Spiny Dogfish/Bluefish Research Track Assessment Peer Review Meeting

Paul A H Medley
April 2023

1 Executive Summary

- Two stock assessments, for Atlantic spiny dogfish and Atlantic bluefish, were reviewed over 4 days of online meetings to determine whether the working groups had met the stock assessment Terms of Reference (ToRs).
- All ToRs were met for both Atlantic spiny dogfish and Atlantic bluefish.
- ToR 1 – 3 were met because both working groups reviewed the available data on environment, catches and abundance indices and carried out necessary evaluations and analyses making them suitable for inclusion in the stock assessments.
- ToR 4 was addressed for spiny dogfish by applying a Stock Synthesis 3 separate-sex model which had been successfully used for Pacific spiny dogfish. For bluefish, the assessment was transferred from an ASAP model to the state space model WHAM in a well thought-out step-wise process, and demonstrated significant improvements.
- For ToR 5-6, on biological reference points and projections, both assessments applied methods used in their respective software. Neither assessment had a reliable stock recruitment relationship, so suitable proxy reference points were proposed.
- ToR 7 covered research recommendations and in both cases the working groups demonstrated progress where it had been possible, as well as proposing further suitable recommendations themselves. These formed a sound basis for ongoing improvements in these stock assessments.
- I agreed with their most important recommendation for spiny dogfish, which was to complete the dorsal spine ageing work, as the lack of these data caused the greatest uncertainties and providing them would have the greatest impact on the stock assessment. Otherwise, I suspect more could be made in the model of contrasting the exploitation rates between males and females, and this might be enhanced by separation of abundance indices into sexes.
- My additional recommendations for bluefish are to explore whether a latent categorical variable might be estimated for the length data compositions to explain their bimodality and to try to capture within year variation by considering a shorter time step.
- This independent report provides my findings, conclusions and recommendations for each of the working group ToRs. There is a separate Review Panel's summary report which describes the consensus findings of the review panel. There were no significant disagreements between review panel members.

2 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 marine living resources based upon the best scientific information available. 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.

The purpose of this meeting was to provide an external peer review of the spiny dogfish and bluefish stock assessments. 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, which forms part of the Northeast Region Coordinating Council stock assessment process. This process consists of stock assessment development and report preparation carried out by Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees, an independent peer review conducted by the peer review panel, public presentations, and document publication. The results of the peer review will be incorporated into future management track assessments, which serve as the basis for developing fishery management recommendations.

I was one of three experts on the review panel were employed as part of the CIE program and charged with conducting their peer review impartially, objectively, and without conflicts of interest. The reviewers were independent of the development of the science, and without influence from any position that the agency or constituent groups may have.

3 Review Activities

Approximately two weeks before the meeting, the assessment documents and supporting materials were made available to the Peer Review Panel through a NEFSC website[1] (

[1] https://appsnefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php

Appendix 1. Bibliography). On December 3, 2022, the review panel members met with Michele Traver and Russell Brown to review and discuss the meeting agenda, reporting requirements, meeting logistics, and the overall process.

The peer review was conducted remotely from December 5 to 8, 2022. The meetings were held remotely from 9:00am to 4:00pm EST from Monday to Thursday to review the working group's progress on each of the terms of reference. The Terms of Reference are in Appendix 2, Annex 1; the agenda is in Appendix 2, Annex 2. The review panel consisted of Joe Powers, Robin Cook, Paul Medley, and Yan Jiao (chair).

On each day, Monday 5th – Thursday 8th December, members of the working group made presentations explaining how they had attempted to address the ToR, with the first two days allocated to spiny dogfish and the next two to bluefish. The presentations during the meeting for each assessment followed the TORs, allowing the review panel to gain a deeper understanding of each assessment. There was an opportunity to ask questions by the review panel and any stakeholders present at the open meeting. Some additional minor work was requested in the form of additional graphs and calculations based on the outputs. Friday December 9th was allocated to report writing and no further information was requested from the assessment working groups.

All sessions were open to the public and no technical problems of note were encountered over the course of the meeting. The meeting software (Webex) performed well. There was the odd drop out or distortion from some speakers dependent on the reliability of the internet connection, but these problems were minor and did not affect the quality of the review.

The chair prepared a summary review panel report that was edited subsequently, as necessary, by the panel members using Google Docs. Each of the CIE Panelists also drafted and submitted an independent reviewer's report to the Center for Independent Experts. This independent report provides my findings, conclusions and recommendations, and can be read independently of the Review Panel's summary report. I had no disagreements with the panel's summary report, and these findings are consistent with it.

4 Spiny Dogfish

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

ToR 1 was met by the working group. The working group had reviewed relevant ecosystem and climate influences on the stock and stock productivity. The working group characterized the uncertainty of the available information and reviewed connections to the other ToRs, primarily covering their potential effects on stock dynamics.

A major finding was the significant changes in the life history, notably a shift in length at maturity to a smaller size. The working group were unable to attribute this change to a

particular cause, but could include it in the stock assessment model as a separate time block. However, confidence in such changes in the biology of the stock would be increased if the causation was clearer, so this issue would benefit from further research.

Strong relationships with other environmental factors, affecting recruitment for example, were not found. It is usually difficult to find simple patterns relating physical oceanographic variables to population dynamics, probably because fish exhibit behavioral changes to mitigate effects and such effects may be non-linear so clear patterns are only exhibited infrequently. Even where a clear relationship is identified, this may simply push the dependence onto another (environmental) variable which itself cannot be predicted.

