SEDAR 65 Highly Migratory Species
Atlantic Blacktip Shark Assessment Independent Peer Review

conducted for the
Center of Independent Experts (CIE)
by
Joseph E. Powers

Encompassing Webinars Oct 29, 30, Nov 2, 4, 5, 2020,
Preparation and Reporting

January 2021

Executive Summary

A Review Workshop (RW) was held via webinar from 12 – 5 pm on October 29, 30 and November 2, 4 and 5 of 2020 in order to evaluate the assessment and status of Atlantic blacktip shark resources off of the east coast of the United States. This report provides the author's findings and recommendations pertaining to those tasks.

The stock assessment was based on an application of the Stock Synthesis 3 modeling platform which was adapted to the data constraints and life history requirements of blacktip sharks. The model development was aligned with normal assessment practices with some unique applications due to blacktip life history.

Key findings were: The stock assessment constitutes the best scientific information available; the stock is not overfished because the current (2018) Spawning Stock Fecundity is 134% of the Minimum Stock Size Threshold; and, the stock is not undergoing overfishing, as the current F_{2018} is 50.9% of F_{MSY}.

A suite of projections and their uncertainties were presented indicating the likely outcomes of status criteria (stock size and fishing mortality rate relative to Maximum Sustainable Yield). These projections indicate that status quo catches are likely to increase Spawning Stock Fecundity over the next few years due to current recruitment and to maintain Spawning Stock Fecundity above MSY-related levels in the longer term

While this blacktip assessment is "data-rich" by typical shark assessment standards, there were still a number of issues arising from limited data. Improvements are needed on basic life history information including natural mortality rates, stock-recruitment processes and spatial distribution and migration. Additional improvements might be achieved in understanding historical catch levels and improved length compositions. Also, future assessment improvements might arise from further methods to characterize uncertainty.

Background

A Review of the stock assessment for Atlantic Blacktip Shark was conducted under the auspices of SEDAR 65 (**S**outh**E**ast **D**ata and **A**ssessment **R**eview). SEDAR 65 proceeded in three stages: Data, Assessment and Review. The Data process was held via a series of webinars held April 2019 – September 2019 and a workshop held October 29 – November 1, 2019 in Charleston, SC. The SEDAR 65 Assessment Process was conducted through a series of webinars held from February 2020 through July 2020. The Review Workshop (RW) was held via webinar from 12 – 5 pm on October 29, 30 and November 2, 4 and 5 of 2020. The Report herein is a result of the Review Workshop. The Assessment and Data Reports and other documents (Appendix 1) were available for the Review Panel sufficiently prior to the Review webinars.

The stock assessment addresses Atlantic Blacktip sharks (or more clearly the stock of blacktip sharks off of the US Atlantic coast extending from Florida through at least New York, although the distribution is centered off of the Southeast United States). Catch time series in the assessment were available through 2018. The assessment analyses integrated catch, index, size and life history data and estimated status and management quantities specified in the HMS Fishery Management Plan. This review examines those (and further) analyses to provide peer review evaluation to further support the management process.

Description of the Individual Reviewers' Roles in the Review Activities

This review evaluates the stock assessment with the goal of providing advice for interpreting those results in determination of status and in the use of that advice by managers. More specifically, the review examined the assessment in terms of nine Terms of Reference (TORs) listed below, the ninth of which required a Peer Review Summary prepared by the Review Panel.

The Review Panel consisted of a Panel Chair and three CIE reviewers (Appendix 3). Additionally, the Panel was assisted by the lead assessment scientists who provided explanations of the assessment analyses, results of further analyses requested by the Panel and interpretations of the input data. Also, observers and the public participated in the webinars (Appendix 3) and were given the opportunity for comment and to answer questions posed by the Panel. All Panel members interacted with the lead assessment scientists through discussions and analyses results to provide responses for each TOR.

Drafting of sections of the Peer Review Summary were assigned to individual members of the Panel according to interest/expertise/need. The final Peer Review Summary was jointly edited by all Panel members for submission to SEDAR.

Finally, the CIE members of the Panel are required to individually report to CIE on their findings and conclusions according to their Performance Work Statement (Appendix 2). This report represents that task.

Summary of Findings for Each Term of Reference (TOR)

1) *Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:*

   Are data decisions made by the DW and AP sound and robust?

   Are data uncertainties acknowledged, reported, and within normal or expected levels?

   Are data applied properly within the assessment model?

   Are input data series reliable and sufficient to support the assessment approach and findings?

The original assessments of sharks under HMS jurisdiction suffered from many data deficiencies, largely due to the lack of species identifications in the catch. For that reason, the original assessment evaluated the status of the aggregate of large coastal sharks, of which Atlantic blacktip was one. As the science and management evolved, data started to become disaggregated enough to address blacktip separately (SEDAR 11). At that time (some 14 years ago) the assessment was still limited by the need for catch series reconstructions and the limited size frequency data. Thus, the assessment models used aggregate biomass methods (Age Structured Production Model and Bayesian surplus production models).

With the ensuing time period since SEDAR 11 (and the accumulation of catch, size, index and life history data since then) SEDAR 65 provided the opportunity to progress to Stock Synthesis 3. In this blacktip application this required a comprehensive set of input data, sex-disaggregated, three commercial fleets, a recreational fleet and ten indices of abundance. The available length-frequency data to support the model still appears to be quite limited, especially during the early years of the time series. Uncertainties arise about assumptions made to reconstruct the commercial fishery catches between 1981 and 1990 and whole weight to landed weight conversions. However, the sensitivity analyses were used to explore ramifications.

