Center for Independent Experts (CIE) Desk Review of the 2017 SEDAR 54 Sandbar Shark (Carcharhinus plumbeus) Stock Assessment

Prepared by:

John D. Neilson, Ph.D.
Independent Fisheries Consultant
1500 Balmoral Avenue
Comox, British Columbia, Canada
V9M 2P6

Email: largepelagicsscientist@gmail.com

Executive Summary

The 2017 SEDAR assessment of sandbar shark *(Carcharhinus plumbeus)* is a logical evolution of the previous assessment, conducted in 2011. Many (but not all) research recommendations from the previous assessment have been acted on, and a new modeling platform has been adopted that addresses concerns with the previous approach, including providing sex-disaggregated results, selectivity computed internally within the model and improved estimates of statistical uncertainty. The perception of stock status using the base case formulation is somewhat more negative than expected based on the results of the previous assessment, but those differences are discussed in the current document.

While the current assessment of sandbar sharks does benefit from seemingly adequate catch data and abundance index availability compared with most other elasmobranch assessments, serious reservations with the quality of those inputs remain. The main issues include the reconstruction of historic catches, and the adequacy of the abundance indices. Issues with the abundance indices include substantial unexplained interannual variation, conflicting trends of abundance indices, the absence of any index covering the more southern range of the distribution, and the absence of any index covering the most important remaining fishery. The interpretation of stock status depends on the grouping of the CPUE series, with greatly divergent conclusions reached depending on the choice of grouping. Such issues are considered to be serious deficiencies that challenge the credibility of the assessment, and make it difficult to conclude that the work constitutes the best science available.

More positive aspects of the analyses and assessment results include good fits to the length composition data, and the reasonable reconstruction of the population numbers at age.

Overall, it was considered that the population model and assessment results in SEDAR 54 should be employed for general guidance on stock status only, taken together with other indicators such as size/age composition information for the catch and population. Judging from the improving age structure and the recent increases in many of the indices of abundance, recent catch levels of < 200 t have not prevented the population from rebuilding following the period of relatively heavy exploitation during the 1980s and 1990s.

Among the key recommendations, more investment should be made in electronic tagging, providing a fisheries-independent method to determine linkages among the regions comprising the very large stock area. As part of the increased knowledge from migration studies, a conceptual model of the life history of the species is needed to help interpret the utility of candidate abundance indices. Abundance indices should be evaluated using "report card" tools similar to those developed by other fisheries management agencies, and in the previous assessment. The historic data reconstruction work is pivotal for this assessment, and should continue, as should continued work to determine post-release survival.

Background

The document presented here contains my independent peer review of the 2017 SEDAR assessment of sandbar shark. It is prepared under contract to the Center for Independent Experts, following the Statement of Work contained in Appendix 2.

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

My role in the process was as a CIE-appointed peer reviewer only, and the work was conducted as a "desk review".

Summary of Findings for each ToR

1. Evaluate the data used in the assessment, addressing the following:
a) Are data decisions made by the assessment panel sound and robust?

I will start off by complimenting the assessment team on the comprehensive review of the management and assessment history for the stock – I found this to be most helpful. However, the absence of a description of the biology of the species (reproduction, movement and migrations, age and growth) including key uncertainties was unfortunate. Having such information would have helped in the interpretation of the assessment results.

Still on the positive, I found the explicit treatment of discard mortality to be well done. While some of the inferences borrow from studies of other species (blue sharks), it was encouraging to see that this important source of mortality was included.

However, I found the treatment and evaluation of other aspects of the input data to be incomplete. For example, in some background Working Papers (such as Walters and Brown WP 03) the authors do provide some justification for data filtering of records that were not shark-directed and did not involve chumming. Such exclusions seem reasonable. The development of the CPUE indices reported in the supporting documents have followed different paths -- some included factors such as surface water temperature (and is found to be significant for proportion positive in Walter and Brown 2017) or management measure considerations, whereas others did not. It would have been helpful for the working papers to provide better background on why such factors were not included in the analyses.