Seasonal patterns are obvious responses to environmental forcing, but as a regular pattern they are often ignored or seen as a complicating rather than helpful factor, as in this case. Nevertheless, seasonal movements and changes in behavior need to be understood to interpret data correctly. Qualitative information has been used to this effect in this assessment.

4.2 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.

ToR 2 was met. The working group provided an outline of all catch data and from where they were derived. This covered a number of fleets operating in USA waters as well as catches from outside reported to the North Atlantic Fisheries Organization. Some basic data were obtained from Canada, although interest in this fishery in the Canadian waters has fallen in recent times. Discard estimates were available and estimated. Uncertainty was characterized, estimated as a CV for discards, but not used in the stock assessment because of convergence issues. This is not a particularly good reason in the longer term to exclude these measures of uncertainty, and therefore solving this lack of convergence will be worth further exploration.

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

ToR 3 was met. The working group reviewed all available data as far as the panel was aware and demonstrated the value of different indices leading to their justified inclusion or exclusion from the model. In particular, the comparison between the combined VAST index and the NEFSC survey index supported NEFSC being the main abundance index in the model. The working group was correct in my opinion not to consider the inshore surveys as independent indices of abundance.

A pup index from the NEFSC spring bottom trawl survey may be considered in the future as a recruitment index. Recruitment indices may provide better catch limit estimates from short term projections.

The effect of schooling on the indices was not accounted for directly, except through general over-dispersion in observation errors. Schooling could affect patterns in length compositions, so might be worth considering as an explicit effect and I believe that this could benefit from more study when standardizing CPUE and length frequencies as well as accounting for uncertainty. For example, dependent on the resolution of the available data, it might be possible to model the contagion as a mixture of Poisson (number of schools) and the gamma distribution (catch weight from each school). A likelihood based on this (e.g., Tweedie distribution) might be tried for both the catch-per-unit-effort and length frequency data where lengths from the same school are likely to be very similar. This could lead to an improvement in the treatment of the survey and other data.

Tagging data have been rejected as an abundance index as this was not the intent of the experiment and tagging only took place in a limited area. However, it may still be informative if sufficient mixing has taken place. Excluding recaptures until a significant time has passed may lead to too few recaptures. If tagging is being undertaken to look at growth and movement, it may be worth seeing whether it might also be useful as an abundance index, particularly for smaller sharks. Tag and release programs might be promoted in the recreational fishery.

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

ToR 4 was met by the working group. The chosen modelling software, SS3, was appropriate because it can model separate sexes with different life history parameters and can fit to only length data if age data are not available. In addition, SS3 has been used successfully for the Pacific spiny dogfish stock which has similar data and life history.

The biological blocks were necessary to fit the data. There seems to have been a genuine change in growth and maturity for the population, and this could underlie additional life history changes due to climate change, fishing or a combination of both. The cause of these changes would be valuable to know, and I would encourage further exploration and research on this issue.

The problems with reading ages have led to the lack of length-at-age data which is the largest uncertainty in the model. For a model that is fitting to length data, the uncertainties in growth are critical in obtaining reliable scientific advice. A number of concerns are only likely to be resolved in the short term by obtaining reliable a length-at-age data set covering a wide range of lengths and age. This would at least pin down the current growth and maturity-at-age which would produce more reliable model results.

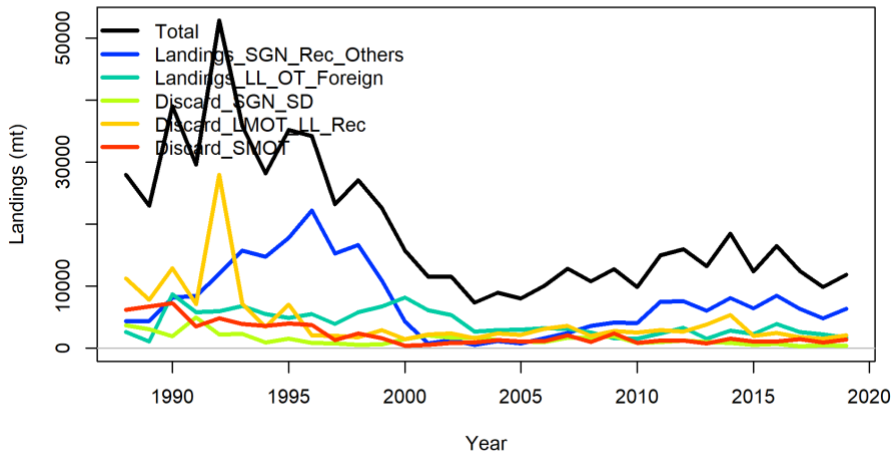
The Dirichlet multinomial is a useful likelihood extension of the multinomial and helps with accounting for overdispersion, but does not otherwise account for inter-correlated observations. There is, therefore, probably room for improvement when examining length compositions in how they are treated in the model. The lack of fit to the abundance indices is not ideal and a more balanced way to fit length, age (when available) and indices would be desirable.