Recognizing this, the Data and Assessment workshops used due diligence to address the data choices and their ramifications (well documented in the working papers available to the assessment team) and were carefully considered by assessment team members. This gives confidence to the credibility and robustness of the conclusions of the assessment team and a better foundation of data to build the assessment compared with SEDAR 11. Overall, **the data decisions appear to be sound and robust and based on the best available information**.

Given the data limitations discussed above and, importantly, the acknowledgement and investigation of those limitations, ***data uncertainties were acknowledged, reported, and within normal or expected levels***, especially for typical coastal shark assessments, noting the available fishery independent indices, relatively complete life history information and gear-specific information concerning post release mortality.

Given these data and their evaluation in the Data and Assessment phases, **the data are applied properly within the assessment model** and the ***input data series sufficiently reliable to support the assessment approach and findings.***

2) *Evaluate and discuss the strengths and weaknesses of the method(s) used to assess the stock, taking into account the available data, and considering the following:*

Are methods scientifically sound and robust?

Are assessment models configured properly and consistent with standard practices?

Are the methods appropriate for the available data?

The model presented by the assessment team for HMS Atlantic Blacktip Shark is the Stock Synthesis 3 assessment model, which is applied ubiquitously in the US and worldwide. Stock Synthesis is one of the most general and detailed assessment models, which is an advantage because it is applicable in many different scenarios and is able to accommodate many different types of observations. Stock Synthesis is capable of including data in its original format.

This assessment configuration has the following characteristics: yearly catches in weight/numbers from four fleets are assumed known without error; indices of abundance from 10 fleets are assumed log-normally distributed with externally estimated CV's; length compositions are assumed multinomially distributed with effective sample sizes; parametric selection curves are estimated if sufficiency length composition data are available, otherwise the selectivity is mirrored from an assumed similar fleet; age-structured, sex-specific Von Bertalanffy growth; a common length-weight relationship and Beverton-Holt stock-recruitment with a specified steepness.

The catches given in weight are included as mass, the recreational catches given in numbers are included as numbers, and the length compositions are included where available. One detail is that the length compositions are included as multinomial, which implicitly assume that the compositions from a fleet within a year are negatively correlated, but the data most often show that such observations are positively correlated across neighboring length groups. This could affect the estimated uncertainties.

There are many possible ways to set up and configure Stock Synthesis which implies alternative assumptions both big and small. One way to address this uncertainty is to conduct the assessment simpler models than Stock Synthesis or simpler renditions of Stock Synthesis as "structural" sensitivity analyses. Such an analysis had been completed by the assessment team in a previous assessment of sandbar shark and found that Stock Synthesis could be configured to be very similar to the ASPM.

In any model configuration process, there is always the tradeoff between adding structural detail at the expense of neglecting sources of variance. This assessment is no different. In order to obtain meaningful model fits, some parameters are fixed, prior distributions are provided on other parameters, variances or effective sample sizes are assumed, and some index data were smoothed across years. While this was necessary, the ensuing estimates of variance and their propagation in the projections are likely underestimated. This needs to be noted when communicating results, especially to managers.

The model has been configured properly and consistently with standard practices. In fact, the configuration options are in some cases inspired by already peer reviewed assessments (SEDAR 39: Smooth Dogfish and ICCAT Shortfin Mako assessment). Noting the limitations above, ***the methods are appropriate for the available data*** and **the methods are scientifically sound and robust** (see TORs 3, 4, 7).

3) *Evaluate the assessment findings and consider the following:*

   Are abundance, exploitation and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

   Is the stock overfished? What information helps you reach this conclusion?

   Is the stock undergoing overfishing? What information helps you reach this conclusion?

   Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

   Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

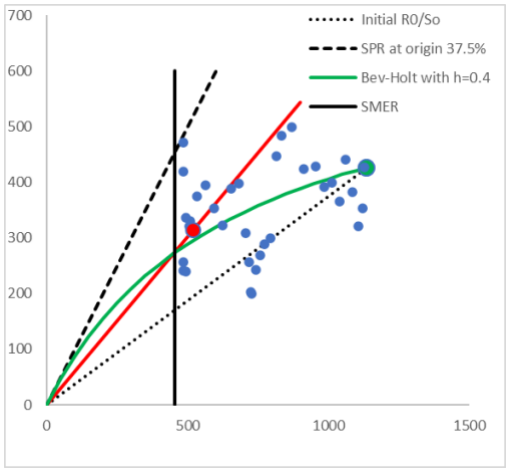
As noted, the stock assessment utilized the Stock Synthesis modeling platform, which integrated survey data, CPUEs, size frequencies, growth and reproductive life history with the catch history of the stock. The final model was selected after extensive examination of the data, alternative model structures and sensitivity diagnostic tests. The final model selected integrated these data in a biologically and statistically appropriate manner such that the ensuing estimates were useful for status inferences. **Thus, abundance, exploitation, and biomass estimates are reliable, consistent with input data and population biological characteristics, and are useful to support status inferences**.