The level of background documentation on the abundance indices is very uneven. For example, while there was a concise working paper provided for the longest running CPUE series (LPS, Walter and Brown 2017, WP 03), it did not contain any information that would allow a thorough evaluation of the utility of the abundance index, such as an Analyses of Deviance table similar to what is provided in WP 01 (Carlson and Mathers 2017). I would have liked to see output that reflected the relative contribution of factors included in the analyses to help evaluate the authors' assertion that the absence of water temperature information for many sets is an important issue that could be mitigated through use of satellite sea surface temperature data.

An explanation of why 2016 data were not included in the analyses would have been helpful.

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

The assessment is incomplete in this regard. As noted above, several of the available indices of abundance are supported with working papers that give some concise background on the development and interpretation of the abundance indices. But there are 11 abundance indices used in this assessment, and many abundance indices are not accompanied with detailed explanation of the series.

Concerning the fishery landings series, there is a pattern to the REC+MEX data that should be checked. Table 2.3 shows a ramping up of REC+MEX fishery over just a few years (15.3 t in year 1978 to 1861.8 t in 1983). The scale and rapidity of this increase does not seem credible.

The average weight of sharks in the catch also changed considerably over the above period (see text table below). This implies that the average weight of a shark caught in the REC+MEX fishery declined considerably over a short period. Some checking of the input data seems to be indicated, and an interpretation of this trend would be very helpful.

	#fish (Table 3.1.1)	mt (Table 2.3)	fish/t
1983	426,979	1861.8	229.34
1977	1079	3.4	317.35

Figure 2.1 shows the catch time series from SEDAR 21 and 54, and the remarkable year 1983 catch stands out. Better explanation of that extreme value is required. How credible are the catch data from that period? Is this an uncertainty that should be explored through sensitivity analyses?

Still on catch, the differences in timing of the peaks in Figure 2.1 and 2.2 (catches in weight and numbers) needs to be reconciled.

In general, there is insufficient discussion of the uncertainties associated with F3, the Recreational and Mexican catch. There is some assumption that the Mexican catch from certain states came from US waters (pdf P. 11). Isn't the important point that the catch came from the stock area? What proportion of the catch is subject to such assumptions? It is difficult to find the answers in the document to such questions concerning the input data. To get a better appreciation for this, I went back to the SEDAR 21 document, which gave a more complete accounting of the issue. It appears that the Mexican catch of sandbar sharks is in fact one of the key sources of uncertainty in the assessment, and its reconstruction is based on several assumptions, including historic species composition of the catch. However, I am still unclear as to the potential scale of this issue.

Finally on catch, I was wondering about catches by other nations in the stock area (for example, Central American countries with an Atlantic coast, and Cuba). Is there anything known about landings from those countries, and if not, this source of uncertainty should be acknowledged.

Considering age-structured data, I recognize that there has been further improvement of the understanding of sandbar shark age and growth, including age validation.

Turning to the indices of abundance, it is very unfortunate that the majority of the 11 included *indices do not cover the periods of development and heavy exploitation of the fishery*. Also, considering the overall range of the stock, the abundance indices are concentrated in the northwest portion of the range (see Figure 3.2.1). It would also have been helpful to have a conceptual model of the life history and migrations of sandbar sharks to help understand the applicability of the various indices of abundance. I understand the migration patterns are at least partially known, and include both seasonal and ontogenetic components. However, the migrations literature for this species does appear to be incomplete, and does not include some of the newer approaches for studying movements, such as satellite archival tagging. I will have more to say about this under research recommendations.

The uncertainties in the abundance indices include substantial interannual variation, which is more than expected for this relatively slow-growing species. It is unfortunate that the recommendation from SEDAR 21 for a power analysis of the ability of these surveys to detect changes in abundance apparently was not followed up. As noted in SEDAR 21, a power analysis would have allowed a determination of the adequacy of the survey effort, and could have led to increased or re-distributed resources in order to be able to evaluate the effectiveness of rebuilding strategies.

c) Are data applied properly within the assessment model?

The assessment is weak in this regard. As noted by the authors, the presence of large interannual variation in the abundance indices (inconsistent with the life history of sandbar sharks) may mean that the standardization process is not accounting for other significant factors that affect catch rates. The lack of concurrence among the indices also suggests to me that more work is required to assess the validity of the various series. For example, it is possible that a better understanding of the movement and migrations of sandbar shark would allow more informed choices of what abundance indices should be included in the population model.