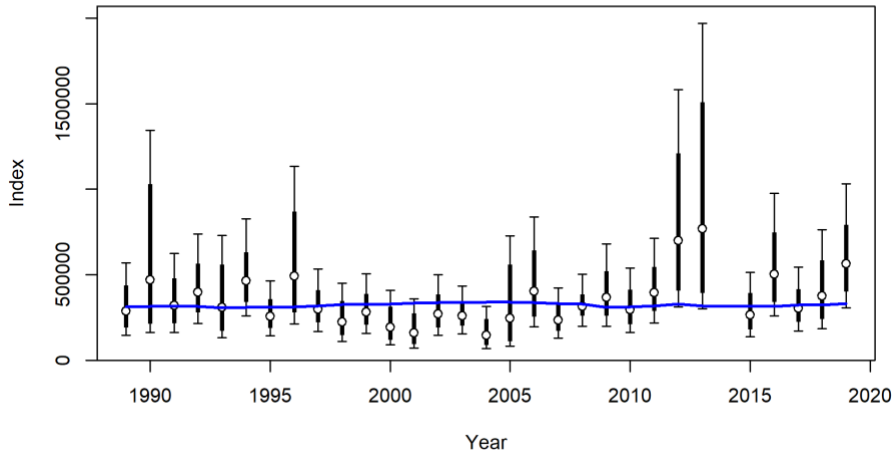
I believe the working group was correct to estimate the stock recruit (SR) function outside the model in this case, or to use a fixed steepness appropriate for the species. The aim of an integrated stock assessment should be to fit to all observations in a single model, but this is often not possible. It is therefore reasonable to estimate some model parts separately, such as standardizing abundance indices, or, as in this case, the SR function.

The density dependent pup production might well explain a SR of the types used, but it would be reasonable for there to be other density dependent effects on adults as well. There are significant differences between sharks and bony-fish population dynamics, which may need to be more explicitly accounted for in the stock assessment model. This would also have implications for reference points (ToR 5). In the longer term, perhaps in co-operation with the Pacific spiny dogfish assessment, alternative models to SS3, more explicitly accounting for viviparous birth in particular, might be considered.

In SS3, discards are treated as a separate “fleet” for model fitting purposes. It was pointed out in the meeting that this is not how discarding works. In my experience, trying to fit discarding as a process applied to the catch is problematic because parameters are more likely to be heavily correlated, making it difficult to estimate them and decreasing the chance of successful model convergence. I was not involved in decisions to use this fleet approach in modelling discards in SS3, but I would imagine for a general modelling software, to increase robustness, something like this approach, where selectivity is estimated independently for different length and age compositions, would be necessary. In practice, it makes little difference to estimation as long as the data are well fitted and appropriate links between landings and discards are maintained in projections, reference points and other uses of the results.

In general, there is weak correspondence between the various data sources. The decrease in catch in the 1990s should have seen an increase in population size and in mean length after an appropriate delay. There is evidence for an initial increase in mean length, but the survey abundance index remains relatively flat and appears to be almost tracking the catches rather than responding to them (Figure 1). It is unlikely that the model will be able to explain these patterns well without more help. One option might be to produce separate survey indices for each sex which would allow the model to contrast the different fishing mortality applied between males and females (see ToR 7). This would help the model attribute the changes in observations between adjustments in fishing mortality and recruitment.





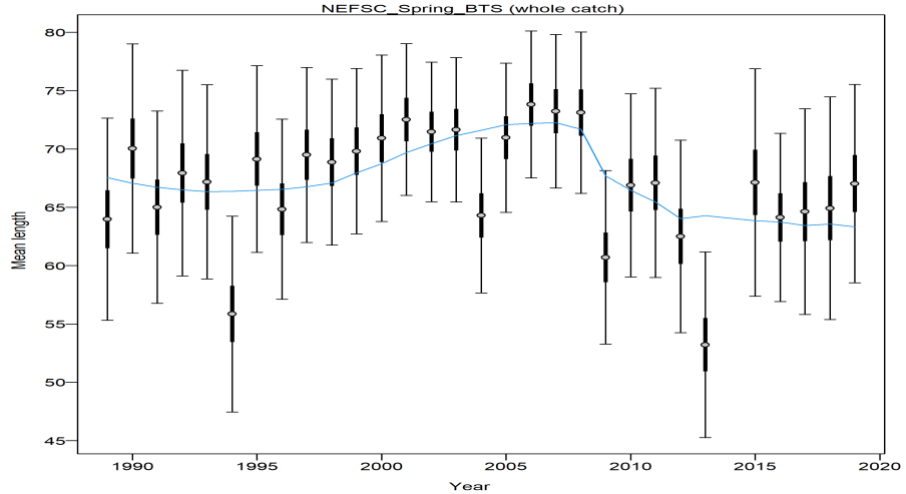


Figure 1 Top: Annual spiny dogfish catch. Middle: Observed and model-predicted abundance index (1,000s) for the NEFSC spring bottom trawl survey. Lines indicate 95% uncertainty interval around index values based on the model assumption of lognormal error. Thicker lines indicate input uncertainty before addition of estimated additional uncertainty parameter. (Figure 4.13 from main report). Bottom: Mean spiny dogfish length.

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

This ToR was met. Estimates of the reference points were provided and the basis for these examined. Analytical model-based estimates were available but based on a stock-recruitment (SR) function that was questionable because the fitted SR functions were highly uncertain.

As was suggested by the review panel, it may be better in these circumstances to justify a fixed steepness for the Beverton-Holt SR and use that as the basis to define reference points. Some standard, agreed-upon precautionary method to define these reference points, such as precautionary fixed steepness, may be a better approach than attempting fitting SR parameters when these will be highly uncertain. A target of SPR60% is also reasonable for this species.

It was suggested by the working group that a primary density dependent reproduction effect may manifest itself in reproduction frequency and litter size. These ideas might be tested by observing female reproductive rates, so it may be possible, but difficult, to detect this if pregnancy rates can be detected in catches.

The working group defined MSY in terms of the retained catch and did not include discards. This was discussed on the basis that MSY should be an attribute of the stock rather than the fishery, so discards should be retained in the calculation. However, it is difficult to separate the two because the calculation will always include the fishery selectivity and the idea of exploitable biomass. This perhaps should be addressed more formally based on the management goals for the fishery, but I would tend to support the working group decision to consider the retained catch as the yield calculation in MSY as it is generally the management objective.