The determination of status criteria is dependent on a key parameter of the stock recruitment relationship, h the steepness, or equivalently the slope of the stock recruitment curve at the origin. Steepness is also directly related to the stock size (SS) that produces maximum excess recruitment (MER: the stock size where the slope of the stock-recruitment relationship equals R0/S0 (Brooks, E. N., and Powers, J. E. 2007. Generalized compensation in stock-recruit functions: properties and implications for management. – ICES Journal of Marine Science, 64: 413–424). This, in turn, is related to stock size at MSY, being adjusted by the selectivity of the various fisheries a stock is subjected to. The point to noting these relationships is to emphasize that the steepness specification drives the determination of status criteria.

In typical groundfish cases, FSPR30%-40% is specified as a proxy for FMSY and $SS_{MSY}=SS_{FSPR30\%}$. In many of those cases steepness is fixed at levels associated with the chosen SSMSY proxy. The basis for h selection is often a quasi-meta-analysis leading to best practices leading to h's of ~0.7 to 0.9.

In the case of sharks, and in the case of Atlantic blacktip sharks in particular, the early life history can be better quantified than for groundfish. This was done for this assessment by examining the range of pups/gestation period, the length of the gestation period, age at maturity and other life history factors. Evaluating the variability of the life history parameters by simulation, the selection of a Beverton-Holt steepness of 0.4 was made. This was an appropriate conclusion. The specification of h=0.4 is the philosophical equivalent of specifying SPR30% in groundfish. However, I believe there is more basis for the 0.4 selection for blacktip than typically one uses for groundfish ~SPR30%. But as expected, the plot of S-R points from the base-case (Figure 1) indicates that the data themselves do not provide much information about h.

Figure 1 Stock-Recruitment plot from base case. The red dot and diagonal denote current (2018) conditions. The green dot denotes initial conditions with no fishing, R0 and S0



Note that an h=0.4 for a Beverton-Holt relationship implies that the SPR at the origin (R0/S0)/(Slope at origin)=SPR_{origin}=37.5%, S_{MER}/S0=0.40, SPR_{MER}=0.61.

The next question arose over the choice of the Beverton-Holt functional form and the effects on the scale and dynamics. Alternative forms were investigated during the assessment workshops. Additionally, in this review workshop a sensitivity was run in which h was estimated. Because the data were not informative, this was equivalent to specification of h=0.99 or that the R/S was a constant horizontal line over the range of S observed. The test was to determine the effect on the dynamics of a neutral S-R choice. Results indicate that the scale of the SS shifts with the SR choice, but the dynamics are similar (Figure 2). Additionally, the R deviations from the base case h=0.4 show a better pattern than the alternative SRs. This gives further support for the h=0.4 B-H choice.

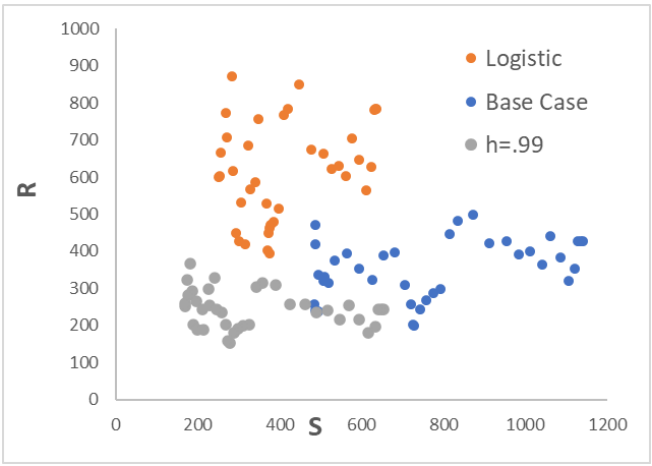
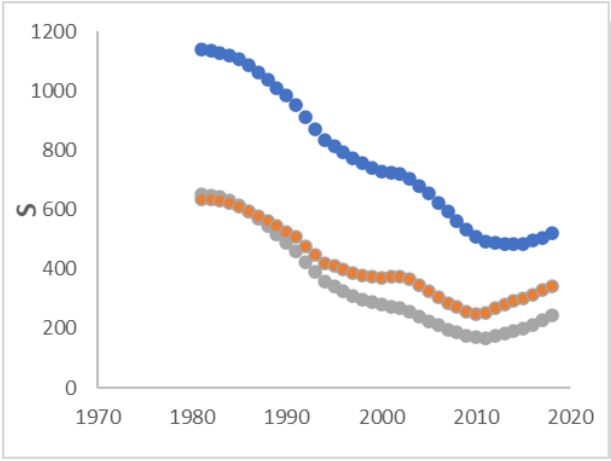
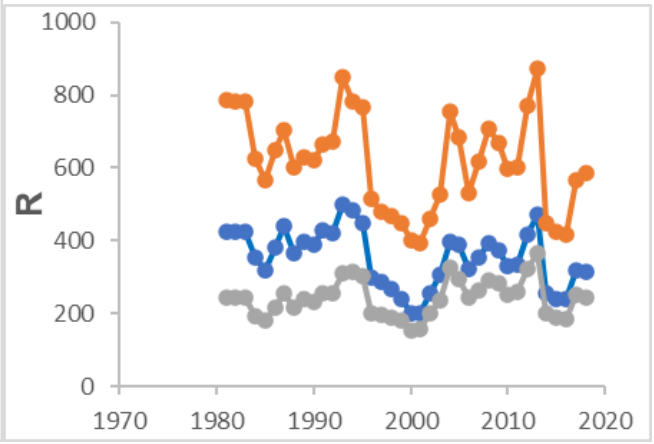


Figure 2. Plots of stock and recruitment from the base case and two sensitivities.