The method for standardization of the abundance indices (delta log-normal/Poisson) is well accepted by fisheries scientists and appears to be correctly applied. However, as mentioned above, important factors that affect the catch rate series may still be unaccounted for in the standardization process.

I would have liked to see a model run that included only the two longest-running series (LPS and VALL), which are positively correlated. In my view, inclusion of many of the shorter time series, often with contradictory trends, is adding only noise to the reconstruction of the population.

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

For the reasons noted above, I don't think that the input series are sufficient to support the quantitative assessment approach and findings. Better resolution and interpretation of the conflicting CPUE series is required.

Also of note, most of the indices of abundance do not cover the period of rapid development of the fishery.

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

   a) *Are methods scientifically sound and robust?*

Inclusion of SS3 into the modeling work is a significant advance, and it represents the first time this modeling platform is used for this assessment, and only the second application for an elasmobranch species by NMFS scientists. The earlier approach (ASPM) used for sandbar sharks in SEDAR 21 had limitations recognized during earlier reviews, and SS3 allows for greater biological realism and more accurate depiction of uncertainties in the analyses.

Given that elasmobranch fisheries are often considered data-deficient, it is fortunate that sufficient data exist to employ the SS3 platform. However, as I explain elsewhere in my report, the real question concerns the quality of the input data.

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

The SEDAR 54 document frames the assessment as a standard analysis using the SS3 platform (p. 33). Given the description of the model formulation, I think this is a fair characterization. However, the choice of abundance indices to be included in the analysis is part of the configuration, and the decision to include all eleven indices in the base case is questionable. The various sensitivity analyses undertaken by the authors show that the choice of CPUE grouping profoundly influences the inference of stock status (Figure 3.2.13).

c) Are the methods appropriate for the available data?

No. Given the problems I perceive with the both the catch time series and the conflicting indices of abundance, I consider that they require resolution before application of the model. The choice of abundance index grouping as described profoundly influences the results of the assessment, and this reduces the utility of the assessment for management purposes.

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

The assessment team has done a thorough job in this regard. The replication and continuity analyses appear on P. 29 and 30 respectively in the report. Table 3.1.6 give evidence that the new modeling platform was able to closely approximate results from the ASPM runs for those parameters that can directly be compared.

3. *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?

The reconstruction of population numbers at age show some patterns which I find interesting. The bubble plot on Fig. 3.2.4 shows a "banding" pattern, with fewer fish surviving into older ages starting around the 1980 cohort. This, of course, corresponds with the period of higher exploitation. More recently, cohorts are starting to experience better survival (see, for example, the pattern with ages 10+). This could reflect the success of management interventions. Two small concerns with this figure: the 2016 numbers at age are truncated in my pdf version, and the red horizontal lines are not identified in the figure caption. I assume it represents the average age in the population?

To evaluate the consistency of the model outputs to the input data, I referred to Fig. 3.2.3. There are several issues in the CPUE fits, see in particular, S2, S5, S6, S7 where the predicted fits do not capture observed increases in the CPUE in the later years in the series. The period of initial decline is also not well captured by the model (see S1). Only two of the indices cover the period of relatively high exploitation: S1 LPS and S4 VALL, so the relatively poor fit of S1 is of particular concern.

As noted in the assessment, there is considerable sensitivity of the analyses to CPUE groupings. The authors appear to support the base case findings compared with the two alternative CPUE groupings, but I find rather limited support for this conclusion. I agree with the conclusion that the variation in estimates of spawning output depletion (SFF/SSF_0, Figure 3.2.12) are less variable for the base case compared with the results from the positive CPUE grouping. While this speaks to the precision of the estimates, there is no real support for the accuracy of the estimates.

My main issues with the assessment include the reconstruction of historic catches, and the adequacy of the abundance indices. Issues with the abundance indices include substantial unexplained interannual variation, conflicting trends of abundance indices, the absence of any index covering the more southern range of the distribution, and the absence of any index covering the most important remaining fishery. These are all very important deficiencies, in my view.

Considerations to the contrary include good model fits to the length composition data, and the reasonable reconstruction of the population numbers at age.

Overall, I consider that the existing population model and assessment results should be employed for general guidance on stock status only, taken together with other information such as size/age composition information for the catch and population. I use a "weight of evidence" approach to address the questions below, and do not consider the base case results to have the greatest credibility.

