.

The WG explored a range of state, regional and federal surveys. Extensive analyses (GAMs and GLMs) of the fishery dependent landings per unit effort (LPUE) data were conducted. Annual standardized indices from the GLM were significantly correlated with the NEFSC fall trawl survey biomass indices since 2008. Economic factors were considered in the GAM and average weekly price was identified as the most important of the variables considered, but unlike the GLM, the GAMs were run separately for each of the two fleets rather than combining both in the same model run. The analyses used were appropriate to address the ToR.

4.3. Utilize the age, size and maturity dataset, collected from the 2019 landings, to identify the dominant intra-annual cohorts in the fishery and to estimate growth rates and maturity ogives for each cohort. Also use these data to identify fishery recruitment pulses.

This ToR was adequately addressed.

Data were collected during 2019 and 2020. Daily age readings of statoliths were undertaken and two dominant modes were identified corresponding with a dominant winter (Nov.-April) and a summer (May-July) cohort. Due to the timing of sample collection, the summer cohort could be seen in only the 2020 data. A similar finding regarding dominant winter and summer or spring cohort were shown by a May 2000 study. Although these cohorts are dominant, spawning does seem to occur continuously throughout the year.

Sample sizes of mature females from the 2019 and 2020 collection were too small to estimate cohort-specific maturity ogives. *Illex* growth and maturation rates appear to be highly environmentally driven so that measurements of these rates can provide an indication of the observed ranges of cohort-specific rates but might not be strongly predictive of the rate of a cohort that will support fishing in a given future season.

4.4. Characterize annual and weekly, in-season spatio-temporal trends in body size based on length and weight samples collected from the landings by port samplers and provided by Illex processors. Consider the environmental factors that may influence trends in body size and recruitment. If possible, integrate these results into the stock assessment.

This ToR was adequately addressed.

Data from processors and port samples collected during 1997-2019 were considered along with samples taken during surveys. The fishery mean body weight measurements do not show the gradual decline seen in the survey series. The survey estimates of mean body weight are positively correlated with relative abundance.

Environmental factors were also considered in an LPUE model of the study fleet and observer data, and ten influential factors were identified. Of those, the significant environmental factors were bottom temperature, ring footprint index, ring orientation, salinity at the 222 meter isobath, chlorophyll frontal activity, and standard deviation in sea surface temperature. The differing influence of the environmental factors on recruitment, body size, and availability could not be teased apart using this approach.

4.5. Develop a model that can be used for estimation of fishing mortality and stock biomass, for each dominant cohort that supports the fishery, and estimate the uncertainty of these estimates. Compare the results from model runs for years with low, medium and high biomass estimates.

This ToR was partially addressed.

The very short life-span of *Illex* present a particularly challenging problem for fishery management, given that the individuals that were observed (i.e., sampled, surveyed or caught by the commercial fishery) are dead by the time management recommendations are made for the next fishing season. Observed characteristics of the cohorts sampled are relevant for the management of subsequent cohorts only if there is strong correlation between successive cohorts and this correlation holds for the four cohorts that are likely to pass between measurement and implementation of management. For this reason, the most appropriate models are likely to be those that attempt to calculate probable envelopes of likely values of quantities of interest. This is the type of modelling that was performed by Rago, which is therefore appropriate to address this ToR.

The general depletion model (GDM) is a somewhat different approach, which attempts to calculate specific quantities of interest for particular cohorts and which is presented as an approach that could be used (given fine scale data collection) for within-season management. Reviewers agreed that the model, as presented to the panel, has convergence issues and is not suitable for management in its current form. My opinion (which differs somewhat from that of Dr Yong Chen) is that the GDM does not present a useful approach for management even if more work is performed to improve its convergence problems or given finer scale data. The need to identify and estimate the size of incoming and outgoing cohorts suggests that reducing the size of the time step will increase the numbers of such cohorts to estimate and will not, therefore, improve the data to parameter ratio. Dr Chen and I agree that it would be appropriate to conduct a cost benefit analysis before moving to within season data collection.

4.6. Describe the data that would be needed to conduct in-season stock assessments for adaptive management and identify whether the data already exist or if new data would need to be collected and at what frequency.

This ToR was adequately addressed.