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

This ToR was met. The projections used the most recent biological and selectivity blocks from the stock assessment to project forward 3 years under the different scenarios (F=0, SPR50%, SPR60% and SPR70%). The projections employed the SS3 standard method which was appropriate given this was used for the assessment. These short term projections should be reasonably reliable as they do not depend too heavily on projected recruitment.

All projections are expected to produce an increasing population size, the stock recovering with larger year classes expected to arrive into the fishery in the next few years. The NEFSC

spring survey abundance for the 25-30cm pups is critical for short term projections, and it may be sensitive to survey timing.

Discarding in projections should probably be handled as a fixed ratio with landings fishing mortality. This is handled in SS3 as part of the mixture of fleets. For short term projections, maintaining the proportional mortality by fleet is appropriate. For longer term projections, the proportion of discards might well change, which would have significant implications for the fishing mortality.

It was not clear that exactly the same assumptions with respect to the SR had been made in fitting the stock assessment, reference point estimates and the projections. For the management track assessment, it will be important to ensure that all calculations are internally consistent, so that the same calculations and reference points are applied to the projections (ToR 5). Again, in the absence of a demonstrably better and more reliable model, it might be more robust to use a standard fixed steepness (say, 0.6 or whatever the working group decides) with a Beverton-Holt SR in fitting the stock assessment, reference point estimates and projections.

4.7 ToR 7: Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 1[2] could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

This ToR was met. The working group reported on their review and evaluation of all previous research recommendations, as well as identifying new research recommendations. Research recommendations were prioritized, and I agreed with the working group that the most critical research would be on the ageing. Completing the extensive work that had already been done on this would have the greatest impact on the current stock assessment and significantly reduce uncertainty. Other work related to ToR 2, notably seasonal movement in response to temperature and the effect of depth, was reviewed, although a clear way forward on environmental effects is less clear at this stage.

It was noticeable that one of the species covered in the VAST forage index developed for bluefish was spiny dogfish (Ng et al. 2021). This information was not used by the working group because they believed that the spiny dogfish diet was too variable and affected by size. Nevertheless, based on the work carried out on the bluefish, it may be worth re-examining this issue if the bluefish forage index proves to be useful in estimating catchability or selectivity for the abundance index. Adapting the bluefish methodology should be straightforward and it is possible this might improve the index fit in the SS3 model.

[2] This was changed from "ToR 2" in the original text as it clearly is referring to ToR 1 and therefore was considered a typing error.

Given the potential useful contrast between the levels of exploitation applied to males and females and the ability to track them separately, I believe it would be potentially useful to completely separate the survey index between males and females. This would help the model fit to different signals, which may be lost when just contrasting sex ratios. This could be done by separating the data into two indices and setting the catchability to zero in the model for each sex as appropriate.

Previous recommendations included developing a length-based and/or state space model which I generally support. However, the SS3 model is the last that was used for the Pacific spiny dogfish stock, so there is some benefit with this approach. In addition, the primary research recommendation is to carry out more ageing using the 2nd dorsal spine, which will make an age-based model more appropriate depending on how much ageing is done. An age-based model may excessively smooth out length information, but a length-based model may similarly ignore age information in the data. For example, the length distribution of a cohort resets each year if using a growth model, whereas with length-based selectivity, the length structure of the cohort will change because lengths will be subject to different mortality rates within the same cohort. So ideally, both length and age are tracked in the population model.

I would therefore suggest exploring the use of the Gadget model (Begley and Howell 2004[3]) or similar, or develop a bespoke model following typical shark life history, to see whether it would be of use in this context. In contrast to the Woods Hole Assessment Model, the Gadget software is more focused on model structure rather than statistics, so it is designed to account potentially for multiple areas and ecosystem effects as well as more detailed stock structure at the cost of less statistical sophistication. Nevertheless, it might be useful at the very least in trying to determine what might be important in terms of model structure when using the SS3 model. Gadget is reportedly available as a R-package from CRAN "gadget2" (and "gadget3" is also now available). However, I cannot whole-heartedly recommend it as although I have reviewed assessments using it, I have not used it myself.

4.8 ToR 8: Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment. A backup assessment approach is required to be in place as a hedge against a scenario where the primary catch-at-age model is not suitable for providing management advice.

This ToR has been met. A backup assessment approach was identified if the main assessment was rejected. However, I believe that the alternative is not a better alternative than the SS3 model, so the SS3 model should be used as the primary model for management advice. It is worth noting that the selected alternative approach (the "Stochastic Estimator") would use the survey data exclusively, which was not fitted well in the SS3 model. Switching to a data limited approach would require a strong argument as to the basis the excluded data (in this case length compositions) were rejected.

[3] Building models using the Gadget framework (gadget-framework.github.io)

However, having this ToR as part of the review process is a good step forward in my opinion. Previously a stock assessment might be rejected with no clear plan of what would be done instead. This additional ToR makes it clear what the best alternative would be and should help inform the review panel when making a decision to reject a management track assessment, what is the alternative model they are rejecting it for.

5 Atlantic Bluefish

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

ToR 1 was clearly met. A wide range of environmental variables was considered, including physical factors (such as movement and distribution related to temperature, depth and coastline), and biological factors (such as life history variability, prey and predation). The clear systematic approach used by the Ecosystem & Socioeconomic Profile gave confidence that all relevant and most possible influences had been reviewed, although conclusions were limited in many cases from a lack of data. The works completed under this ToR were used in other ToRs, particularly the survey indices (ToR 3), life history updates to the model (ToR 4) and further research recommendations (ToR 7).