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

Yes. Without placing extra weight on the base case results, median estimates of SSF/SSF_{MSY} from six of nine alternative states of nature scenarios indicate that the stock remains overfished (estimated spawning output in 2015 relative to MSY < 1, Figure 3.2.13). It is also recognized there is considerable variation around the median estimates (Figure 3.2.14).

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

No. Without placing extra weight on the base case results, median estimates of F/F_{MSY} from six of nine alternative states of nature scenarios indicate that the stock remains overfished (fishing mortality rate in 2015 relative to MSY > 1, Figure 3.2.13). It is also recognized there is considerable variation around the median estimates (Figure 3.2.14). I also note that the population model shows a pattern of recovery of older-aged individuals, which would be inconsistent with overfishing (Figure 3.2.4).

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

No. I was not clear what was meant on P. 35 where the authors stated "Annual recruitment deviates from the recruitment relationship were estimated, but constrained to reflect the limited scope for compensation given the estimates of fecundity." I also understand that the reproductive biology of the species is not well understood, as a research recommendation contained in the assessment. The lack of information on reproductive biology probably negatively impacts the usefulness of the current stock recruitment relationship, and may limit its utility 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?

My response earlier applies here. Overall, I consider that the existing population model and quantitative assessment results should be employed for general guidance on stock status only, taken together with other information such as size/age composition information for the catch and population. I think using a "weight of evidence" approach may be the most appropriate way to use the results from this particular stock assessment. By this, I mean considering the results from the alternative states of nature scenarios equally, without attaching more weight to the base case formulation.

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

Yes, the model runs provided are well-described and reasonable. However, I would have liked to see an additional scenario that included just the two longest-running series (S1 LPS, and S 4 VALL), as they cover the period of relatively heavy exploitation. Some of the shorter series, I believe, may just be contributing noise to the results.

4. 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 projection software is done within the SS3 modeling framework, and is routinely used for numerous stock assessments. However, the reservations I have concerning the population reconstruction (historic catches, problems with abundance indices) also apply to the projections, since they impact the starting point for the projections.

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

Yes. As the projection software is designed as part of the overall SS3 modeling platform, the methods are appropriate for the population reconstruction and outputs. An advantage of the SS3 process, as noted in the report, is that the stock forecast routine is implemented after the variance estimation phase, so that certain aspects of parameter uncertainty calculated using the inverse Hessian method in the maximum likelihood estimation for the population reconstruction through 2015 are propagated into the variance of in forecasts of stock abundance under a chosen TAC.

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

An obvious caveat with the projection results is the very long time frames involved. For the base case, the projection is conducted until 2070, and for the alternative NEG CPUE grouping, the projection is run until 2111.

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

Recruitment variability was not included in the projections, but the authors concluded that given the reproductive biology of this species, variability in recruitment is expected to be low. This is reflected in Figure 3.2.9, that shows the unusually invariant and linear relationship between spawning stock and recruitment. The assessment document notes that based on the observation that the influence of the high and low productivity scenarios had minimal effect on stock status in comparison to the CPUE groupings, projections were only carried out for the base case productivity assumptions. However, I was not sure that the range of productivity scenarios developed for the assessment reflected the uncertainty in the range of reproductive output. The importance of this point appears to be supported by the assessment authors' view that research on this subject is a priority for future work (see p. 47).

I could not find documentation on what values were assumed for future weights at age in the projections.

The uncertainties in the model pertaining to CPUE grouping were also explored in the projection phase of the stock assessment. The conclusions on future harvest levels and rebuilding varied considerably depending on CPUE grouping, and changed the designation of whether the stock was experiencing overfishing (see below).

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

The assessment clearly indicates that under the base case, the stock was estimated to be overfished, but not experiencing overfishing ($F_{2015}/F_{MSY} < 1$). Therefore, as per the TORs, because there is no new or unexpected information about the status of the stock, no new rebuilding schedule was warranted, and projections were implemented consistent with the current rebuilding plan (started in 2005, projected to end in 2070) at a fixed level of removals (TAC on whole weight) allowing rebuilding of the stock by 2070 with 50% and 70% probability.

