Independent Peer Review of SEDAR 50: Atlantic Blueline Tilefish Assessment

Center for Independent Experts (CIE) Independent Peer Review Report

Paul A Medley
Email: paulahmedley@yahoo.co.uk

1 Executive Summary

- SEDAR 50 covers the stock assessment and Center for Independent Experts (CIE) review for U.S. South and Mid Atlantic blueline tilefish stocks. This is CIE reviewer Paul Medley's report for the SEDAR 50 review workshop. The SEDAR process is ultimately responsible for ensuring that the best possible assessment is provided. This independent report is complimentary to the Review Panel's report.
- The data and models were appropriate and sufficient to support the stock assessment and scientific advice to the management councils. The final stock assessment results given in the Assessment Workshop report and addendum, and the Review Workshop report provide the best scientific advice for this stock that is currently available.
- The available evidence indicated that the stock South of Cape Hatteras was not likely to be overfished or undergoing overfishing, despite limitations in the data. Catches in 2013 and 1980-85 may have exceeded MSY. However, the stock assessment is dominated by uncertainty, such as uncertain catches and missing abundance indices in important years.
- It was not possible to assess the stock status for the stock North of Cape Hatteras and so no projections were carried out. However, available information and analysis suggests overfishing may be occurring, at least in the area currently exploited. If appropriate limits are not placed on the fishery, there is a significant chance that the current fishing area will be locally depleted.
- The main uncertainties are due to the missing information (stock structure, ages, accurate historical catch, and recent abundance indices), and uncertainty is a dominating feature of these assessments.
- Before conducting another assessment of this species, the problems with ageing will need to be resolved. In addition, I strongly recommend conducting a longline deep water survey covering blueline tilefish and other species caught in these fisheries, since the lack of recent abundance indices will hamper ongoing monitoring and assessment. I recommend continuing to develop fully Bayesian approaches to modelling, particularly for this type of fishery with missing information.

2 Background

This is one of three independent reports that describes the findings and conclusions of the review workshop for the SEDAR 50 U.S. Atlantic blueline tilefish in accordance with the CIE statement of work (Appendix 2). The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise to conduct independent peer reviews of NMFS scientific projects without conflicts of interest. Each reviewer contracted by the Center for Independent Experts provides an independent peer review report to be approved by the Center for Independent Experts Steering Committee (www.ciereviews.org). In addition to the three independent CIE reports, there is a Review Panel summary report which represents the consensus view for all five panel members.

SEDAR 50 covered the stock assessment and CIE review for U.S. Atlantic blueline tilefish. SEDAR 50 consisted of a series of workshops and stock assessment tasks culminating in submission of a stock assessment for the SEDAR 50 Review Panel. A SEDAR 50 Stock ID Work Group meeting was held 28th-30th June 2016 in Raleigh, North Carolina, followed by Stock ID Joint SSC Sub-panel review webinar in October 2016 and a Stock ID Management and Science Leadership conference call in November 2016. A Data Workshop (DW) was conducted 23rd – 27th January 2017 in Charleston, SC, with additional webinars in November and December 2016 and February 2017 to finalize the data

input to the assessment. An Assessment Workshop (AW) was held 23rd-26th May 2017 in Atlantic Beach, with four supporting assessment webinars April-July 2017, two before and two after the AW. These Workshops produced a report and other material for the Review Workshop (RW), which took place 29th-31st August 2017 in Atlantic Beach, NC. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stock assessed through SEDAR 50 is within the jurisdiction of the Mid-Atlantic and South Atlantic Fishery Management Councils.

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

I conducted an independent peer review in accordance with the Statement of Work (Appendix 2) and Review Panel Terms of Reference.

Eighteen days before the review workshop, the NMFS Project Contact made available the necessary background information and reports for the peer review, which I used to prepare for the review. A one hour pre-planning Webinar took place on Friday 18th August 2017 at 7:00pm (BDT) ensuring all necessary information and giving the opportunity for specific requests to the reviewers. I requested the data input to the stock assessments and an additional presentation on the ageing problem, both of which were subsequently provided. I attended the review workshop (RW), which took place Tuesday 28th August to Thursday 31st August 2017 in Atlantic Beach, NC. Results were presented at the RW of work on the stock identity and ageing errors as well as the stock assessment work. During the workshop, the Review Panel requested some additional analyses, which included changes to several of the base operating models (stock, fleet, and observation components) and also additional sensitivity analyses. The results of those analyses are presented in an addendum to the main assessment report.