Although the evaluation of ecosystem and climate influences was rigorous, a few areas have not been evaluated yet. It may be worth including the consideration of disease and parasitism (e.g., the nematode ovary parasite *Philometra saltatrix*) as well as predation because these might exhibit density dependent mortality which could have a substantial impact on the population dynamics and models. Early life history might also bear some evaluation using plankton surveys. While information may be lacking in these areas, marking them for future research development would be important.

Environmental covariate effects on the population or observations are best linked to a specific cause, rather than relying on correlation. In trying to estimate catchability changes, there is a clear attempt to address this.

5.2 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.

ToR 2 was met by the working group. Catches from all sources, landings and discards, had been compiled to the best of my knowledge. A full description of the spatial and temporal distribution of catches was given, including attempts to compile the older historical catch. Uncertainty in these data sources was fully described. In particular, significant improvements in data sources was demonstrated.

The recreational fishery makes up the largest proportion of the catch with significant discards (catch and release). The literature review indicated a significantly lower release mortality than used previously. I also support including commercial discards which were previously

treated as negligible. Including these small factors allows the model objectively to decide how negligible they are.

The working group was able to demonstrate clear improvements in the recreational (MRIP) data, although there were limited length data for recreational discards. These discard length data will be difficult to collect for obvious reasons but were not a critical problem for the assessment. The new methods seem more reliable for the current fishery and running the new survey system alongside the old allowed a reliable transfer between the systems. Recreational CV for annual catch was around 9% which is very good for this type of survey method. The fisheries are seasonal, with North-South seasonal migration detectable in CPUE by state.

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

ToR 3 was met. A very wide range of indices were considered and there were a number of significant improvements in the indices used in the final stock assessment. Improvements include a Bayesian hierarchical model that was used to combine YoY indices across different local surveys, the MRIP CPUE index using new survey data, and fish-guild based trip selection to standardize effort including null trips. The multinomial model to construct ALKs was in theory an improved method compared to the borrowing algorithm used previously. This allowed consistent application across all age-length data.

While inclusion of the forage index (ToR 1) as a catchability covariate in WHAM is a reasonable aim, and a strength of WHAM is that this sort of covariate can be included, the forage index could bear further examination outside WHAM to see how prey availability might be affecting fishery dependent and independent indices and selectivity. For example, including the forage index in a CPUE standardization model might provide a clearer indication of how it is related to the different abundance indices.

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

ToR 4 was met by the working group. The procedure to move the assessment from the older ASAP model to the state-space model WHAM was thorough and very clear. The process demonstrated the advantages of moving the assessment to a state-space model.

The use of multinomial estimates for the age-length keys made a significant difference to the assessment, particularly in the early part of the series. The multinomial model produced smoother estimates than the borrowing method applied previously and resulted in improved model convergence when fitting. Overall, the multinomial ALK model provided a better fit in my opinion and provided more realistic estimates of ALK probabilities. The final stock assessment outputs were not changed significantly.

The model selectivity was age-based. Length based selectivity may improve the model fit and should be considered for future models as an alternative. The bimodality in length compositions was only partially explained by the bimodal age-based selectivity and selectivity appeared to also have changed over time. Selectivity blocks reduced but did not eliminate the problem. It is possible that the length frequency data result from a mixture of selectivities that may be better modelled through explicit separation (see ToR 7).

Within-year dynamics is also a significant contributor to uncertainty for the model, particularly for younger ages. Growth is relatively fast in the first year, so averaging over the whole year will lose quite a lot of information on recruitment that may be present in the length data. A shorter time step could better account for any seasonal effects, such as changes in catchability based on seasonal movement, and early growth.

I am not sure that all the indices presented should be included in the final model, although the working group demonstrated results were insensitive to index inclusion and exclusion (except the PSIGN index, which was important as an indicator of older fish abundance). In general, indices that do not cover the stock area and are considered suspect, such as the "Bigelow" index, should be considered for removal and only used in sensitivity runs to see if they make any difference. If indices show different trends, they cannot all be tracking abundance of the stock, so including them in base runs does not make much sense. At worst, they should be included in separate runs which are presented as equally likely. However, combining indices that may be locally informative in an additive index like the YoY index or the VAST approach appears to be the best way to include composite CPUE and survey information.

The sensitivity run including the forage index (BF28W_m7ecov) is still considered to be in development, but on the face of it, leads to general fit improvement, particularly to the MRIP index (Figure 2). Exactly how the index is affected by forage availability is not necessarily clear, but presumably this measures how much the stock and fishery overlap when taking into account fish movement. It could be that similar evaluations might explain bimodal selectivity observed in much of the length composition data. The changing catchability over time also provides an explanation for retrospective patterns, so improvements in this regard are to be welcomed.