However, with the scenario using the NEG CPUE grouping, the stock was estimated to be overfished and experiencing overfishing, which triggered the provision of a new rebuilding schedule. Under this scenario, the target rebuilding year would be 2111, and the model was projected with a fixed TAC strategy that would attain rebuilding by the designated year with 50% and 70% probability. These TACs were 71 mt and 53 mt (whole weight), respectively. Compare these TACs with the recent five year average catch (=171 t, Table 2.3), and the improving age structure of the population (Figure 3.2.4). With that comparison, it would seem that the TAC levels associated with this scenario may be unnecessarily restrictive.

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*

The methods selected reflect and capture significant sources of uncertainty. The SS3 software is particularly useful for capturing statistical imprecision. However, the main sources of uncertainty associated with this assessment relate to the input data, as explained earlier.

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

The Discussion contains some insights into why the authors supported the model run that included all eleven CPUE indices as the base case. They note that: "the retrospective analysis found no systematic pattern of over- or under-estimation of abundance, relative abundance, or fishing mortality for the base case, which is as close as possible to the previous benchmark assessment base case configuration. The base model configuration, parameter values and input data are based on the best available information, and stock status results based on the base case run should thus be considered the most credible."

The logical inconsistency I find here is that as stated in Courtney (2017, WP06), there is literature that indicates that combining multiple conflicting indices (either explicitly or combined) into a stock assessment model is ill advised, and may result in biased parameter and uncertainty estimates. Yet the base case model does precisely that. No convincing argument is provided in the stock assessment document that addresses this apparent inconsistency.

6. *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.*

- There is a need for improved information concerning movements and migrations. The available indices cover only a fraction of the range of the sandbar sharks, and the absence of information concerning the connectivity of the population over the range is problematic. New tagging technologies, such as pop-up satellite archival tags, are now used routinely for other shark species, and would likely provide significant new information that would lead to a conceptual model for movements and migrations, comparable to what has been done for other large pelagic species, such as bluefin tuna.

- Much of the stock range is not covered by any index. In a changing ocean environment, this is a particular problem. Consider starting a new index of abundance further south and east, perhaps off Puerto Rico. The main fishery component (REC + MEX) in recent years is also not covered.

- Some of the recommendations from the SEDAR 21 report appear to have merit, but have not been dealt with. In particular, I am not sure that estimates of post-release survival have been improved since SEDAR 21.

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

As a desk reviewer, my exposure to the SEDAR process for this stock was limited. However, I would comment that the conclusions appeared transparent, and supported by the evidence. The process itself was well-documented.

One improvement in the process that might be helpful would be to list recommendations from the previous assessment, and note progress against the recommendations, or provide an explanation of why the recommendation was not followed.

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

- To the extent possible, it would be very helpful to review and evaluate the highly uncertain Mexican removals, including the assumptions concerning historic species composition.

- It was interesting and encouraging to note that there has been “cross-fertilization” of ideas for handling data and analyses from other organizations involved with stock assessment, notably the International Commission for the Conservation of Atlantic Tunas (ICCAT). This is apparent in WP 06, which uses an ICCAT approach for a hierarchical cluster analysis and cross-correlations of selected CPUE indices for the SEDAR 54 assessment was conducted to identify conflicting information among CPUE indices.

  On the other hand, it would have been useful to see other best practices from ICCAT followed for this assessment. In particular, I was thinking of the recent work of ICCAT for evaluation and comparison of abundance indices. Given the plethora of indices available for this assessment, having some organized way to compare them would be very helpful. Some possible approaches follow below.

While much of the process refers to abundance indices from commercial fishery catch rate series, most of the steps could also apply to fishery independent surveys as well.