This work primarily focused on the GDM and its data requirements. I suggest a Management Strategy Evaluation approach (see Research recommendations). An MSE would test alternative Management Strategies (MS), and the GDM could be one candidate MS. However, alternative measures including responses to quantities measured during surveys or during fishing that are less arduous to collect than those suggested to support in-season management using the GDM could also be considered.

4.7. Update or redefine Biological Reference Points (BRP point estimates for BMSY, BTHRESHOLD and FMSY) or BRP proxies, for each dominant cohort that supports the fishery, and provide estimates of their uncertainty. If analytical model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing and recommended BRPs or their proxies.

The WG attempted to address this ToR. Too few mature females were available to use the Hendrickson and Hart (2006) per-recruit model to estimate %MSP-based BRPs. In the absence of a reliable analytical stock assessment, it was not possible to identify meaningful BRPs.

4.8. Recommend a stock status determination (i.e., overfishing and overfished), for each dominant cohort supporting the fishery, based on new modeling approaches developed for this peer review.

This ToR was adequately addressed.

In the absence of BRPs and a definitive stock assessment, a formal stock status determination was not possible. However, the working group concluded that the indications from the various assessment approaches were that the stock was lightly fished in 2019. The Panel agreed that this was likely to be the case, but that the term "lightly fished" needs to be interpreted with caution since it has no specific definition relating to sustainable exploitation.

4.9. 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, and maturity.

This ToR was adequately addressed.

Given the biology of the species and current models, short term projections are problematic. The status of current stock assessment models for this stock does not provide a basis for such projections. The WG suggested using Plan B smooth as an alternative and provided examples of the catch multiplier for 2019 that would be estimated from a range of abundance indices. These multipliers (from different indices) were all close to each other and imply that the best estimate of next year's catch is the last observed catch.

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

This ToR was adequately addressed.

The WG considered the recommendations and ranked these in priority based on a poll of members. The Panel supports these recommendations. In view of the extensive research that has been done on the biology of the stock and the limitations of conventional stock assessment models, the Panel suggests the following additional research recommendations:

- (Highest priority) An operating model for the stock and fishery should be developed to allow the testing of potential assessment models and of simple harvest control rules based on abundance indices that would promote sustainable exploitation. The Panel recommends developing the model around a set of hypotheses of *Illex* and fishery dynamics. This platform could also be used to test the GDM. An example of in-season management using an empirical Harvest control Rule based on the result of a pre-recruit survey is given in de Oliveira *et al.* (1998).

- Consider methods for developing projections using environmental correlates, and test their potential performance using an operating model.

- Conduct a study to improve understanding of stock structure (e.g., statolith microchemistry, genetics).

- A cost-benefit analysis of real-time management should be considered.

4.11. Develop a "Plan B" alternate assessment approach to providing scientific advice to managers if the analytical assessment does not pass review.

This ToR was adequately addressed given the limitations of the analyses available. The WG notes that the SSC has used the Rago indirect approach to provide annual ABC and OFL advice.

The application of the Plan B smooth method is only useful with the most up-to-date abundance indices. Thus, the use of the previous year's indices to set limits for the projection year (i.e., 2 years beyond the last abundance index) is probably not appropriate for such a short-lived species. A more responsive approach to make best use of current data is required.

5 Reference list

De Oliveira JAA, Butterworth DS, Roel BA, Cochrane & Brown JP (1998) The application of a management procedure to regulate the directed and bycatch fishery of South African sardine *Sardinops sagax. South African Journal of Marine Science* 19: 449-469.

Lee H, Piner KR, Taylor IG & Kitikado T (2019) On the use of conditional age at length data as a likelihood component in integrated population dynamics models. Fisheries Research 216: 204-211.

6 Appendix 1. Bibliography of materials provided for review

6.1. Provided to the Review Panel for the Butterfish Review

Adams C.F. 2018. Butterfish 2017 Stock Assessment Update. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 18-05 31 p. Available from: http://www.nefsc.noaa.gov/publications/

Adams, C. F. 2017. Age-specific differences in the seasonal spatial distribution of Butterfish (Peprilus triacanthus). ICES Journal of Marine Science (2017), 74(1), 170–179. doi:10.1093/icesjms/fsw128

Adams, C. F. 2022. Update on the Spatial Distribution of Butterfish, 1982–2019

Additional TOR 1: Butterfish Condition

Additional TOR 1: Butterfish Habitat

Butterfish 2021 Research Track Assessment Report

Butterfish (*Peprilus triacanthus*) 2022 Research Track Assessment Report

Dayton, J. E. Robillard. 2021. Age determination, validation and growth of Butterfish *Peprilus triacanthus* along the Mid-Atlantic Bight.