As a member of the Review Panel, I contributed to a separate summary report of the Review Workshop, which was a consensus report of the entire panel. In accordance with the Statement of Work, the findings, conclusions and recommendations of this independent review report do not necessarily represent the views of other members of the Review Panel.

4 ToR 1: Evaluate the data used in the assessment, addressing the following:

a. Are data decisions made by the DW and AW sound and robust?

Most decisions made were sound and robust, but some decisions require comment. The key decisions made by the DW and AW were on the stock structure, the rejection of the age data, and the decision to support the ASPIC model.

The decision on the stock structure (independent assessments for the Atlantic stocks North and South of Cape Hatteras) was not supported by the scientific evidence, as fully acknowledged by both the DW and AW. It is highly likely that stocks are linked through recruitment, but the degree to which they might be linked is unclear. Conversely, the adult populations are likely to have more limited ranges, so discrete populations may form on a smaller spatial scale than used for this assessment. However, I agree that there was insufficient information to support an alternative assessment stock structure to that which was done. The decisions made by the review panel were likely robust to the stock structure assumption, given there was no explicit test of the assumption in a sensitivity analysis.

I support the rejection of the age data for this assessment. The lack of age data degrades the accuracy of the assessment, but the ageing errors meant no reliable length-age relationship could be established. It might have been possible to use these data more to inform on the decision on the growth model, as had been done with SEDAR 32 assessment of these stocks, but this would have led

to a loss of confidence in the results. Before any further stock assessment work is undertaken on this stock, the ageing problem needs to be solved.

I did not support the choice of the ASPIC model as the base model for the scientific advice. The decision made by the AW was based on a lack of information on life history. This decision was understandable, but in principle, I believe was not the best decision.

The logistic model which underpins ASPIC is highly informative by nature of its structure. This can give the impression that its estimates are more precise than they are because of the structural limits on the model's behaviour. In particular, the logistic population growth model in ASPIC assumes that the population is closed (no immigration or emigration) and that population growth (individual fish growth and recruitment) is controlled by a single parameter (r). While such a model can often fit past single depletion-recovery events quite well, in my experience this does not lead to an ability to predict future status. This is because stock changes cannot be so easily explained as they are determined by highly stochastic events, such as strong or weak year classes.

In its favour, the estimate of MSY the logistic model provides tends to be precautionary when used in the right way. Note that the MSY estimate (50% B_0) is for the exploitable biomass, not SSB, so the SSB will be less than this. For short term projections, the model is accurate enough to provide advice. In the absence of data other than abundance indices and total catches (removals), this model is appropriate.

The fact that the ASPIC model is likely to be precautionary is in its favour, and would justify its use if no length composition data were available. However, it was not appropriate to ignore the length composition data as it was possible to make use them. Also, it is highly likely that the stock status bounds estimated by the ASPIC model would not include the stock status estimated in future by an age structured model. Such large and significant changes in advice undermine the scientific process. Stock assessments should lead to clear reductions in uncertainty, so data poor methods should give ranges in status, which it is highly likely a stock will be within. As data and methods improve, the assessment should narrow down within this range rather than move outside of it.

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

With a notable exception, data uncertainties were acknowledged and reported. Full explanations were provided for critical issues related to stock structure and ageing errors. Uncertainty with stock structure was not unusual. However, the problem with ageing was clearly unexpected and needs to be resolved.

Uncertainty in historical catches South of Cape Hatteras were not acknowledged or reported in the original assessment. The assessment models used (production models) are very reliant on total catch data to scale biomass, so it was found that results were very sensitive to alternative catch history scenarios. The uncertainty in past catch was raised by a stakeholder during the RW. The reason it was not raised earlier was, perhaps, because the unexpected reversion back to simpler production models left the new assessment more sensitive to this issue than it might otherwise have been.

Recreational catches were also highly uncertain, but have a much lower impact because the quantities are lower. However, it is already well-known that it is difficult to obtain accurate catch estimates for recreational fisheries and uncertainty was reported and acknowledged.

c. Are data applied properly within the assessment model?

The fit and diagnostics confirmed that the data were used correctly within all the assessment models after some changes were made as recommended during the RW. For the stock South of Cape Hatteras, abundance indices were fitted in a single ASPIC model with equal weighting, whereas all data, including the length compositions, were used in the ASPM assessment, which the RW preferred.

There were some gaps in the documentation for the CPUE index standardisation, so it was unclear

what process had been applied in developing the standardisation models. The documentation appeared to have been primarily produced for the DW rather than for external review. It might be expected that the same procedures could be scripted (e.g. in RMarkdown/RStudio: see Grolemund and Wickham 2016) that would make more complete documentation possible for all species caught in these fisheries as the analyses would be very similar. However, the information that was presented did not indicate any problems with the standardisation, and the resulting indices were smoother than the nominal indices. So, it was likely that they were better representations of changes in abundance.