One thing to note is that in the blacktip model the SS is measured in the number of pups produced at birth (referred to as fecundity in the assessment) versus recruitment as the number of pups at recruitment time, both are on the same scale. This is not unique in fisheries to use fecundity as SS. However, since in blacktip, the R's and S's are closed in scale, then the S-R scaling parameters must be chosen under the constraint that S>R. It appears (Figure 2) that for the sensitivity runs this was not the case. What this means is that the scales of the sensitivity results are suspect and those might be viewed as indices. In the future this should be checked within the assessment model code. This issue was not constraining in the base case, where all S>R.

Another issue is the timing of recruitment and the duration of the stock-recruitment process. The assessment model as constructed assumed that the duration of the density dependent S-R model was very short (instantaneous?) followed by a density-independent period where both F and M were active (noted by having significant catches at sizes just greater than the size at birth). The M age 0 value in the density-independent period simply affects the scaling of R0, but in the future some effort needs to be given to aligning the duration of M0 and the M0 value itself.

A more important issue that could not be resolved with this assessment (and was not investigated) was the timing and duration of the density-dependent recruitment phase. In typical fish stocks all of the density dependence is assumed to occur in the recruitment phase prior to fishing. While density-dependence may occur later, it is probably negligible. But if density-dependence occurs jointly with fishing then the catch equation should be modified accordingly (SEDAR65-RD20, SEDAR65-RD21). In

the case of blacktip shark, it is less apparent when density-dependence is in effect (is it occurring in nursery areas?) and whether catches are occurring simultaneously. This needs to be investigated in the future. However, the base model SR assumptions are the most appropriate at this time.

Based on this discussion, I conclude that **there is there an informative stock recruitment relationship that is reliable and useful for evaluation of productivity and future stock conditions.**

Noting the above conclusions,

The quantitative estimates of the status determination criteria for this stock are reliable. The status determinations as required in the fisheries management plan (FMP) are estimated in this assessment, as well as estimates of variance. These should be the scientific basis for managers' decisions.

The stock is not overfished. The definition of an overfished condition is when the Spawning Stock Fecundity (SSF) is less than the Minimum Stock Size Threshold (MSST) where MSST is (1-M)SSF_{MSY}. M is the mean natural mortality rate of 0.139 in the base run. Thus, MSST=0.861 SSF_{MSY} in the base run. The assessment estimates that the current SSF_{2018} is 1.344 SSF_{MSST}. Thus, the stock is not overfished. Sensitivity analyses also found that SSF was greater than MSST, providing some supporting information to reach this conclusion. Also, note that SSMSY/SS0 =0.39 which is very close to SSMER/SS0=0.40

The stock is not undergoing overfishing. The definition of overfishing condition is when the fishing mortality rate (F) is greater than F_{MSY}. The assessment estimates that the current F_{2018}/F_{MSY} is 0.509 in the base case. Sensitivity analyses also found that F was less than $F_{MSY.}$ providing some supportive information to reach this conclusion.

4) *Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:*

   Are the methods consistent with accepted practices and available data?

   Are the methods appropriate for the assessment model and outputs?

   Are the results informative and robust, and useful to support inferences of probable future conditions?

   Are key uncertainties acknowledged, discussed, and reflected in the projection results?

The methods used in the projections were consistent with accepted practices and available data. The projections were done using the standard methods available within the SS3 modeling software. A complete MCMC to determine status probabilities was not available due to a lack of time. Instead, the status probabilities were calculated from the assumption of status statistics being normally distributed around their MLE value with a standard deviation equal to the estimated standard error based on the likelihood. This approach has been used before for sandbar sharks. For mako and sandbar sharks, the method was found to be consistent with MCMC but slightly more pessimistic about the TAC that would allow rebuilding. But even the MCMC applications use some fixed

parameters and model structure. Thus, it is not clear how best to characterize uncertainty of status. This is a common debate in the profession and there are no pat answers. Given that, the normal approximation around the *relative* statistics is most appropriate at this time.

The projection methods are appropriate for the assessment model outputs. Projections were made for a range of future constant catch scenarios to the year 2043. That year was chosen as a point where equilibrium was established and at a time of 2X mean generation time into the future. While this time horizon is useful to establish equilibrium conditions, in practice shorter term projections of abundance are more useful for management. The assessment approach built upon this fact when interpreting the projections.

The catch scenarios suite was conducted for the base case and several sensitivities. For each sensitivity, parameter/model changes were implemented into the assessment model and then the results of that assessment model were carried through into the projections. This was done to provide some characterization of the uncertainty in status criteria into the future.

The sensitivities were: base case, "logistic", "remove the CPUE", high and low catch scenarios, low catch scenario, high and low productivity scenarios. In all the projections, future selectivity was assumed to be the same as the average in recent years, recruitment was generated from the stock recruit relationship with deviations. The choice of these sensitivities was designed to provide a plausible range of future stock conditions. It is noted that this may not encompass all the uncertainty in projection outcomes (see TOR 5). However, it is an adequate first approximation. This is especially so, because the outcomes are measured in relative status (e.g., ratio of SSF/SSFMsy). In other words, there may be uncertainties in parameters leading to SSF, but those uncertainties enter into both the SSF and SSMSY such that the variance in the ratio should be less than otherwise.

