Center for Independent Experts (CIE) External Independent Peer Review

SEDAR 54 HMS Sandbar Shark Assessment Review

By

Jean-Jacques Maguire
1450 Godefroy Ave
Quebec City, Qc, Canada
G1T 2E4

jeanjacquesmaguire@gmail.com

For the
Center for Independent Experts

December 2017

Executive Summary

This CIE review is for the Southeast Data, Assessment, and Review (SEDAR) 54 assessment of the Highly Migratory Species (HMS) sandbar sharks.

The data decisions made by the assessment panel appear sound and robust. Generally speaking, they are consistent with those made in the previous assessment in SEDAR 21 and, where they were changed, the changes were relatively small or fully justified with new research results and or new data becoming available. Future assessment should investigate how much sandbar shark is being caught in the Caribbean that could belong to the same stock, but is not currently included in the assessment. Data are used correctly in the assessment model. For 1960-1990, total commercial landings were apportioned into Gulf of Mexico (GOM) and Atlantic (ATL) using the average percent composition by region for the first five years with more reliable data (1991-1995). An alternative would have been to use a single commercial fleet for the two regions prior to 1991. This could be tested in the next assessment. In all, eleven indices of stock size are used in the assessment. Most of them are surveys in different locations on different life stages, except indices 2 and 3, which are based on the commercial fishery. The various stock size indices may be monitoring different age / size portions of the stock . This should be investigated further in the next assessment.

SS3 is a scientifically sound assessment method, but depending on what data are used and how it is configured, results can be volatile. In the case of SEDAR 54 sandbar shark, the results appear robust. The replication analysis, the base case and the sensitivity analyses suggest that the assessment model is configured adequately and consistent with standard practice. SS3 is a very appropriate assessment tool for the data available for sandbar shark, and the differences between the previous model and the current one are clearly documented and described.

Assessment outputs, except for parameter estimates and recruitment deviations, are presented as aggregate or in figures. This makes it difficult to evaluate the consistency of the assessment with input data. However, there is sufficient information in the report to conclude that stock status is highly uncertain, mostly because of inconsistent trends in the indices of stock size. All model runs suggest that the stock decreased substantially after catches increased markedly starting in the early 1980s, but the extent of the decline and rate of recovery after catches were decreased from about 2008 vary between the base case and the sensitivity cases using subsets of stock size indices. All the results are consistent with the input data and population biological characteristics, but there is no basis in the report to choose which one is most reflective of the true stock trends. Under the base case scenario, the stock is overfished, but overfishing is not occurring. The quantitative estimates of status determination show conflicting trends depending on the subset of stock size indices that is used. Therefore, they cannot be considered as entirely reliable. Further examination of the stock size indices, perhaps particularly the fishery independent ones, to try to identify which ones can be considered reliable indices of changes in stock size and for what size / age range might provide an alternate basis to inform managers about stock trends and conditions.

The projections are done internally from the SS3 software including the estimates of uncertainties. The results are informative and useful to support inferences of probable changes in relative abundance and key uncertainties are acknowledged, discussed, and reflected in the projection results.

The base and sensitivity cases, extended to projections, capture and reflect the significant sources of uncertainty.

Background

This CIE review is for the Southeast Data, Assessment, and Review (SEDAR) 54 assessment of the Highly Migratory Species (HMS) sandbar sharks. The desk review provides an independent peer review of SEDAR stock assessments. The review is responsible for ensuring that the best possible assessment is provided through the SEDAR process and provide guidance to the Southeast Fisheries Science Center (SEFSC) to aid in their review and determination of best available science, and to HMS when determining if the assessment is useful for management. The Terms of Reference (ToRs) of the peer review are attached in **Appendix 2**.