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

For South of Cape Hatteras, the data were sufficient for the age structured model used. It appears that reported data was sufficiently reliable, whereas the main problem was missing data. In some cases, input data to the assessment are actually themselves estimates (e.g. early removals and abundance indices). Whether growth parameters and recruitment deviations could be estimated was checked through sensitivity analyses. This indicates whether there was sufficient support for these parameters in the data. There was sufficient information for growth parameters, but not recruitment deviations. The assessment approach and reported findings match this level of information.

For North of Cape Hatteras, the data were insufficient to conduct a standard stock assessment. The available information used was appropriate for the DLMtools software. This is appropriately reflected in the limited findings and advice.

5 ToR 2: Evaluate the methods used to assess the stock, taking into account the available data.

a. Are methods scientifically sound and robust? Do the methods follow accepted scientific practices?

All the methods applied for both North and South of Cape Hatteras were scientifically sound and robust and followed accepted scientific practices. ASPIC and DLMtools methods are taken from the NMFS toolbox. The BAM software implemented in ADMB has a long history of successful use, including for golden tilefish. The software has therefore been well tested. ASPIC and BAM (used for ASPM) apply good sound techniques for stock assessment. DLMtools, used for the stock North of Cape Hatteras is also well tested. DLMtools is not a stock assessment method, but a management strategy evaluation tool, but applies scientifically sound procedures and provides robust advice.

b. Are assessment models configured appropriately and applied consistent with accepted scientific practices?

All models applied were configured correctly and applied within accepted scientific practices. This was demonstrated by the diagnostics as well as the documentation.

However, the approach of averaging models for the separate abundance indices used by the ASPIC assessments was not endorsed by the RW. A mixed model approach can be used under specific circumstances (alternative Bayes models fitted to the same data), but one needs to ensure that the posteriors are normalised correctly and are compatible. In this case, there is no theory to back up averaging as carried out by the AW. Again, this could be a Bayesian approach if *a priori* both models were considered equally likely, but if this was the case, then the assessment should have been more explicit concerning assumptions. The RW recommended the simpler solution of weighting the indices effectively equally within a single fit, which made more sense as the data rather than the model was changed.

The original fit used the parameter standard errors from the standardisation to weight the index

points. This standard error only accounts for the difference between the observed CPUE and the mean CPUE. There is an additional error accounting for the difference between the mean CPUE and the expected abundance index value (q * B), which was not estimated. This may particularly be a problem with more than one abundance index since the standardisation only measures goodness of fit to the standardising covariates and factors, not the underlying population trends. Unless there is an argument or evidence to the contrary, the best default is to assume equal weighting of different abundance indices.

c. Are the methods appropriate for the available data?

Both ASPIC and BAM were appropriate for the catch and abundance index data, but the BAM model was able to use the length composition data, making it more appropriate for the assessment. The age-structured model also was able to capture uncertainty better and will provide a better understanding of the fishery in the longer term.

ASPIC only requires total catch and abundance indices. However, it does not describe population dynamics processes, but only a population's response to depletion. This provides a fairly crude description of population behaviour that can seem to fit quite well to past data, but provide poor description of future behaviour because the estimated parameters can change over time. Periods of low or high recruitment or growth can mislead the assessment. This is not to say ASPIC is not useful if it is used in a precautionary way, as it provides useful advice for short term decisions. ASPIC may also give a false impression of certainty because the model structure constrains population behaviour that may not be supported by the available information.

5.1 North of Cape Hatteras.

The DLMtools methods are scientifically sound and robust, and follow accepted scientific practices. The models were configured appropriately and applied consistently with accepted scientific practices. The purpose of the DLMtools is primarily to conduct simplified management strategy evaluations on stocks with little information to support stock assessments. However, in this case, the objective was to evaluate a recent increase in catches relative to likely long term catches that meet the US National Standard 1 Guidelines. The lack of data, notably any abundance index, makes this the best option for this fishery.

6 ToR 3: Evaluate the assessment findings with respect to the following:

a. Are population estimates (model output – e.g. abundance, exploitation, biomass) reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

The model output was consistent with the input data and the biological characteristics. The dominant characteristic of this (and other) data poor assessments is the uncertainty. In absolute terms, estimates may not be reliable and could change significantly with new and better data. Absolute status cannot be estimated for the stock North of Cape Hatteras. For South of Cape Hatteras, stock size estimates vary widely. However, relative estimates (e.g. B/B_{MSY}) are sufficiently reliable to be used for setting precautionary catch limits.