The **results are informative and robust, and useful to support inferences of probable future conditions** in that they provide estimates of the probability SSF>SSF_{MSY}, SSF>MSST, and F>F_{MSY} based on the normal approximation to the distribution of the ratios. This allows evaluation of whether the standard of 70% has been met. The sensitivity analyses imply that the findings are generally robust to uncertainty.

Key uncertainties are acknowledged, discussed, and reflected in the projection results through the sensitivity analysis. The sensitivities range from the most optimistic high catch and low productivity cases, which imply that catches could more than double while still achieving management targets to the more pessimistic remove CPUE sensitivity case, which would require reduction in catches. Also, the MLE standard errors are perpetuated through the projections to approximate parameter uncertainty. The possibility that the normal approximation may underestimate the uncertainties that could be estimated by MCMC was adequately discussed. The effect of fixing parameters in the model (e.g., steepness, some selectivity parameters) was also discussed (see TORs 2,3,5), especially related to underestimation of uncertainty.

5) *Consider how uncertainties in the assessment, and their potential consequences, are addressed.*

Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.

Ensure that the implications of uncertainty in technical conclusions are clearly stated.

My ***comments on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods*** are as follows.

As in all assessments, the base case assessment model was constructed using some simplifying assumptions. Then the consequences of some of those choices and the uncertainty in the outcomes were evaluated by sensitivity analysis. As noted below, an MCMC analysis might have elucidated the variation in estimates better, but it could not be done in the time available. However, even using MCMC, some base model assumptions are assumed, so the uncertainty would still be "uncertain".

In terms of data weighting, key points were the selection of CPUEs/indices and their weighting and the weighting of length composition data. In the model construction, the variability of indices was computed and they were used as weights. Then length composition data were adjusted for effective sample size. This procedure carried some of the known uncertainties into the assessment model. The weighting balance between indices and length composition is a key factor of this (and any) assessment. The balance chosen here seems appropriate, but in the future, alternative decisions based on additional information/opinion might be useful.

Selection of indices was also an important component. The assessment team did explore the different ways of combining indices, as recommended from the data workshop. For age-0 the hierarchical Bayesian and dynamic factor analysis produced similar indices, so the latter was used. The assessment team also investigated the sensitivity of the results to alternative groupings of the indices. Finally, the impacts of uncertainty in the input data on stock assessment results and projections were investigated using sensitivity analyses.

Other parameters (M at age, sigmaR, h) were fixed, which inherently reduces the perception of uncertainty.

When assessment parameters were not fixed by the analysts, uncertainty in estimated and derived parameters was obtained from the Stock Synthesis output as the asymptotic parameter standard deviations at the converged solution. Time series trajectories of the two stock status metrics (SSF/SSFMSY, F/FMSY) are provided with approximate 95% asymptotic confidence intervals for the population and the projections.

A number of sensitivity runs were conducted some of which were diagnostics of individual parameters and others that used combinations of parameter sensitivities which were done to provide a perspective on the overall uncertainty. As always, the S-R parameters and selectivity assumptions were key.

The assessment team planned to provide estimates of credible intervals for reference points using MCMC techniques, but constraints associated with telework interfered with that plan, and only MLE results were available. However, the assessment team presented results for other shark species

assessments (sandbar and shortfin mako) that indicated that MCMC and MLE results were comparable, but the MLE estimates were slightly more conservative for the two examples provided.

The above discussion indicates **that the implications of uncertainty in technical conclusions are clearly stated.**

Having said all this, there is always a question about best ways to characterize uncertainty. "Best practices" have not been established yet, especially for shark life histories. This assessment uses the MLE estimates. In my opinion, it is likely that best practices will evolve toward external determination of status CVs based on life history and some basic data characteristics. We shall see. But in any case, whatever method is chosen it must be linked to the management harvest control rule and should be evaluated in an MSE context.

6) *Consider the research recommendations provided by the Data Workshop and Assessment Process and make any additional recommendations or prioritizations warranted.*

   Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.

   Provide recommendations on possible ways to improve the SEDAR process.

I find this TOR to be redundant to TOR 8. Thus, the discussion here is limited and the crux of my discussions on research and modeling are included in TOR 8. Thus, **Research and monitoring recommendations that could improve the reliability of, and information provided by, future assessments are clearly denoted in TOR 8.** An overview is provided here.

Possible data improvements include: fleet specific weight conversions, historical catch reconstructions, and length compositions from the recreational sector.

Biological/ecological improvements are needed on longevity/M, possible climate related changes in migration and relative abundance at age.

Possible assessment evaluation: explore alternative S-R approaches (models, timing, occurrence period of density-dependence); bootstrapping indices, further MCMC including in projections; further index standardization/selection criteria, especially multi-species approaches for recreational CPUE; alternative approaches to selectivity time blocks.

Provide recommendations on possible ways to improve the SEDAR process. The SEDAR process for this meeting was well organized. The meeting was efficient. The assessment panel was able to quickly answer questions and produce new runs and requested diagnostics. So, within the constraints imposed by Covid-19, this meeting was close to optimal. The support staff was excellent and very helpful.