Deroba, J.J. 2022. ReadMe Roadmap for documents prepared during the 2022 Butterfish Research Track Stock Assessment

Didden, J. 2022. Working Paper: Fishery Background and Fishing Industry Perspectives on Butterfish.

Jones, A. et al. 2022. Estimating Butterfish released discards 2014-2019 using observer records.

Northeast Fisheries Science Center. 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-01; 41 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Northeast Fisheries Science Center. 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-03; 383 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/ publications/

Northeast Fisheries Science Center. 2014. 58th Northeast Regional Stock Assessment Workshop (58th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-03; 44 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa. gov/publications/

Northeast Fisheries Science Center. 2014. 58th Northeast Regional Stock Assessment Workshop (58th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-04; 784 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/publications/

Smith, L., M. Lyssikatos, F. Wenzel. 2022. Consumptive Removals of Butterfish by Marine Mammals

Smith, B. E. 2022. Consumption of Butterfish at various life stages by fishes of the Northeast US continental shelf
Spatial distribution of Butterfish commercial catch
Stock, B., T. J. Miller. 2021. A state-space assessment of Butterfish using the Woods Hole Assessment Model (WHAM). NEFSC NMFS, Woods Hole, MA.
Suca, J.J. et al. 2018. Feeding dynamics of Northwest Atlantic small pelagic fishes. Progress in Oceanography 165:52-62.
Suca, J.J. et al. 2021. Environmental drivers and trends in forage fish occupancy of the Northeast US shelf. ICES Journal of Marine Science (2021), https://doi.org/10.1093/icesjms/fsab214
Vincent,R. 2022. Seabird Predation on Butterfish in the Northeast, a Summary of Available Information. MIT Sea Grant, January 2022
Wilberg, M., E. Houde, and F. Serchuk. 2020. 2020 Management Track Peer Review Committee Report.

6.2. Provided to the Review Panel for the *Illex* Review

Hendrickson et al. 2022. Report of the *Illex* 2021 Research Track Assessment Working Group
Hendrickson, L. C. 2004. Population biology of northern shortfin squid (*Illex illecebrosus*) in the Northwest Atlantic Ocean and initial documentation of a spawning area. ICES Journal of Marine Science, 61: 252-266.
Hendrickson, L.C. and D. R. Hart. 2006. An age-based cohort model for estimating the spawning mortality of semelparous cephalopods with an application to per-recruit calculations for the northern shortfin squid, *Illex* illecebrosus 78:4-13
Hendrickson, L.C. 2020. Characterization of Body Weight Data from the Landings of Northern Shortfin Squid (*Illex illecebrosus*) and Preliminary Annual Landings-per-Unit-of-Effort for the Southern (USA) Stock Component
Jones, J. 2022. The Falkland Islands squid fisheries: stock assessment, management and how similar methods can be applied to the US squid fisheries
Lowman, R. A., A.W. Jones, A.J.Mercer. 2022. Integrating economic factors and fleet dynamics into the standardization of northern shortfin squid landings per unit effort.
Manderson et al., 2022. Plausible bounds for availability of and net efficiency for northern shortfin squid in the US fishery and Northeast Fishery Science Center Bottom Trawl Survey
Manderson J.P. A. M. Mercer. 2022. On the application of generalized depletion modeling with an open population assumption to the Northeast US *Illex* illecebrosus fishery.
Manderson et al., 2021. Fishery & Survey availability & net efficiency for *Illex* illecebrosus
Manderson et al. 2022. Impacts of processors on effort and catch.Working paper: Technical and economic aspects of northern shortfin squid (*Illex illecebrosus*) processing and marketing essential for interpreting of fishing effort and catch as indicators of population trend and condition
Manderson, J.P. 2022. Update: generalized depletion modeling
Mercer et al. 2022. Technical and economic aspects of northern shortfin squid (*Illex illecebrosus*) harvesting, processing, and marketing essential for interpreting of fishing effort and catch as indicators of population trend and condition. A Working Paper for the 2021 *Illex* Research Track Stock Assessment
Mercer, A.J. and J. P. Manderson. 2022. Comparative anatomy of the oceanography, ecology and fisheries as they relate to *Illex illecebrosus* and *Illex argentinus* and implications for the

application of real time management approaches developed in the Southwest Atlantic *Illex argentinus* fishery to the Northwest Atlantic *Illex illecebrosus* fishery. A Working Paper for the 2021 Illex Research Track Stock Assessment