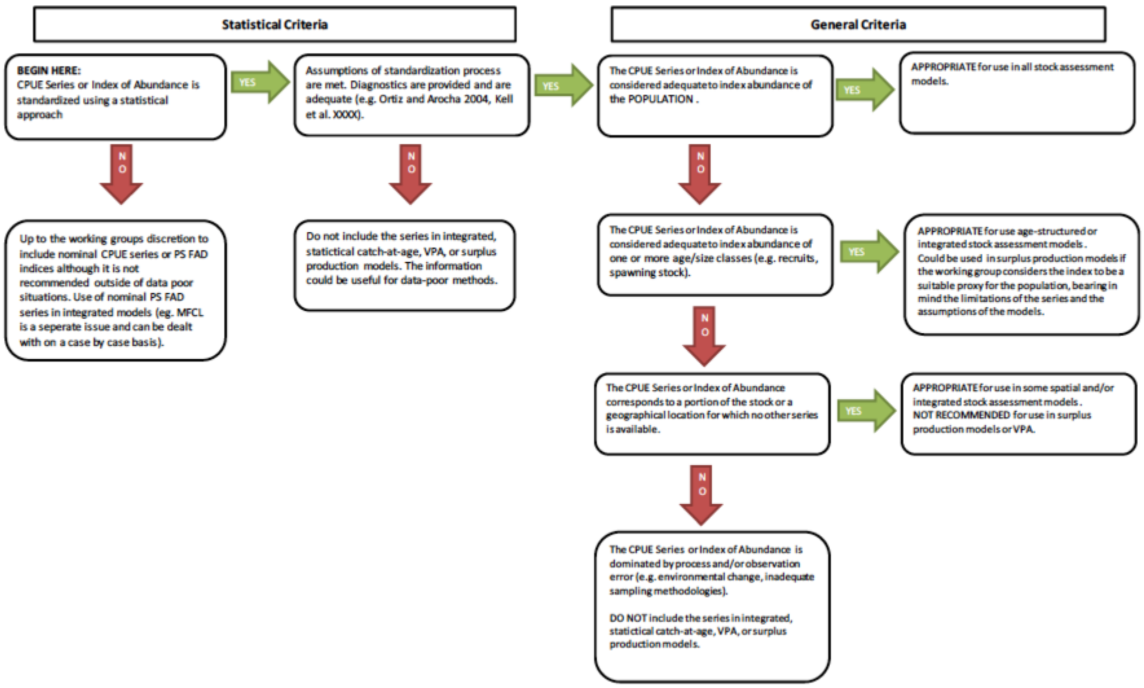

Figure 2. Flow chart for the evaluation of potential indices of abundance (ICCAT 2013).

Another tool that can be used to evaluate the sufficiency and quality of the proposed abundance indices is provided below:

ELEMENT	*DESCRIPTION*	*SUFFICIENCY SCORE (1 is poor, 5 is best)*				
		1	*2*	*3*	*4*	*5*
1	Diagnostics	No diagnostics or assumptions clearly violated				Full diagnostics and assumptions fully met.
2	Appropriateness of data exclusions and classifications (e.g., to identify targeted trips).	Not appropriate				Fully appropriate
3	Geographical coverage	Small localized fishery/survey				Represents geographic range of population
4	Catch fraction	Small				Large
5	Length of time series relative to the history of exploitation.	Short				Long
6	Are other indices available for the same time period?	Many				It is the only available index
7	Does the index standardization account for known factors that influence catchability/selectivity?	No				Fully
8	Are there conflicts between the catch history and the CPUE response?	Yes				No
9	Is the interannual variability outside biologically plausible bounds (e.g., SCRS/2012/039)	Frequently				Seldom
10	Are biologically implausible interannual deviations severe? (e.g., SCRS/2012/039)	Very severe				Minimal
11	Assessment of data quality and adequacy of data for standardization purposes (e.g., sampling design, sample size, factors considered)	Low				High
12	Is this CPUE time series continuous?	Very discontinuous				Completely

Figure 3. Template used to grade the quality of proposed abundance indices (ICCAT 2013).

- A recommendation from SEDAR 21 was to conduct a power analysis for the available abundance indices, noting the large interannual fluctuations that are unaccounted for with the existing approaches for standardization. I would reiterate the recommendation, as I feel it is particularly important.

- I would recommend including a scenario that includes only indices of abundance that cover the period of relatively high exploitation (LPS and VA_LL).