The meeting would have benefited from a focus on presentations that are easier to follow in a remote meeting using larger fonts and more figures, rather than scrolling text. Meeting planners

should be cognizant of this because remote attendees will have different screen capabilities with home-covid offices.

Having an assessment review online is not a complete substitute for an actual review meeting. In-person meetings provide some benefits that cannot be readily duplicated. Typically, in remote meetings the discussion is slower, and hence fewer issues are raised. Also, you cannot easily stand up and make an illustrative drawing where needed. Furthermore, the sharing of knowledge, which for other review meetings has been substantial (e.g., sharing tips and tricks of modelling, or introduction to new tools or software) does not happen if all breaks are in isolation. Having informal discussions in person is much better for networking between assessment panel and reviewers, and overall makes the meetings more pleasant and productive. Nevertheless, we all realize the Covid-19 situation we are now in.

7) *Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.*

The stock assessment constitutes the best scientific information available. The assessment has gone through several stages of peer review through the SEDAR 65 process, including reviews of data inputs, assessment model structure and application, and the interpretation of results in terms of status determinations. These reviews provided public participation for transparency and comment and were inclusive of a wide array of contributing scientists. These processes promoted objectivity and verification/validation. The assessment is relevant to the management needs of the FMP.

The timeliness is limited in that the last year of data is 2018 and the assessment/review process is lengthy. That is the tradeoff of having detailed reviews. The life history of Atlantic blacktip is such that large annual changes in biomass are not expected, nor are the annual catches expected to change much. Nevertheless, managers should keep that in mind when scheduling future assessments.

Note the Peer Review Summary response to TOR 7 is virtually identical to my answer above, precisely because the Summary adopted my wording. In addition, my conclusions arise from my responses to the TORs above, especially related to TORs 3, 4 and 5.

8) *Provide suggestions on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.*

A number of recommendations for improvement of the assessment and general scientific understanding were discussed at the Review Panel meeting. These were listed in the Peer Review Summary submitted by the Review Panel. I concur with each of those recommendations. Those recommendations are listed below with the addition of my explanatory or mitigating comments, as appropriate.

Also, as noted previously, I find this TOR to be redundant to TOR 6. Thus, the discussion in TOR 6 is limited and the crux of my discussions on research and modeling are included here in TOR 8.

Recommendations for research activities:

- Species and fleet specific conversions between dressed weight and whole weight should be considered. These might also be time dependent in that standard practices may have evolved.
- A multi-species analysis of catch rates in the recreational fishery might be useful to extract an abundance index that is not biased by the issues with identification of sharks that are released alive. This is a general recommendation, but it implies the development of models of targeting and other species-interaction issues which are not insignificant.
- Longevity is poorly estimated and is one basis for estimates of M. Better estimates of longevity, and an independent estimate of natural mortality, for example from a tagging study, would be useful. Tagging is almost always useful but is often expensive.
- The data workshop discussed whether blacktip sharks may be migrating northward. This migration could be modeled in a spatially explicit assessment. Spatially explicit models might also be useful for explaining differences in trends in indices from different locations.
- The apparent dome shaped selectivity in several gears implies that there are sharks in the population that are older than the oldest individual observed. Whether this is realistic could be validated with fishery independent research. Again, tagging would be useful in this regard.

Recommendations for improvements to data for the assessment:

- The lack of data on catches and size distribution of catches during the peak of the fishery in the 1980s remains a key uncertainty in this assessment. Future work to improve catch reconstruction or evaluate model sensitivity to the catch reconstruction is recommended. It is unclear how this might be done.
- There is a need to better characterize the length composition, particularly in recreational fisheries, which may be influenced by both state and federal regulations.

Recommendations to the assessment methods:

- Model runs that do not fix parameters should be explored to more accurately characterize the uncertainty in parameter estimates. For example, if there is not enough data to estimate a selectivity parameter for two time blocks, rather than estimating it in one time block and applying the estimated value as a fixed parameter in the other, the data from both time blocks can be pooled to estimate the parameter.
- Bootstrapping the data could be used to quantify the uncertainty contained in the data. Current estimates of uncertainty are conditional on the full dataset and the modeling assumptions.
- Projections should be done using MCMC or profile likelihood methods to evaluate whether the normal approximation was adequate.

- Further research is needed on inconsistency of indices including the hierarchical models considered in this analysis.
- Improved model diagnostics are needed, as recommended in the assessment panel report and as described under TOR 6.
- Explore whether some other functional form of the stock recruit relationship would be more appropriate for this species, such as the low fecundity model that was used in the low productivity sensitivity. Explore using reference points that do not depend on MSY such as SPR-based reference points. As I note in TOR 3, effectively the specification of h for a B-H curve is the same thing as specifying SPRMSY. However, this might be explored with other SR functions.
- Investigate the timing and duration of the recruitment period, the duration of age 0 natural mortality, and the possibility of age 0 catches occurring during the recruitment period. I believe this is an important point that affects status criteria.

9) *Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.*

The Peer Review Summary was prepared and submitted to SEDAR.

Conclusions and Recommendations in Accordance with the TORs

The assessment and status of Atlantic blacktip shark resources off of the east coast of the United States was based on an application of the Stock Synthesis 3 modeling platform which was adapted to the data constraints and life history requirements of blacktip sharks. The model development was aligned with normal assessment practices with some unique applications due to blacktip life history.