Northeast Fisheries Science Center. 1996. Report of the 21st Northeast Regional Stock Assessment Workshop (21st SAW): Stock Assessment Review Comntittee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref Doc. Q6-05d; 200 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

Northeast Fisheries Science Center. 1999. Report of the 29th Northeast Regional Stock Assessment Workshop (29th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref Doc. 99-14; 347 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

Northeast Fisheries Science Center. 2003. Report of the 37th Northeast Regional Stock Assessment Workshop (37th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-16; 597 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026

Northeast Fisheries Science Center. 2006. 42nd Northeast Regional Stock Assessment Workshop (42nd SAW) stock assessment report, part A: silver hake, Atlantic mackerel, and northern shortfin squid (CRD 06-09a). U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 06-09a; 284 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

Rago, P. 2022. Biological Reference Points for *Illex* Squid: Extensions of the Hendrickson and Hart Approach

Rago, P. 2021. Coherence and Persistence of Survey Indices for *Illex* in NAFO Areas 3-6 *Illex* Research Track Assessment Working Group

Rago, P. 2022. Indirect Methods for Bounding Biomass and Fishing Mortality for *Illex* Squid and Implications of an Alternative Quota in 2022

Rago, P. 2022. Mass Balance Approaches for *Illex* Assessment—Stochastic Methods Working Paper to *Illex* Research Track Assessment

Rago, P. 2020. Part 1: Application of Cusum Method for In-Season Detection of Fishery Condition for *Illex* Squid; Landings, 1996-2019

Rago, P. 2020. Part 2: Application of Cusum Method for In-Season Detection of Fishery Condition for *Illex* Squid; Mean Weight, 1997-2019 Working Paper #16 Part 2 for *Illex* Working Group

Salois, et al. 2022. Oceanographic indicators for Northern shortfin squid (*Illex illecebrocus*) in the Northwest Atlantic. A Working Paper for consideration by the 2021 *Illex* Research Track Assessment Working Group

Working paper: Analyses of *Illex* Weight frequencies for evidence of cohorts

Working paper: Goodness of fit analyses for VMS data by 3x3 nm sqr grids Report of Preliminary Analyses to *Illex* Working Group

7 Appendix 2. Performance Work Statement provided to Reviewers

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

Butterfish and Northern Shortfin Squid (Illex)
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

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

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 Butterfish and northern shortfin squid (*Illex*) 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 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. Modifications to the PWS and ToRs cannot be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. 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 both index-based and age-based 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 Reports of the Butterfish and *Illex* Research Track Working Groups.

The Peer Review Panel 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.

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

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 May 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 7-11, 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.

8 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. Research Track Terms of Reference

Butterfish

1. 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.

2. Present the survey data available (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.), and describe the basis for inclusion or exclusion of those data in the assessment. Characterize the uncertainty in these sources of data.

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

4. Update or redefine status determination criteria (SDC point estimates or proxies for BMSY, $B_{THRESHOLD}$, FMSY and MSY) and provide estimates of their uncertainty. If analytic model based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the “new” (i.e., updated, redefined, or alternative) BRPs.

5. Make a recommended stock status determination (overfishing and overfished) based on new modeling approaches developed for this peer review.

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, and maturity.

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, as well as the most recent management track assessment report. Identify new research recommendations.

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

Additional Terms of Reference

1. Describe life history characteristics and the stock's spatial distribution, including any changes over time. Describe ecosystem and other factors that may influence the stock's productivity and

recruitment. Consider any strong influences and, if possible, integrate the results into the stock assessment.

2. Evaluate consumptive removals of Butterfish by its predators, including (if possible) marine mammals, seabirds, tunas, swordfish and sharks. If possible, integrate results into the stock assessment.

Illex

1. Estimate catches from all sources, including landings and discards, and characterize their uncertainty.

2. Evaluate indices used in the assessment, including annual abundance and biomass indices based on research survey data and standardized industry CPUE data. Characterize the uncertainty of the abundance and biomass index estimates. Explore the relationship between fishing effort and economic factors (e.g., global market price) in order to determine whether the addition of an economic factor will improve the fit of the CPUE standardization model.