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

The stock South of Cape Hatteras was likely not overfished. The assessment provides quantitative evidence to support this conclusion, but there is also some qualitative information. With respect to the quantitative information, point estimates should be considered with great care. The bootstrap estimates will underestimate uncertainty because they only account for observation error. Other sources of error (process and structural) are usually much greater. In this case, the sensitivity range is a better indicator of the true uncertainty for this stock.

The abundance indices only cover the period well after the initial depletion and two years into the recovery. Confirmation over whether the production models are properly determining stock size would require the abundance indices to cover the periods of highest contrast in the removals, which is not the case here.

There is no direct quantitative information on stock status North of Cape Hatteras. The assessment was primarily undertaken to determine whether overfishing was occurring or not. There are no abundance indices for north of Cape Hatteras. The DW noted that there is increased targeting of blueline tilefish, so any increased catch rate is probably not tracking abundance.

Any determination of stock status for both North and South of Cape Hatteras should probably be accounting for stock structure. It is likely that recruitment links a number of sub-populations. All these sub-populations would probably need to be depleted for the stock to become overfished overall. Given that blueline tilefish is mostly caught as bycatch, the population overall is probably more robust than would otherwise be the case. However, this is only weak evidence without additional supporting data, such as from a deep water survey.

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

For the stock South of Cape Hatteras, all the ASPM runs and majority of ASPIC runs indicated that recent fishing mortality is likely to be less than F_{MSY}. For some ASPIC runs fishing mortality was estimated to be at or just over F_{MSY}. Removals have been declining throughout the 1990s and 2000s, and the abundance indices during the mid-period seem to show an increasing trend. However, the catches have increased in the last few years and were high in 2013, so in this particular year it is more likely overfishing did occur, but the other years should give an indication of safe catch limits.

For the stock North of Cape Hatteras there is possible local depletion. Based on the DLMtools simulation work, current catches could easily exceed sustainable levels while the local population is depleted. This cannot be confirmed without more information however, so it can only be stated as a risk.

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

There is no information on a stock recruitment relationship for any of these stocks and the stock recruitment relationship is not really estimated. The productivity, as indicated by average recruitment and growth, is estimated, but the response of recruitment to spawning stock depletion, steepness, is not. Steepness is primarily used to estimate reference points.

Reasonable assumptions are made for the ASPM model and mean recruitment (a measure of productivity) was estimated. The stock recruitment relationship in the ASPIC model is not really estimated, but assumed within the model structure. A stock recruitment relationship would only be estimable with good ageing data, and most importantly, a better understanding of the stock structure. For these assessments, default spawning stock reference points ($F_{30\%}$, $F_{40\%}$) are suitable for estimating stock status.

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

The quantitative point estimates are not precise, but can be used by management taking into account the uncertainty.

For South of Cape Hatteras, the population estimates across the sensitivity runs are reliable enough to support status inferences. They are consistent with available information and data. However, I would place more emphasis on the risks. My personal preferred run, based on information at the meeting, was the ASPM 0.1 catch 1980-1986 ("S09 0.1HLRemo80to85"). This is a worst-case scenario, however, as it makes sense that increased catches of blueline tilefish could have occurred

with increased targeting of golden tilefish.

Production models are particularly sensitive to the total removals. Overestimating catches cause the fit to create imaginary biomass to cover this loss from the population. The imaginary biomass adds to the potential yield of the fishery, and could lead to over-optimistic, biased advice.

I suggest precautionary catch limits are set, taking into account run "S09 0.1HLRemo80to85" (Report Addendum), which considers blueline tilefish catches are lower than those reported during 1980-85. As better data, particularly the age data, become available the model should become less reliant on these past catches. The model will be able to estimate an initial state or the higher catch may be partially explained by recruitment deviates rather than steady-state biomass. This last will add to the recruitment variance, increasing uncertainty in projections which is perhaps a better representation of this potential error in the catches. If the catches have been incorrectly recorded, this would still be incorrect, but the implications would be less severe for the assessment results.

The ASPIC model gave more precautionary advice than ASPM, which given that much data are missing, may seem a better model to use for providing advice. The ASPIC was the original preferred model by the AW. However, there is sufficient information to suggest the ASPM is useful for management decision making. Specifically, the species appears to mature before being caught, which should make it relatively resilient. While the ASPM advice does make sense, any management action should be cautious.

For the stock North of Cape Hatteras, DLMtools only provide an indication of whether overfishing might be occurring. Information is sufficient to guide appropriate adjustments to current catch limits to reduce risks of overfishing. However, it is not possible to determine stock status without additional data.