Originally, management of the shark fisheries was the responsibility of the five fishery management councils on the Atlantic and Gulf coasts. Fishery Conservation Amendments of 1990 (Pub. L. 101-627) amended the Magnuson Fishery Conservation and Management Act to give authority (effective January 1, 1992) to the Secretary of Commerce to manage fisheries on Highly Migratory Species in the exclusive economic zone (EEZ) of the Atlantic Ocean, Gulf of Mexico, and Caribbean Sea under authority of the Magnuson-Stevens Act (16 U.S.C. §1811). The Secretary delegated authority to manage Atlantic fisheries on HMS to the National Marine Fisheries Service (NMFS).

The sandbar shark was first assessed individually in 1998 and later in 2002, 2006, and 2011. Prior to that, it was part of the Large Coastal Shark complex, which was first assessed in 1991 and subsequently updated in 1994, 1996, and 1998. The first assessment of sandbar sharks under the SEDAR framework took place in 2006 (SEDAR 11) and the most recent one was in 2011 in SEDAR 21.

This review is for the stock assessment in the SEDAR 54 assessment Process which was held via a series of webinars between May 2017 and August 2017.

Description of the Individual Reviewer's Role in the Review Activities

The assessment document was received at the end of the day on October 23, 2017. I undertook to review the document shortly after, consulting background working papers and the previous SEDAR 21 assessment as necessary. I found the documents complete and informative, but would have appreciated easier access to data inputs and model results such as is provided by

the North-East Fisheries Science Center (NEFSC) for the Groundfish species it assesses (https://www.nefsc.noaa.gov/saw/sasi/sasi_report_options.php).

Summary of Findings for each ToR in which the weaknesses and strengths are described

1. Prepare a Peer Review Report that summarizes the Reviewer's evaluation of the stock assessment and addresses each of the following Terms of Reference.

My review of the assessment is described below under ToRs 2 to 8. In reviewing the report, I found a number of typos, but did not keep track of them. Two that may warrant attention are:

- In Section I Assessment history and review, paragraph 3 erroneously states that overfishing was occurring, when in fact F2009/FMSY ranged from 0.29 to 0.93.
- On page 14 of Section II in the paragraph on the VIMS Longline survey, reference is made to "The previous assessment (in 2004) …". Presumably the intention was to refer to "A" previous assessment, as "The" previous assessment was in 2011.

2. Evaluate the data used in the assessment, addressing the following:

a) Are data decisions made by the assessment panel sound and robust?

The data decisions made by the assessment panel appear sound and robust. Generally speaking, they are consistent with those made in the previous assessment in SEDAR 21, and where they were changed, the changes were relatively small, e.g., changing the maximum age from 29 to 31 or steepness from 0.29 to 0.30, or fully justified with new research results and/or new data becoming available, e.g., the new von Bertalanffy growth coefficient, new estimates of M, the integration of the three fisheries and eleven CPUE's in the assessment, etc.

The assessment covers the US Atlantic seaboard from Cape Cod south and includes the Gulf of Mexico. The FAO Fact Sheet (http://www.fao.org/fishery/species/2807/en) states "*In the Western Atlantic Springer (1960) suggested that there are two stocks or subpopulations of sandbar sharks, a northern major one from the US Atlantic seaboard to the Gulf of Mexico and the eastern Caribbean, and a minor South American one from Trinidad eastwards and southwards to Brazil.*" If this is true, it implies that the assessment may not cover the entire stock area. This is a potential shortcoming / source of uncertainty, and it is acknowledged in the last paragraph of the report on page 53. Future assessment should investigate how much sandbar shark is being caught in the Caribbean which could belong to the same stock, but is not currently included in the assessment.

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

Reporting on data uncertainties was not specified in SEDAR 54 terms of reference, therefore uncertainties in the data are not highlighted in the report. It is however possible to get a sense

of the uncertainties related to the indices of stock size from the first part (Model Fits to Abundance Indices) of section 3.2.5 Base Case Model Results. In the last part of the report, uncertainties in the data, including in the sampling for length composition and for relative changes in abundance (the stock size indices) are acknowledged. This is somewhat late in the report and this discussion could have been included earlier in the report.

c) Are data applied properly within the assessment model?