Conclusions and Recommendations

There are many aspects of this stock assessment that are logical and well executed. However, there are several important concerns. One of the most striking is the observation that combining multiple indices of abundance is not advised when the information contains conflicting trends. WP-06 makes this recommendation, and refers to the work of Schnute and Hilborn (1993) to substantiate the point. Schnute and Hilborn (1993) showed that when conflicting indices are included in a stock assessment model, the most likely parameter values are usually not intermediate but occur at one of the apparent extremes. Including conflicting indices in a stock assessment scenario may also result in residuals not being identically and independently distributed (IID) and so procedures such as the bootstrap cannot be used to estimate parameter uncertainty. Yet, the assessment continues to include eleven indices of abundance in the base case that contain contradictory trends that are not reconciled. The stock assessment fails to provide convincing arguments that the concerns of Schnute and Hilborn (1993) should be disregarded. Other important concerns with the indices of abundance include poor spatial and temporal coverage, as well as large interannual variations that are not accounted for in the catch rate standardizations.

I also have reservations about the adequacy of the historic catch information, particularly the REC+MEX component that has some important underlying assumptions. It is noteworthy that the REC+MEX component now comprises the largest fraction of the total catch.

Taking these concerns together, I am not convinced that the assessment provides a robust basis for quantitative fisheries management advice. I reiterate my earlier points that the existing population model and assessment results should be employed for general guidance on stock status only, taken together with other information such as size/age composition information for the catch and population.

I recommend that greater efforts be made towards reconciling the differences in the available abundance indices, as well as validating the historic catch data to the extent possible.

Literature Cited

ICCAT (International Commission for the Conservation of Atlantic Tunas) 2013. Report of the 2012 Meeting of the ICCAT Working Group on Stock Assessment Methods. Collect. Vol. Sci. Pap. ICCAT, 69(3): 1354-1426 (2013).

Schnute J.T., and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 50 (9): 1916-1923.

Appendix 1: Bibliography of materials provided for review

Carlson, J.K. and Alyssa N. Mathers. 2017. Updated catch rates of sandbar sharks (*Carcharhinus plumbeus*) in the northwest Atlantic Ocean from the Shark Bottom Longline Observer Program, 1994-2015 SEDAR54-WP-02. SEDAR, North Charleston, SC. 10 pp.

Cortés, E. and Xinsheng Zhang. 2017. Standardized catch rates for sandbar sharks from the U.S. pelagic longline observer program using generalized linear mixed models. SEDAR54-WP-05. SEDAR, North Charleston, SC. 13 pp.

Courtney, D. 2017. Example Implementation of a Hierarchical Cluster Analysis and Crosscorrelations of Selected CPUE Indices for the SEDAR 54 Assessment. SEDAR54-WP-06.SEDAR, North Charleston, SC. 11 pp.

Driggers III, W.B., Bryan S. Frazier, John K. Carlson, Bethany M. Deacy, Michael P. Enzenauer and Andrew N. Piercy. Updated life history parameters for sandbar sharks, Carcharhinus plumbeus. SEDAR54-WP-01. SEDAR, North Charleston, SC. 11 pp.

Pollack, Adam G., David S. Hanisko and G. Walter Ingram, Jr.. 2017. Sandbar Shark Abundance Indices from NMFS Bottom Longline Surveys in the Northern Gulf of Mexico. SEDAR54-WP-04. SEDAR, North Charleston, SC. 19 pp.

SEDAR 21 Stock Assessment Report. HMS Sandbar Shark, September 2011. 459 p.

SEDAR 54 Stock Assessment Report. HMS Sandbar Shark, October 2017. 179 p.

Walter, J. and Craig A. Brown. 2017. Standardized catch rates of sandbar sharks from the Large Pelagics Rod and Reel Survey 1986-2015. SEDAR54-WP-03. SEDAR, North Charleston, SC. 9 pp.

Appendix 2: 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.
   e) Are methods scientifically sound and robust?
   f) Is the assessment model configured properly and used consistent with standard practices?
   g) Are the methods appropriate for the available data?
   h) Are differences between the current model and the previous model clearly documented and described?
4. Evaluate the assessment findings with respect to the following:
   f) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   g) Is the stock overfished? What information helps you reach this conclusion?
   h) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   i) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   j) 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:
   b) Are the methods consistent with accepted practices and available data?
   c) Are the methods appropriate for the assessment model and outputs?
   d) Are the results informative and robust, and useful to support inferences of probable future conditions?
   e) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
   f) 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.
    - 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.
8. Provide guidance on key improvements in data or modeling approaches that could be considered when scheduling the next assessment.