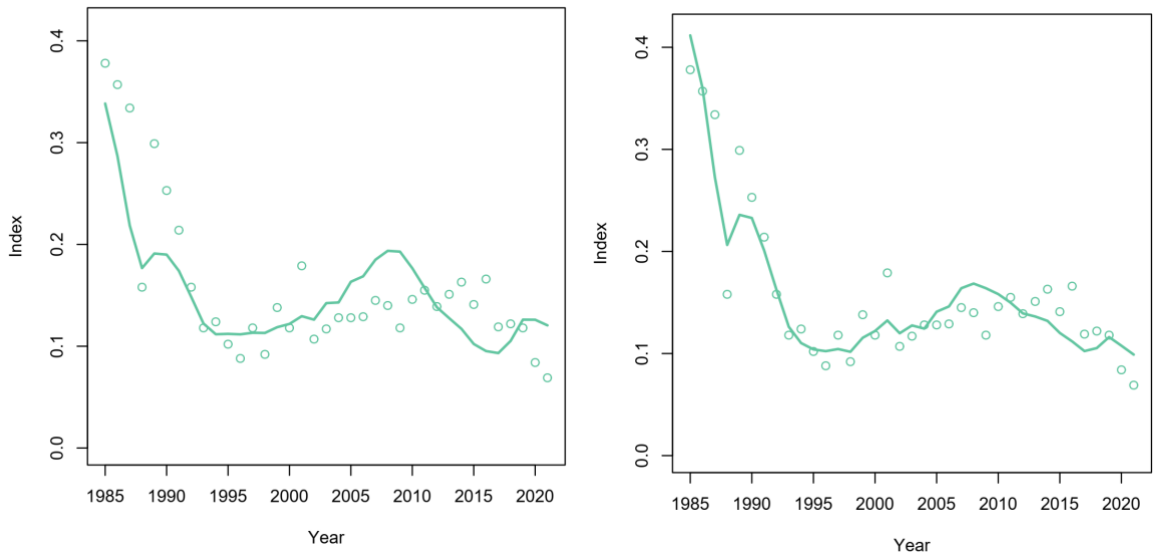


Figure 2 Model fit to the MRIP index for the base model (right) and including the forage index (left).

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

This ToR was met. The status criteria were estimated appropriately for the new model, with confidence intervals where appropriate. The use of a target SPR35% for deriving catch limits was well justified and I believe is robust. Using the last five-year averages for selectivity and weights at age for calculations seemed reasonable. Estimates of the stock status were provided and the Kobe plot, in particular, provided a good visualization of results. Further exploration of SR function and MSY reference points is required, but the proposed reference points provide a good foundation for management advice.

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

This ToR was met. Three-year projections are carried out in WHAM using 5-year average fitted selectivity, and weights at age, with other parameters fitted or used in the assessment model. Sensitivities were also conducted on 3 and 10-year averages, and in my opinion the 5-year average appeared most appropriate. This was the same methodology as applied for reference points (ToR 5). For short term projections, the method applied is reliable, does not heavily depend on the SR function for example, and can account for observation and process errors appropriately. The new model should in theory provide more reliable short term

projection estimates, including estimates of uncertainty, than the previous maximum likelihood model because it includes autoregressive terms among other things.

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

This ToR was met by the working group. The working group identified and reviewed all previous research recommendations and had taken action to resolve many. Some, like revisiting scale ageing, were not done because the original scales could not be found. Some, like extending the SEAMAP age structure to age-1 and importantly, improving the MRIP index using fish-guild associations to define trips, were completed. New recommendations were provided, primarily being further data collection (rather than research as such), which will help better inform the stock assessment model. The research recommendations were prioritized.

The most promising research line, related to ToR 1, was the forage index. I agree that this has the potential to improve the stock assessment fit to survey indices and in general provide a better explanation for observations leading to a more reliable stock assessment. This should remain a high priority research recommendation.

In addition to the BWG recommendations, I believe the following, in order of priority, may be worth considering:

1. The bimodal length frequencies may be better explained in the model with different data treatment. It is not clear how these modes are formed, but, dependent on the data resolution, it may be possible to disaggregate them into separate fleets each with their own selectivity. I have done some preliminary work on this for tuna which looked promising (Medley et al. 2021). The idea is to estimate an unobserved latent categorical variable that can be linked to separate selectivities. Medley et al. (2021) used a simple method to cluster similar length frequency samples, but more sophisticated methods could be proposed, and the method might also be linked to the forage index and environmental covariates. If the modes cannot be separated by clustering samples, then this approach will not help.
2. The current annual time step may be missing contrasts in the data which could help the model fit. There are suspected within-year changes to the population in distribution and life history which may be better explained using a shorter half-year or quarter-year time step. This can help, particularly if the model is able to use within year depletion effects to improve estimates of F or explain changes in catchability and selectivity. Whether this works will also depend on the data being allocatable to the relevant periods and the within-year changes being useful. In some cases, within-year changes need to be explained by nuisance parameters, which offer no improvement

to the assessment. Therefore, in the first instance, whether this is worth pursuing should be explored.

3. Length-based selectivity may provide a better fit for the length compositions.

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

This TOR was met. As for spiny dogfish, an abundance index method was chosen as most appropriate. Given the model fitted the abundance indices reasonably well, this could be justified as a strong indicator for abundance. However, to favor this method as opposed to the integrated stock assessment, some argument is required that the catch, length and/or age data are inappropriate. On the whole, removing the suspect data from the stock assessment is preferred wherever possible, rather than resorting to data-limited methods, in my opinion. As the WHAM stock assessment was accepted, this approach should not be required.

6 General Conclusions

All ToRs were met for both assessments. The proposed models should be used as the basis for stock assessments, and the BRPs and short-term projections can be used for management advice.

Like all stock assessments, they should continue to improve wherever possible, but both are sufficiently well developed to provide good management advice in terms of sustainable catch and reasonably reliable short-term projections.

The ToRs were clearly written and so made it easier to review the stock assessments and apply less subjectivity. Some of the ToRs were very general, so they are relatively easy to meet. ToRs requiring the working group to review available information or to characterize uncertainty were met if these issues were addressed. For example, ToR 1 simply required that environmental data and potential effects be considered and evaluated. The level that this evaluation was taken to was quite different between the spiny dogfish and bluefish assessments, with bluefish able to evaluate environmental effects much more rigorously. However, I am not sure anything might be done to improve that as it is unclear what is possible before the working group can attempt the task or whether equal priority is required between stocks. The level of detail reflects the available data collection and research priorities.