Data are used correctly in the assessment model. The assumption in the replication analysis and in the base case that the stock was close to virgin stock in 1960 is probably reasonable. Unlike in more northern waters where large stocks of commercial species (e.g., cod, herring, mackerel) have been the subject of important fisheries for centuries, fisheries in the southern end of the range of sandbar shark and in the Gulf of Mexico are less likely to have caught large quantities of shark. According to (https://www.nefsc.noaa.gov/nefsc/Narragansett/sharks/sandbar-shark.html), the northern end of the area of distribution of sandbar shark is Cape Cod, MA. Relatively large demersal fisheries developed off and south of Cape Cod from the mid 1950s mostly by European distant water fleets. These however were mostly offshore fisheries and sandbar shark is an inshore species unlikely to have been seriously impacted by the offshore distant water fleet fisheries.

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

Generally speaking, the input data series appear sufficient to support the assessment approach and findings. However, the commercial landings are split between the Gulf of Mexico and the Atlantic because of different sizes being caught in each area. For 1991 – 2015, the split used the percentage by region and year from the general canvass data (1991-2012) or from the HMS eDealer database (2013-2015). Prior to 1991 regional landings data for 1987-1990 fluctuated widely from one area to another. Therefore, for 1960-1990, total commercial landings were apportioned into GOM and ATL using the average percent composition by region for the first five years with more reliable data (1991-1995). This is consistent with what was done in SEDAR 21, but the observation that the percentages fluctuated widely from year to year during 1987-1990 suggests that the allocation of commercial catch by region prior to 1991 is more uncertain. An alternative would have been to use a single commercial fleet for the two regions prior to 1991. This could be tested in the next assessment.

At the end of the Section F3 Recreational and Mexican catches, the report says that Mexican catches for 2014 and 2015 were assumed equal to the mean of those in 2011-2013. In table 2.3, recreational and Mexican catches are combined in a single column and the 2014 and 2015 values differ slightly. Presumably, the statement in the report that Mexican catches in those years are the average for 2011-2013 is correct, but there is no way this can be checked.

In all, eleven indices of stock size are used in the assessment. Most of them are surveys in different locations on different life stages, except index 2 and 3 which are based on the commercial fishery. All indices are standardised, except possibly index 4, the VIMS longline

survey, where the text states that nominal and standardised indices were presented, but there is no indication of how the index was standardised. For the other indices, various approaches have been used to standardise the indices, e.g., GLM, delta-lognormal, two-steps delta-lognormal, delta-lognormal with stepwise forward incorporation. A short statement of why a given approach was preferred in each case would have been useful.

The caption of figure 3.2.2 says "Available length frequency data by fishery and survey, aggregated across years, used in the base case model configuration". It is not clear from the caption or from the text if yearly values or the aggregates were used in SS3, but based on figures and text later in the report, it seems that SS3 was fit to size compositions aggregated over year as the shape of the size compositions suggests. It is not clear what the numbers in each cell of figure 3.2.2 represent: in some cases, they are close to the effective sample size given in table 3.2.2, but in other cases they are different.

Page 34 paragraph 1: "*Cross-correlations identified strong autocorrelation in some CPUE indices over 2 to 3 years, which could indicate a year-class effect. Cross-correlations also identified strong cross correlation of lagged values of some CPUE indices (at lags between 2 to 10 years) with the current values of other CPUE indices, which could indicate that some CPUE indices represent younger age-classes than others. However, the specific lagged relationships with high correlation were not consistent among the series. Further information can be found in section 3.2.7 and SEDAR54_TEMP1*". This suggests that the indices may not all index the same size / age range. That the various stock size indices may be monitoring different age / size portions of the stock is apparent in the selectivity estimates shown in figure 3.1.3. This should be investigated further in the next assessment.

3. Evaluate the methods used to assess the stock, taking into account the available data.

a) Are methods scientifically sound and robust?

SS3 is a scientifically sound assessment method, but depending on what data are used and how it is configured, results can be volatile. In the case of SEDAR 54 sandbar shark, the results appear robust.