Key findings from the review were: the stock assessment constitutes the best scientific information available; the stock is not overfished because the current (2018) Spawning Stock Fecundity is 134% of the Minimum Stock Size Threshold; and the stock is not undergoing overfishing, as the current F_{2018} is 50.9% of F_{MSY}.

A suite of projections and their uncertainties were presented that indicated the likely outcomes of status criteria (stock size and fishing mortality rate relative to Maximum Sustainable Yield). These projections indicate that status quo catches are likely to increase Spawning Stock Fecundity over the next few years due to current recruitment and to maintain Spawning Stock Fecundity above MSY-related levels in the longer term.

While this blacktip assessment is "data-rich" by typical shark assessment standards, there were still a number of issues arising from limited data. Improvements are needed on basic life history including natural mortality rates, stock-recruitment processes and spatial distribution and migration. Additional improvements might be achieved in understanding historical catch levels and improved length compositions. Also, future assessment improvements might arise from further methods to

characterize uncertainty. These are the most striking limitations. Nevertheless, the current assessment provides an adequate scientific basis for management decisions.

This review was conducted remotely due to Covid-19-related issues familiar to us all. The remote review functioned reasonably well with the obvious limitations due to local connections, having only one screen for a presenter and one-person conversations. Thus, I believe the Review was conducted successfully. Nevertheless, there are added benefits to an in-person meeting that were lacking here: namely, the ability to interact more informally with participants and to generate *ad hoc* discussions through drawing on a white-board. But all in all, the review worked well.

Appendix 1: Bibliography of materials provided for review

Documents Prepared for SEDAR 65 Review Workshop			
SEDAR65-RW01	Updated Commercial Gillnet Length Composition Data for use in SEDAR 65	Dean Courtney, Alyssa Mathers, and Andrea Kroetz	9/18/2020
SEDAR65 RW02	Projections Conducted for the Atlantic Blacktip Shark Stock Synthesis Base Model Configuration at Alternative Fixed Total Allowable Catch (TAC) Limits	Dean Courtney	10/5/2020
Reference Documents			
SEDAR65-RD15	Marine Recreational Information Program Transition to Improved Survey Designs	John Foster and Kelly Denit	10/22/2020
SEDAR65-RD16	APAIS At-a-Glance	NOAA Fisheries, Marine Recreational Information Program	10/22/2020
SEDAR65-RD17	Field Procedures Manual: Access-Point Angler Intercept Survey	Atlantic Coastal Cooperative Statistics Program	10/22/2020
SEDAR65-RD18	National Marine Fisheries Service's Marine Recreational Information Program Survey Design and Statistical Methods for Estimation of Recreational Fisheries Catch and Effort	Katherine J. Papacostas and John Foster	10/22/2020
SEDAR65-RD19	Review of the Marine Recreational Information Program. .	The National Academies of Sciences, Engineering, and Medicine	10/22/2020
SEDAR65-RD20	Age-specific natural mortality rates in stock assessments: size-based vs. density-dependent	Joseph E. Powers	10/30/2020
SEDAR65-RD21	Modelling the effects of density-dependent mortality in juvenile red snapper caught as bycatch in Gulf of Mexico shrimp fisheries: Implications for management	Robyn E. Forrest, Murdoch K McAllister, Steven J.D. Martell, Carl J. Walters	10/30/2020

Appendix 2: Performance Work Statement

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

SEDAR 65 HMS Atlantic Blacktip Shark Assessment 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. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The **SouthEast Data, Assessment, and Review (SEDAR)** is the cooperative process by which stock assessment projects are conducted in NMFS' Southeast Region. SEDAR was initiated to improve planning and coordination of stock assessment activities and to improve the quality and reliability of assessments.

SEDAR 65 will be a CIE assessment review conducted for HMS Atlantic Blacktip Shark. The review workshop provides an independent peer review of SEDAR stock assessments. The term review is applied broadly, as the review panel may request additional analyses, error corrections and sensitivity runs of the assessment models provided by the assessment panel. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stocks assessed through SEDAR 65 are the Atlantic stock of blacktip sharks in U.S. federal waters from Maine through Florida. The specified format and contents of the individual peer review reports

are found in **Annex 1**. The Terms of Reference (TORs) of the peer review are listed in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3** and the technical specifications required for this review are listed in **Annex 4.**

Requirements

NMFS requires three (3) reviewers to conduct an impartial and independent peer review in accordance with the Performance Work Statement (PWS), OMB guidelines, and the TORs below. The reviewers shall have a working knowledge in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of providing peer-review advice in compliance with the workshop Terms of Reference fisheries stock assessment. It would be preferable for reviewers to have an expertise in shark population dynamics and/or shark assessments.

Tasks for Reviewers

1) Two weeks before the peer review, the Project Contacts will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the Project Contacts will consult with the contractor on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the PWS scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

2) Additionally, two weeks prior to the peer review, the CIE reviewers will participate in a test to confirm that they have the necessary technical specifications provided in Annex 4 prepared in advance of the panel review meeting.

3) 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 answer any questions from the reviewers, and to provide any additional information required by the reviewers.

4) After the review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.

5) Each reviewer should assist the Chair of the meeting with contributions to the summary report. The Chair is not provided by the CIE under this contract.

6) Deliver their reports to the Government according to the specified milestones dates.

Place of Performance

The place of performance shall be online via gotowebinar.