7 References

Begley, J, and D Howell. 2004. An overview of Gadget, the globally applicable area-disaggregated general ecosystem toolbox. ICES CM, 1–5. http://www.hafro.is/gadget/.

Medley P, Defaux V and Huntington T 2021. Exploratory Analysis of Tropical Tuna Longline Selectivity and its Implications for Stock Assessment. IOTC-2021-WPTT23(DP)-11_Rev1.

Ng, E. L., J. J. Deroba, T. E. Essington, A. Grüss, B. E. Smith, and J. T. Thorson. 2021. Predator stomach contents can provide accurate indices of prey biomass. ICES Journal of Marine Science 78(3):1146–1159I.

Appendix 1. Bibliography

Documentation for the meeting was provided via the NEFSC Data Portal[4].

2022 Spiny Dogfish Research Track Assessment

Last edited: 23 November 2022

This document serves as a broad Table of Contents of the materials on the data portal.

The main assessment document (Spiny_Dogfish_SAW_SARC_2022_FINAL.pdf) addresses the Terms of Reference, with additional details, supplementary work, or main text abbreviated for readership ease are provided in working papers. Model details for the base case model are also provided.

Main Assessment Report

• "Spiny_Dogfish_SAW_SARC_2022_FINAL.pdf"

Background

"Read Me.pdf" – Document and materials guide, as well as a repository of any report revisions.

"plots_v3.6.2_1.5_fnum_a12.zip" - This zip file contains the base case model figures and files produced from SS3. Within this zip file, there is a file labeled '_SS_output.html', which is an html that allows for viewing SS3 produced plots and results in an organized fashion (i.e. by various data type or model result).

Working Papers

• "Anstead 2022a.pdf" - Anstead K. 2022a. Natural mortality estimates for spiny dogfish.

• "Anstead 2022b.pdf" - Anstead K. 2022b. Two data poor methods applied to spiny dogfish.

• "Chang et al. 2022.pdf" - Chang J-H, Hart D and McManus MC. 2022. Stock synthesis for Atlantic spiny dogfish.

• "Hansell and McManus 2022.pdf" - Hansell A and McManus C. 2022. Spatio-temporal dynamics of spiny dogfish (Squalus acanthias) in US waters of the northwest Atlantic.

• "Hart and Chang 2022.pdf" - Hart DR, and Chang J-H. 2022. Per recruit modeling and reference points for spiny dogfish.

• "Hart and Sosebee 2022.pdf" - Hart DR, and Sosebee K. 2022. Length/Weight/Fecundity relationships for Atlantic spiny dogfish.

• "Jones 2022.pdf" - Jones AW. 2022. Exploring vessel trip report and observer based fishery information for spiny dogfish.

• "Jones et al. 2022.pdf" - Jones AW, Didden JT, McManus MC, and Mercer AJ. 2022. Exploring commercial CPUE indices for the spiny dogfish in the northeast U.S.

• "McCandless 2022.pdf" - McCandless C. 2022. Preliminary spiny dogfish movements and growth estimates from NEFSC mark recapture data.

• "McManus et al. 2022.pdf" - McManus MC, Sosebee K, and Rago P. 2022. Biological Reference Points for Spiny Dogfish: Revisiting Rago and Sosebee (2010).

[4] Northeast Stock Assessment Workshop :: SASINF Options (noaa.gov) (Assessment Year: 2022, Species Name: Atlantic Bluefish / Spiny Dogfish, Stock Area: UNIT, Review Type: Research Track, Information Type: Select All)

• “Neiland and McElroy 2022.pdf” - Neiland JL and McElroy WD. 2022. NEFSC Gulf of Maine Bottom Longline Survey Data and Analyses for Spiny Dogfish

• “Passerotti and McCandless 2022.pdf” - Passerotti MS, and McCandless CT. 2022. Updated age and growth estimates for spiny dogfish Squalus acanthias.

• “Sosebee 2022a.pdf” - Sosebee KA. 2022a. Maturity of spiny dogfish in US waters from 1998-2021.

• “Sosebee 2022b.pdf” - Sosebee KA. 2022b. Spiny dogfish catch summary and derivation of catch at length and sex.

2022 Bluefish Research Track Assessment

Last edited: 30 November 2022

This document serves as a broad Table of Contents of the materials on the data portal.

The assessment document (Bluefish_SAW_SARC_2022_FINAL.pdf) includes all information to address the Terms of Reference, but we include additional detail in a series of working papers itemized below.

The working papers are available on the NEFSC data portal and on our GitHub site.

“Assessment Report”

• Bluefish_SAW_SARC_2022_FINAL.pdf = Main assessment document

“Background”

• 01 readme.docx – document guide and repository of any report revisions.

• Background_documents.zip -> a small collection of grey and peer reviewed literature that the working group thought was especially helpful to accompany the assessment. Includes:

  o The 2015 bluefish peer review report

  o Stock and Miller 2021 -> describes the Woods Hole Assessment Model (WHAM) framework and software package

  o Technical documentation for ASAP 3 -> technical documentation related to ASAP software used as part of early model exploration

  o User manual for ASAP 3 -> technical documentation related to ASAP software used as part of early model exploration

  o Ng et al. 2021 -> used as a guide for early development of the forage fish index

  o Thorson 2019 -> used to help guide VAST modelling

• WP 01 Tyrell et al. 2022. Bluefish Ecosystem and Socioeconomic Profile.