It is relatively rare that the Terms of Reference for a stock assessment specifically state that a different assessment method should be used and what that new method should be. Yet, this is done here where ToR 1 for the assessment process specifies that Stock Synthesis is to be used. The explanation for this can be found at the end of the assessment report, page 52 where it is stated

"*The use of Stock Synthesis as a modeling platform is due to the recommendations of the CIE Reviewers from SEDAR 21, which did not specifically recommend Stock Synthesis but did recommend the following:*

• *Estimating the fishery and index selectivities within the assessment model.*

• Development of a two sex model for more direct estimation of the spawning stock.

• Fitting the model to either length or age data. In addition to being necessary in order to estimate selectivities, these data can be informative about changes in age-specific abundance.

• Exploration of models that do not require an assumption that the population is at virgin levels at some point in time."

SS3 was meeting the first three criteria. The why SS3 was chosen over the previous assessment model, which had been accepted, could have been given earlier in the document.

b) Is the assessment model configured properly and used consistent with standard practices?

The previous assessment in SEDAR 21 used a state-space, age structured production model (ASPM) while the current one uses Stock Synthesis. The approach of doing a replication analysis, i.e., a reproduction of the previous base case assessment as well as the continuity analyses where new information is sequentially incorporated into the assessment, is a good way of identifying the effect of each data source on the assessment results.

Selectivity (page 35, last paragraph) was assumed to be time-invariant, yet there were considerable management changes over time that could be expected to have had an effect on selectivity. However, the length frequency sample sizes may be insufficient to estimate changes in selectivity and if SS3 was fit to size compositions aggregated over time, there was no data to evaluate changes in selectivity.

The replication analysis, the base case and the sensitivity analyses suggest that the assessment model is configured adequately and consistent with standard practice.

c) Are the methods appropriate for the available data?

Stock Synthesis is a standard assessment tool for many USA West Coast stock assessments that is being used increasingly in the ICES and ICCAT areas. It is a highly flexible assessment tool in which it is possible to use several sources of information (growth information, catch, length and age frequencies, indices of stock sizes, etc.) to evaluate stock status. Stock Synthesis is highly structured with many options and built-in assumptions; it can be configured to mimic several other types of assessment approaches. Because of its structure and underlying assumptions, Stock Synthesis can provide stock estimates and fisheries management benchmarks even when very little data are available. It is sometimes difficult to ascertain the most important influence on the assessment results: the data or the assumptions in the assessment model. SS3 is therefore very appropriate for the data available on sandbar shark.

d) Are differences between the current model and the previous model clearly documented and described?

Yes, the replication analysis compares the data and assumptions made in the previous assessment with those in the current one. However, I do not necessarily agree with the

conclusions: on page 30, end of second paragraph the report states that Figures 3.1.6 and 3.1.7 show a relatively good fit (compared to SEDAR 21) of the model to all the indices. This is not obvious when comparing these to figures 3.6 in SEDAR 21. The fit to the indices appears to me equally poor in both SEDAR 21 and SEDAR 54.

In addition to the replication analysis, the continuity runs changed input values one at a time to evaluate the effect of updating the catch series (Update_Catch), extending the catch series to 2015 (Cont_1), changing the longevity to 31 years (Cont_2), and using the new life history parameters (Cont_3). The results are shown in figure 3.1.8, page 85, although the choice of colors and the scale of the graph make it difficult to identify the runs other than SEDAR_21 and replication, it seems that updating the catch made little difference, extending the catch to 2015 resulted in higher biomasses from 1960 to the mid-1980s or so but lower biomass estimates from the mid-1990s to 2015, changing the longevity to 31 years made little difference compared to extending the catch to 2015. Using the new life history parameters resulted in yet lower biomass estimates for 1980 to 2015 and intermediate between SEDAR 21 and the replication analysis for 1960 to 1980. Similar to the replication analysis, differences in biomass seem larger at the beginning of the time series than towards the end.