Period of Performance
The period of performance shall be from the time of award through January 2021. Each CIE 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 two weeks of award	Contractor selects and confirms reviewers
2 weeks prior to the panel review	Contractor provides the pre-review documents to the reviewers
October 29, 30 and November 2, 4, 5 2020	Panel will attend and participate in review webinars lasting approximately four and a half hours each day held between the hours of 8 am -8 pm CT
Approximately 3 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

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; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
Since this is a remote panel review, travel is neither required nor authorized for this contract.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contacts:
Larry Massey – NMFS Project Contact
150 Du Rhu Drive, Mobile, AL 36608
(386) 561-7080
larry.massey@noaa.gov

Kathleen Howington - SEDAR Coordinator
Science and Statistics Program
South Atlantic Fishery Management Council
4055 Faber Place Drive, Suite 201 North Charleston, SC 29405
Kathleen.howington@safmc.net

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

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.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, 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 summary report that they believe might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

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 2: Terms of Reference for the Peer Review

SEDAR 65 Atlantic Blacktip Shark Assessment
Review Workshop Terms of Reference

Review Workshop Terms of Reference

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
   a. Are data decisions made by the DW and AP sound and robust?
   b. Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c. Are data applied properly within the assessment model?
   d. Are input data series reliable and sufficient to support the assessment approach and findings?
2. Evaluate and discuss the strengths and weaknesses of the method(s) used to assess the stock, taking into account the available data, and considering the following:
   a. Are methods scientifically sound and robust?
   b. Are assessment models configured properly and consistent with standard practices?
   c. Are the methods appropriate for the available data?
3. Evaluate the assessment findings and consider the following:
   a. Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b. Is the stock overfished? What information helps you reach this conclusion?
   c. Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d. Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e. Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:
   a. Are the methods consistent with accepted practices and available data?
   b. Are the methods appropriate for the assessment model and outputs?
   c. Are the results informative and robust, and useful to support inferences of probable future conditions?
   d. Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   a. Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.
   b. Ensure that the implications of uncertainty in technical conclusions are clearly stated.
6. Consider the research recommendations provided by the Data Workshop and Assessment Process and make any additional recommendations or prioritizations warranted.

a. Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
b. Provide recommendations on possible ways to improve the SEDAR process.

7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.
8. Provide suggestions on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.
9. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

Annex 3: Tentative Agenda - SEDAR 65 Atlantic Blacktip Shark Assessment Review

Via webinar

October 29 - November 5, 2020

Each day will consist of a 4.5 hour long webinar held between the times of 8 am and 8 pm CT
The start and end times of each webinar are dependent on CIE and analyst availability

October 29- Introductions and Opening Remarks		Coordinator
	- Agenda Review, TOR, Task Assignments	
	Assessment Presentations	Dean Courtney
October 30 – Assessment Presentation continued		Dean Courtney

October 29 and 30 Goals: Initial presentations completed, sensitivities and modifications identified.

November 2 -	Panel Discussion	Chair
	- Review additional analyses, sensitivities	
	- Consensus recommendations and comments	Chair

November 2 Goals: Final sensitivities identified, preferred models selected, projection approaches approved, Summary report drafts begun

November 4 - Panel Discussion		Chair
	- Final sensitivities reviewed.	
	- Projections reviewed.	
November 5 Panel Discussion or Work Session		Chair
	- Review Consensus Reports	

November 4 and 5 Goals: Complete assessment work and discussions. Final results available. Draft Summary Report reviewed.

Annex 4: SEDAR 65 HMS Atlantic Blacktip Shark Review workshop minimum technical requirements

1. Computer
2. Microphone and speakers (headset recommended)
3. GoToWebinar desktop app (JavaScript enabled) available for download here: https://support.goto.com/webinar/help/download-now-g2w010002
4. Internet: 1 Mbps or better (wired preferred)
5. Web browser:
    a. Google Chrome v57 or later
    b. Mozilla Firefox v52 or later
    c. Internet Explorer v10 or later
    d. Microsoft Edge v12 or later
    e. Apple Safari v10 or later
6. Operating system
    a. Windows 7 - Windows 10
    b. Mac OS X 10.9 (Mavericks) - macOS 10.15 (Catalina)
7. 2GB of RAM (minimum), 4GB or more of RAM (recommended)
8. Smart phone for use as audio backup and internet hotspot (recommended)

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

List of Participants

Review Panelist

Beth Babcock	Chair	University of Miami: RSMAS
Anders Nielsen	CIE	DTU-Aqua Technical University of Denmark
John Neilson	CIE	Independent fisheries Scientist
Joseph Powers	CIE	Independent fisheries Scientist

Analytical Representatives

Dean Courtney	Lead Assessment Representative	NMFS: HMS
Xinsheng Zhang	Assessment Representative	NMFS: HMS
Enric Cortes	Assessment representative	NMFS: HMS

Council and Agency Staff

Kathleen Howington	Coordinator	SEDAR
Karyl Brewster-Geiz	HMS Management	NMFS: HMS
Clifford Hutt	HMS Staff	NMFS: HMS
Heather Baertlein	HMS Staff	NMFS: HMS

Review Workshop Attendees

Catherine Puma	Observer	University of Miami
Chip Collier	Observer	SAFMC
John Carlson	Observer	NMFS
Julie Neer	Observer	SEDAR
Manoj Shivani	Observer	NTVI Federal
Rusty Hudson	Observer	DSF