• WP 02 Valenti 2022a. The Spatial Distribution of Bluefish (Pomatomus saltatrix): Insights from American Littoral Society Fish Tagging Data

• WP 03 Tyrell 2022. Bluefish VAST Index Exploration.

• WP 04 Gaichas et al. 2022. Vector Autoregressive Spatio-Temporal (VAST) modeling of piscivore stomach contents, 1985-2021.

• WP 05 Truesdell et al. 2022. Life History Analyses for Bluefish.

• WP 06 Tyrell and Truesdell 2022. Natural mortality of bluefish.

• WP 07 Celestino et al. 2022a. Index of abundance exploration and development by the Bluefish Working Group’s Fishery Independent Data Group.

• WP 08 Wood 2022a. TOR 2: Commercial and Recreational Data Collection and Analysis.

• WP 09 Drew 2022a. Recreational Data Changes for Bluefish, 2012-2021.

• WP 10 Drew 2022b. The Spatial Distribution of Bluefish (Pomatomus saltatrix): Insights from MRIP Data.

• WP 11 Valenti 2022b. Catch-and-Release Recreational Angling Mortality of Bluefish (Pomatomus saltatrix): Updated Analysis for 2022

• WP 12 Drew 2022c. Development of the Composite YOY Index for Bluefish.

• WP 13 Drew 2022d. A Fishery-dependent CPUE index for bluefish derived from MRIP data.

• WP 14 Celestino et al. 2022b. Development of Bluefish Age-Length Keys.

• WP 15 Wood 2022b. Bluefish Model Bridge-Building in ASAP.

• WP 16 Wood 2022c. ASAP diagnostic plots.

• WP 17 Wood 2022d. WHAM diagnostic plots.

• WP 18 Truesdell 2022. Alternative assessment plan.

• WP 01: A reference was missing from the ESP WP: Conover, D.O., Gilmore, T. and Munch, S.B., 2003. Estimating the relative contribution of spring-and summer-spawned cohorts to the Atlantic coast bluefish stock. Transactions of the American Fisheries Society, 132(6): 1117-1124.

• WP 05: Section 3.2 in the life history working paper, the sentence: The “Midyear model” (that used all available data and where sample timing by wave was included and the average maturity for ages 3 and 4 was used); Should read: The “Midyear model” (that used all available data and where sample timing by wave was included and the average maturity for waves 3 and 4 was used)

• WP 08: Section 1.1.2, page 2 of WP08, “During this time commercial landings have been consistently lower than the recreational catch and accounted for on average ~XX% of the total catch on (Figure X).” -> should have indicated: 14% and Figure 51 o “Six statistical areas, however, collectively accounted for more than 75 % of VTR reported landings in 2001, with individual areas contributing 6% to 18% of the total (Table X).” -> should have indicated Table 3. Section 1.1.5, page 4 of WP08, “This pattern in bluefish length frequency has been observed to a lesser degree in some years of the recreational landings length frequencies (Figure 25), and the recreational discard length frequencies (Figure X).” -> should have indicated Figure 50

Appendix 2. Individual Independent Peer Reviewer Report Requirements

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

Bluefish and Spiny Dogfish Research Track Peer Review
December 5-9, 2022

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[5].

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

The purpose of this meeting will be to provide an external peer review of the spiny dogfish and bluefish stocks. The requirements for the peer review follow. This Performance Work

[5] https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

Statement (PWS) also includes: **Annex 1**: TORs for the research track, which are the responsibility of the analysts; **Annex 2:** a draft meeting agenda; **Annex 3:** Individual Independent Review Report Requirements; and **Annex 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 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 and elasmobranchs 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 spiny dogfish and bluefish 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 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. Again, the CIE reviewers are not required to reach a consensus. The chair will take the lead in editing and completing this report. The chair may express their opinion on each research track Term of Reference, either as part of the group opinion, or as a separate minority opinion. The Peer Reviewer Summary Report will not be submitted, reviewed, or approved by the Contractor.

Foreign National Security Clearance

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

Place of Performance
The place of performance shall be hybrid, both at the contractor's facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts, and via WebEx video conferencing. **CIE reviewers may attend virtually dependent on conditions of the COVID 19 pandemic**

Period of Performance
The period of performance shall be from the time of award through February, 2023. Each reviewer's duties shall not exceed **14** days to complete all required tasks.

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

Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
December 5-9, 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
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $15,000.00.

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

Annex 1. Generic Research Track Terms of Reference

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under the TORs.
2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.
3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.
4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.
5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.
6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.
7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.
8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.
9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

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.

Annex 2. Draft Review Meeting Agenda

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

Spiny Dogfish and Bluefish Track Assessment Peer Review Meeting

December 5-9, 2022

WebEx link: TBD

DRAFT AGENDA* (v. 6/21/2022)

**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, December 5, 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 Panel Chair	
9:30 a.m. - 10:30 a.m.	TOR #1		Spiny dogfish
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2		
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		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #4		
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, December 6, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Spiny dogfish cont.
9:15 a.m. - 10:30 a.m.	TOR #5		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #6		
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 #7		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #8-9		
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, December 7, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Bluefish
9:15 a.m. - 10:30 a.m.	TOR #1		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #2		
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		

3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #4		
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, December 8, 2022

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	Bluefish
9:15 a.m. - 10:30 a.m.	TOR #5		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #6		
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 #7		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #8-9		
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		

Friday, December 9, 2022

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

Annex 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.

Annex 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. It should also include whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.)

   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.

4. 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.