4. Evaluate the assessment findings with respect to the following:

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

Assessment outputs, except for parameter estimates and recruitment deviations, are presented as aggregate or in figures. This makes it difficult to evaluate the consistency of the assessment with input data. This could be greatly improved by implementing something like the very useful web site (https://www.nefsc.noaa.gov/saw/sasi/sasi_report_options.php) of the NEFSC, where all the assessment information for Groundfish stocks is available. This makes it considerably easier to evaluate model configurations and results.

The above being said, there is sufficient information in the report to conclude that stock status is highly uncertain, mostly because of inconsistent trends in the indices of stock size. All model runs suggest that the stock decreased substantially after catches increased markedly starting in the early 1980s, but the extent of the decline and rate of recovery after catches were decreased from about 2008 vary between the base case and the sensitivity cases using subsets of the stock size indices. The sensitivity case with Neg_CPUE shows the largest decrease to about 10% SSF/SSF0 and nearly no increase after 2008. The sensitivity case with Pos1_CPUE shows a modest decrease to about half SSF/SSF0 followed by an increase to about 60% SSF/SSF0 in 2015. The base case is closer to Neg_CPUE with a decrease to 20% SSF/SSF0 and modest increase post 2008. All these results are consistent with input data and population biological characteristics, but there is no basis in the report to choose which one is most reflective of the true stock trends. It would have been informative to see the fits to the indices for the sensitivity cases, particularly in the case of POS1_CPUE.

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

As indicated above, there are considerable uncertainties about stock status. The base case suggest that the stock is overfished but that overfishing is not occurring. The sensitivity case with Neg_CPUE suggests that the stock is overfished and that overfishing is occurring, while the sensitivity case with Pos1_CPUE suggests that the stock is not overfished and that overfishing is not occurring. As indicated under 4a) above, there is no strong basis in the report to choose which of the base case and the two sensitivity runs is most reflective of the true stock trends. However, the panel did produce a base case which is the usual basis for status determination and under the base case, the stock is overfished.

There is a potential for a shifting baseline with models starting well before the fishery started. The replication analysis produced almost identical biomass estimates 1990 - 2015, but a markedly higher initial biomass estimate in 1960 with the difference becoming smaller and smaller over time. For this assessment, this does not make a difference in status determination for the base case, but it could have resulted in a change in status determination without any real change in actual stock status simply by estimating higher initial biomass and MSY biomass reference points.

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

As indicated above, there are considerable uncertainties about stock status. The base case suggests that the stock is overfished but that overfishing is not occurring. The sensitivity case with Neg_CPUE suggests that the stock is overfished and that overfishing is occurring, while the sensitivity case with Pos1_CPUE suggests that the stock is not overfished and that overfishing is not occurring. As indicated under 4a) above, there is no strong basis in the report to choose which of the base case and the two sensitivity runs is most reflective of the true stock trends. However, the panel did produce a base case which is the usual basis for status determination and under the base case, overfishing is not occurring.

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

The Beverton and Holt stock and recruitment relationship is shown in figure 3.2.9. There is no indication that the curve has reached an asymptote at the estimated R0 of 533000 pups. Recruitment deviations are estimated with relatively strong constraints only for a subset of years. While the relationship between spawning output and recruitment is very strong in figure 3.2.9, this is based on assumptions more than on data. In that context, I conclude that there is not an informative stock recruitment relationship, although one would be expected with sandbar shark's life history.

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?

As discussed under 4a-c above, the quantitative estimates of status determination show conflicting trends depending on the subset of stock size indices that is used. Therefore, they cannot be considered as entirely reliable. Further examination of the stock size indices, perhaps particularly the fishery independent ones, to try to identify which ones can be considered reliable indices of changes in stock size and for what size / age range, might provide an alternate basis to inform managers about stock trends and conditions. SEDAR54-WP-06 Example Implementation of a Hierarchical Cluster Analysis and Cross-correlations of Selected CPUE Indices for the SEDAR 54 Assessment is a good start in that direction.

f) Are base model runs, sensitivity runs, and alternate states of nature runs clearly described and reasonable?