3. Utilize the age, size and maturity dataset, collected from the 2019 landings, to identify the dominant intra-annual cohorts in the fishery and to estimate growth rates and maturity ogives for each cohort. Also use these data to identify fishery recruitment pulses.

4. Characterize annual and weekly, in-season spatio-temporal trends in body size based on length and weight samples collected from the landings by port samplers and provided by Illex processors. Consider the environmental factors that may influence trends in body size and recruitment. If possible, integrate these results into the stock assessment.

5. Develop a model that can be used for estimation of fishing mortality and stock biomass, for each dominant cohort that supports the fishery, and estimate the uncertainty of these estimates. Compare the results from model runs for years with low, medium and high biomass estimates.

6. Describe the data that would be needed to conduct in-season stock assessments for adaptive management and identify whether the data already exist or if new data would need to be collected and at what frequency.

7. Update or redefine Biological Reference Points (BRP point estimates for BMSY, $B_{THRESHOLD}$ and FMSY) or BRP proxies, for each dominant cohort that supports the fishery, and provide estimates of their uncertainty. If analytical model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing and recommended BRPs or their proxies.

8. Recommend a stock status determination (i.e., overfishing and overfished), for each dominant cohort supporting the fishery, based on new modeling approaches developed for this peer review.

9. 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, and maturity.

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

11. Develop a “Plan B” alternate assessment approach to providing scientific advice to managers if the analytical assessment does not pass review.

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. Draft Review Meeting Agenda

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

Butterfish and *Illex*
Research Track Assessment Peer Review Meeting

March 7 - 11, 2022

WebEx link: https://www.google.com/url?q=https://noaanmfs-meets.webex.com/noaanmfs-meets/j.php?MTID%3Dm8a1062743b689f38d340622b4c9367ff&sa=D&source=calendar&ust=1635700378125529&usg=AOvVaw3Ehp4lawC73ceuFcaRdmaC
Phone: +1-415-527-5035 US Toll

DRAFT AGENDA* (v. 11/18/2021)

**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, March 7, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:30 a.m.	Welcome/Logistics Introductions/Agenda/ Conduct of Meeting Butterfish	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief Mike Wilberg, Panel Chair	
9:30 a.m. - 10:30 a.m.	TORs #1 and A1	Charles Adams, Andrew Jones, Kiersten Curti	Catch Spatial Distribution
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TORs #2 and A2	Charles Adams, Laurel Smith, Rob Vincent	Survey Data Consumptive Removals
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. - 3:30 p.m.	TORs #3 and A1	Charles Adams, Laurel Smith	F, R, SSB Productivity
3:30 p.m. - 3:45 p.m.	Break		
3:45 p.m. - 4:45 p.m.	TOR #4	Charles Adams	BRPs
4:45 p.m. - 5:15 p.m.	Discussion/Summary	Review Panel	
5:15 p.m. - 5:30 p.m.	Public Comment	Public	
5:30 p.m.	Adjourn		

Tuesday, March 8, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Mike Wilberg, Panel Chair	
9:15 a.m. - 10:30 a.m.	TORs #4 cont. - 5	Charles Adams	BRPs Stock Determination
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TORs #5 cont. - 6	Charles Adams	Stock Determination Projections
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. - 3:30 p.m.	TORs #6 cont. - 8	Charles Adams	Projections Research Recommendations Alternative Approach
3:30 p.m. - 3:45 p.m.	Break		
3:45 p.m. - 4:45 p.m.	TOR #6 - 8 cont.	Charles Adams Jason Didden	Projections Research Recommendations Alternative Approach

			Outreach
4:45 p.m. - 5:15 p.m.	Discussion/Summary	Review Panel	
5:15 p.m. - 5:30 p.m.	Public Comment	Public	
5:30 p.m.	Adjourn		