Yes, all the runs are clearly described and reasonable. As indicated in the report, the different states of nature are probably more realistic and the sensitivity runs using subsets of stock size indices than in the low, medium or high productivity regimes.

5. Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:

a) Are the methods consistent with accepted practices and available data?

Yes, the projections are done internally from the SS3 software including the estimates of uncertainties. The projections did not forecast recruitment variability which is probably reasonable for a species such as sandbar shark.

b) Are the methods appropriate for the assessment model and outputs?

Yes, as indicated above, the projections were done internally with the SS3 software.

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

Yes, the results are informative and useful to support inferences of probable changes in relative abundance.

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

Yes, Projections were done for the base case, the POS1_CPUE and the NEG_CPUE. In that sense, they are very likely to bracket future stock conditions. And the MCMC projections done in each case for one of the TAC scenarios illustrate likely variability in future stock trends.

e) If the results indicate a new rebuilding schedule is required, are the scientific/technical reasons for the new schedule clearly articulated and appropriate?

Not applicable except for the sensitivity case with NEG_CPUE where the reasons for the new schedule are appropriate and clearly articulated.

6. 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

The base case along with sensitivity case POS1_CPUE and NEG_CPUE are likely to bracket past and current stock sizes. Deterministic constant TAC projections were done for each of the three scenarios, and MCMC projections were done with the TAC resulting in a 50% rebuilding by 2070, also for each of the scenarios. In my view, this captures and reflects the significant sources of uncertainty.

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

It is not possible in a desk review like this one to **ensure** that the implications of uncertainty in technical conclusions are clearly stated if they are not in the report. However, in this case, the implications of uncertainty in technical conclusions **are** clearly stated.

7. Consider the research recommendations provided 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.

Of the research recommendations listed in section 3.3, it would be particularly important to follow up on the stock structure. If a sizable and variable portion of the stock is outside the assessment area, changes in the assessment could be due to variability in presence instead of changes in stock size.

The report states in a few places that many stock size indices show interannual variability incompatible with the life history of the species. Clearly, more research into the factors influencing the presence and availability of sandbar shark to the surveys would be beneficial. As indicated above, a close examination of the available stock size indices should be undertaken to identify what size / age range they are monitoring and possibly eliminate those that do not appear to be indexing any of the size / age range.

The first of the research recommendations in section 3.3 is also important to estimate the growth potential of the stock, similar to what is done for marine mammal populations.

b) Provide recommendations on possible ways to improve the SEDAR process.

S54_Final_SAR.pdf section 1 states: "*SEDAR emphasizes constituent and stakeholder participation in assessment development, transparency in the assessment process, and a rigorous and independent scientific review of completed stock assessments*" (paragraph 1). Further down, in the same section, paragraph 4 the report states: "*SEDAR workshops are public meetings organized by SEDAR staff and the lead Cooperator. Workshop participants are drawn from state and federal agencies, non-government organizations, Council members, Council advisors, and the fishing industry with a goal of including a broad range of disciplines and perspectives. All participants are expected to contribute to the process by preparing working papers, contributing, providing assessment analyses, and completing the workshop report*". I thought I was a participant in the SEDAR 41 assessment process for red snapper as a member of the Science Center for Marine Fisheries, but my experience does not conform with the statements above. I attended webinars, but my microphone was not open, and could not be opened. If I wanted to make comments or ask questions, I had to use the chat box with the Coordinator. Only formal Panel members had their mikes open and could intervene directly. My second point is that scientific expertise on that particular assessment panel was limited. Judging from the comments and questions, there may have been 1 or 2 knowledgeable stock assessment panel members other than those directly involved in the stock assessment.

The SEDAR 54 Assessment Process was held via a series of webinars between May 2017 and August 2017. While webinars are a cost-effective way of keeping people involved and informed, they are not the best way to achieve a thorough and effective peer review – in person meetings are a better way of achieving in-depth peer review.

From the list of participants, there seems to have been little involvement of people not working for NMFS.

As indicated above under 4a) the review of assessment methods and results could be greatly improved by implementing something like the very useful web site (https://www.nefsc.noaa.gov/saw/sasi/sasi_report_options.php) of the NEFSC, where all the assessment information for Groundfish stocks is available.

8. Provide guidance on key improvements in data or modeling approaches that could be considered when scheduling the next assessment.

Stock assessments using the SS3 software can be volatile, with small changes in supposedly unimportant parameters causing large changes in stock size estimates. The report does not discuss if it was difficult to obtain a stable configuration for sandbar shark, but the sensitivity cases with POS1_CPUE and NEG_CPUE suggest that the assessment is not volatile.

Several assumptions can be built in the SS3 framework and it is sometimes difficult to evaluate if the results are influenced more by the assumptions than by the data. Again, the sensitivity cases indicate that the model is responsive to changes in the data.

In ICCAT, some of the shark stocks are assessed using Bayesian surplus production models. It could be informative to use such method on sandbar shark and see how the results compare with those of SS3.

Conclusions and Recommendations in accordance with the ToRs

2. Evaluate the data used in the assessment

The data decisions made by the assessment panel appear sound and robust. Future assessments should investigate how much sandbar shark is being caught in the Caribbean that could belong to the same stock, but is not currently included in the assessment. For 1960-1990, total commercial landings were apportioned into GOM and ATL using the average percent composition by region for the first five years with more reliable data (1991-1995). An alternative would have been to use a single commercial fleet for the two regions prior to 1991. This could be tested in the next assessment. The various stock size indices may be monitoring different age / size portions of the stock. This should be investigated further in the next assessment.

3. Evaluate the methods used to assess the stock, taking into account the available data

SS3 is a scientifically sound assessment method and its application to sandbar shark appears robust. SS3 is a very appropriate assessment tool for the data available for sandbar shark, and the differences between the previous model and the current one are clearly documented and described.

4. Evaluate the assessment findings

There is sufficient information in the report to conclude that stock status is highly uncertain, mostly because of inconsistent trends in the indices of stock size. The quantitative estimates of status determination show conflicting trends depending on the subset of stock size indices that is used. Further examination of the stock size indices, perhaps particularly the fishery independent ones, to try to identify which ones can be considered reliable indices of changes in stock size and for what size / age range, might provide an alternate basis to inform managers about stock trends and conditions.

5. Evaluate the stock projections, rebuilding timeframes, and generation times

The projections are done internally from the SS3 software including the estimates of uncertainties. The results are informative and useful to support inferences of probable changes in relative abundance and key uncertainties are acknowledged, discussed, and reflected in the projection results.

6. Consider how uncertainties in the assessment, and their potential consequences, are addressed

The base and sensitivity cases, extended to projections, capture and reflect the significant sources of uncertainty.

7. Consider the research recommendations provided and make any additional recommendations or prioritizations warranted

See main text.

8. Provide guidance on key improvements in data or modeling approaches that could be considered when scheduling the next assessment

See main text.

Appendix 1: Bibliography of materials provided for review

SEDAR54-WP-01. Updated life history parameters for sandbar sharks, Carcharhinus plumbeus. William B. Driggers III, Bryan S. Frazier, John K. Carlson, Bethany M. Deacy, Michael P. Enzenauer and Piercy, Andrew N.

SEDAR54-WP-02. Updated catch rates of sandbar sharks (Carcharhinus plumbeus) in the northwest Atlantic Ocean from the Shark Bottom Longline Observer Program, 1994-2015. John K. Carlson and Alyssa N. Mathers.

SEDAR54-WP-03. Standardized catch rates of sandbar sharks from the Large Pelagics Rod and Reel Survey 1986-2015. John Walter and Craig A. Brown.

SEDAR54-WP-04. Sandbar Shark Abundance Indices from NMFS Bottom Longline Surveys in the Northern Gulf of Mexico. Adam G. Pollack, David S. Hanisko and G. Walter Ingram, Jr.