Wednesday, March 9, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics *Illex*	Michele Traver, Assessment Process Lead Mike Wilberg, Panel Chair	
9:15 a.m. - 10:45 a.m.	TORs #1 and 2	Lisa Hendrickson	Landings and Discards Surveys and Fishery CPUE
10:45 a.m. - 11 a.m.	Break		
11 a.m. - 12 p.m.	TORs #2 cont. and 3	Brooke Lowman Lisa Hendrickson	Surveys and Fishery CPUE 2019 age, size and maturity data
12 p.m. - 12:30 p.m.	Discussion/Summary	Review Panel	
12:30 p.m. - 12:45 p.m.	Public Comment	Public	
12:45 p.m. - 1:45 p.m.	Lunch		
1:45 p.m. - 3:30 p.m.	TORs #3 cont. and 4	Lisa Hendrickson Kim Hyde and Sarah Salois	2019 age, size, maturity data Fishery body size
3:30 p.m. - 3:45 p.m.	Break		
3:45 p.m. - 4:45 p.m.	TOR #5	Paul Rago	Stock size Fishing mortality
4:45 p.m. - 5:15 p.m.	Discussion/Summary	Review Panel	
5:15 p.m. - 5:30 p.m.	Public Comment	Public	
5:30 p.m.	Adjourn		

Thursday, March 10, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Mike Wilberg, Panel Chair	
9:15 a.m. - 10:45 a.m.	TOR #5 cont. and 6	Lisa Hendrickson	Fishing mortality In-season data
10:45 a.m. - 11 a.m.	Break		
11 a.m. - 12 p.m.	TORs #7 and 8	Lisa Hendrickson	BRPs Stock status
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 # 9 - 11	Lisa Hendrickson	Projections Research Recommendations Alternative approach
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 3:45 p.m.	Discussion/Summary	Review Panel	
3:45 p.m. - 4 p.m..	Public Comment	Public	
4 p.m. - 5p.m.	Follow-ups/Key Points	Review Panel	
5 p.m.	Adjourn		

Friday, March 11, 2022

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

Appendix 3. 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. 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.

9 Appendix 3. Panel membership and attendees

Illex/Butterfish Research Track Peer Review Attendance
March 7-11, 2022
Attendance

NEFSC - Northeast Fisheries Science Center
GARFO - Greater Atlantic Regional Fisheries Office
NEFMC - New England Fisheries Management Council
MAFMC - Mid-Atlantic Fisheries Management Council
SMAST - University of Massachusetts School of Marine Science and Technology
MIT - Massachusetts Institute of Technology
VIMS - Virginia Institute of Marine Science
SSC - Science and Statistical Committee
~~~~~~
Mike Wilberg - Chair
Robin Cook - CIE Panel
Robin Thomson - CIE Panel
Yong Chen - CIE Panel

Russ Brown - NEFSC
Michele Traver - NEFSC

Abigail Tyrell - NEFSC
Alan Bianchi -North Carolina Department of Environmental Quality
Alex Hansell - NEFSC
Andrew Jones - NEFSC
Anna Mercer - NEFSC
Ben Levy - NEFSC
Brandon Muffley - MAFMC Staff
Brian Linton - NEFSC
Brian Smith - NEFSC
Brooke Lowman - Virginia Marine Resources Commission
Carly Bari - GARFO
Charles Adams - NEFSC
Chris Legault - NEFSC
David Richardson - NEFSC
Eric Reid - Fisheries Consultant
Eric Robillard - NEFSC
Greg DiDomenico - Lunds Fisheries
~~~~~~

Jason Boucher - NEFSC
Jason Didden - MAFMC Staff
Jeff Kaelin - Lunds Fisheries
Jessica Jones - NEFSC post doc
Jim Gartland - VIMS
Jon Deroba - NEFSC
John Manderson - Open Ocean Research
Katie Almeida - Town Dock
Kathy Sosebee - NEFSC
Kiersten Curti - NEFSC
Kim Hyde - NEFSC
Larry Alade - NEFSC
Laurel Smith - NEFSC
Lisa Hendrickson - NEFSC
Mark Terceiro - NEFSC
Meghan Lapp - Sea Freeze Ltd.
Michelle Duval - MAFMC Member/private consultant for Mellivora Consulting
Mike Simpkins - NEFSC
Noelle Olsen - Maryland Sea Grant
Paul Rago - MAFMC SSC
Rob Latour - VIMS
Rob Vincent - MIT
Sam Schiano - Maryland Sea Grant
Sarah Salois - NEFSC
Steve Cadrin - SMAST
Tim Miller - NEFSC
Thomas Swiader - NEFSC
Toni Chute - NEFSC
Tony Wood - NEFSC
Victoria Kentner - NEFSC