SEDAR54-WP-05. Standardized catch rates for sandbar sharks from the U.S. pelagic longline observer program using generalized linear mixed models. Enric Cortés and Xinsheng Zhang.

SEDAR54-WP-06. Example Implementation of a Hierarchical Cluster Analysis and cross-correlations of Selected CPUE Indices for the SEDAR 54 Assessment. Dean Courtney.

SEDAR 54 Stock Assessment Report HMS Sandbar Shark.

Appendix 2: A copy of the CIE Statement of Work

Statement of Work

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

SEDAR 54 HMS Sandbar 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

Southeast Data, Assessment, and Review (SEDAR) 54 will be a compilation of data, a standard assessment of the stock, and CIE assessment review conducted for Highly Migratory Species (HMS) sandbar sharks. The desk review provides an independent peer review of SEDAR stock assessments. The review is responsible for ensuring that the best possible assessment is provided through the SEDAR process and will provide guidance to the Southeast Fisheries Science Center to aid in their review and determination of best available science, and to HMS when determining if the assessment is useful for management. The stock assessed through SEDAR 54 are within the jurisdiction of the Highly Migratory Species Division of NOAA Fisheries and the states of Texas, Louisiana, Mississippi, Alabama, Florida, Georgia, South Carolina, North Carolina, Virginia, Maryland, Delaware, Pennsylvania, New Jersey, New York, Connecticut, Rhode Island, Massachusetts, New Hampshire, and Maine. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements

NMFS requires three reviewers to conduct an impartial and independent peer review in accordance with the Statement of Work (SOW), OMB Guidelines, and the ToRs below. The reviewers shall have working knowledge and recent experience 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 ToRs. Experience with elasmobranches assessment methods would be preferred. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for reviewers

Each CIE reviewers shall complete the following tasks in accordance with the SOW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: Review the following background materials and reports prior to the review:

Working Papers, Reference Documents, and the Assessment Report will be available no later than 23 October 2017. All materials will be available on the SEDAR website:

http://sedarweb.org/sedar-54-dataassessment-process

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SOW and ToRs, and shall not serve in any other role unless specified herein. Modifications to the SOW and ToRs cannot be made during the peer review, and any SOW or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SOW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Place of Performance

Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Period of Performance

The period of performance shall be from the time of award through December 2017. Each reviewer's duties shall not exceed 10 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
Within four weeks of award	Contractor provides the pre-review documents to the reviewers
October 2017	Each reviewer conducts an independent peer review as a desk review
Within two weeks after review	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government
Within two weeks of Government receiving final reports	Government distributes final reports to Project Contact and SEDAR

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

Since this is a desk 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.

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 or not the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role 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.

3. The reviewer report shall include the following appendices:
    a. Appendix 1: Bibliography of materials provided for review
    b. Appendix 2: A copy of the CIE Statement of Work

Annex 2: Terms of Reference for the Peer Review

SEDAR 54 HMS Sandbar Shark Assessment Review

1. Prepare a Peer Review Report that summarizes the Reviewer's evaluation of the stock assessment and addresses each of the following Terms of Reference.
2. Evaluate the data used in the assessment, addressing the following:
    a) Are data decisions made by the assessment panel 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?
3. Evaluate the methods used to assess the stock, taking into account the available data.
    a) Are methods scientifically sound and robust?
    b) Is the assessment model configured properly and used consistent with standard practices?
    c) Are the methods appropriate for the available data?
    d) Are differences between the current model and the previous model clearly documented and described?
4. Evaluate the assessment findings with respect to 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?
    f) Are base model runs, sensitivity runs, and alternate states of nature runs clearly described and reasonable?
5. Evaluate the stock projections, rebuilding timeframes, and generation times, addressing 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?
    e) If the results indicate a new rebuilding schedule is required, are the scientific/technical reasons for the new schedule clearly articulated and appropriate?
6. 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.

7. Consider the research recommendations provided 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.

8. Provide guidance on key improvements in data or modeling approaches that could be considered when scheduling the next assessment.