7 ToR 4. Evaluate the stock projections, addressing the following:

7.1 Stock South of Cape Hatteras

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

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

The approach used for projections was appropriate for the models and available data and represents good practice. Projections are based on the stock assessment models, so estimated parameters from the bootstrap simulations can be used to project the populations forward under different management scenarios. This is standard practice and suitable for the data and models used.

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

The short term projection results are informative and reasonably robust. The results are described in relation to the most important indicators and the analytical team demonstrated that parameter estimates are robust to uncertainty.

The short term projections (<5 years) are useful for indicating probable future conditions. Longer term projections within the life span of the species require recruitment deviations to be estimated, which would require age data.

The estimated F_{MSY} should not be used as a target reference point, as the estimate is not precautionary and not reliable. It would imply that depleting the stock to less than 20% of the unexploited state would be safe, and clearly it would not. In my opinion, there is insufficient information on this stock to apply higher target fishing mortalities than $F_{40\%}$.

It is worth noting that the abundance indices are truncated and stop in 2006/7. Without an abundance index, the stock assessment becomes increasingly unreliable particularly for projections.

Some sort of confirmation that the model is tracking abundance would be most valuable, otherwise the model could have undetectable errors. Such a confirmation could come from valid catch-effort data from the commercial fishery or a scientific survey.

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

The key uncertainties are the structural uncertainties covered by the sensitivity analyses and the observation errors covered by the bootstrap estimates. When completed, the bootstrap simulations should provide a reasonable estimate of uncertainty caused by observation error. However, I would expect the sensitivity analyses to bracket the uncertainty. Specifically, the alternative catch history with 10% of the previous tilefish catch 1981-85 allocated to blueline tilefish probably represents the worst case scenario. Process errors, primarily present in recruitment variation, were not estimated and therefore not present in the projections. This was made clear and limited the valid projection length.

7.2 Stock North of Cape Hatteras

No projections were run for the fishery North of Cape Hatteras. Given the available data and methods, it was not appropriate to carry out projections for this fishery. The DLMtools used implement simulations to test management procedures, not provide any estimate of stock status.

8 ToR 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 all sources of uncertainty in the population, data sources, and assessment methods

b. Are the implications of uncertainty in technical conclusions clearly stated?

The methods capture uncertainty in the estimates of observation errors through bootstraps, and structural errors were explored through sensitivity analyses. Process errors (intrinsic random variation in population dynamics most apparent in recruitment and mortality) are not modelled explicitly, but implicitly accounted for in the observation and structural errors. The structural errors are most important and lead to greatest changes in the stock status indicators.

The sensitivity analyses were used to address likely errors in life history parameters and data sources. As far as possible, the RW reviewed all likely sources of uncertainty and these are now reported. The implications of uncertainty are clearly stated in the AW report and addendum and the information provided should allow decision-makers to apply an appropriate level of precaution.

The focus of the DLMtools used for the stock North of Cape Hatteras is dealing with uncertainty. It is captured primarily through parametric probability density functions defined for model parameters and simulated observations, but also by the different models and management procedures that can be applied. The parameters used were generally accepted as being reasonable estimates through expert consensus (DW and AW). The implications of uncertainty are clearly stated and captured in the reported performance indicators. In my opinion, this sort of simulation / MSE approach is much better than trying to develop new data poor stock assessment methods that tend to hide the uncertainty.

9 ToR 6. Consider the research recommendations provided by the Data and Assessment workshops 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.

The key problem with research recommendations is that they become wish-lists. All the recommendations made by the AW and DW are useful and would lead to improved stock assessments, but lack focus. Any future assessment of blueline tilefish requires two actions:

1. The problems with aging this species need to be resolved so that the historical age data are available for analysis.
2. Abundance indices are required that extend to the current year. Given the problems with the fishery dependent indices, I would strongly recommend a deep water survey is undertaken every 1-3 years. A survey should be cost effective because not only can it be used to address research regarding blueline tilefish, it can also be used to provide information for the full range of deep water species.

Data obtained from a fishery independent survey would also provide data that could be used to address a number of research recommendations made by the AW and DW:

- Improve estimation of selectivity using a range of hook sizes.
- Provide a better understanding of the spatial distribution of species by size and age.
- Provide information on the maturity and reproductive biology of species. Sampling commercial catches may still be required to identify spawning seasons.

The Management Councils appear to be applying controls without regard to the impact on information requirements. This undermines the management process. Either decisions on management controls applied to the fisheries need to take account of their impact on the abundance indices, measuring catch, selectivity and so on, or the required information needs to come from alternative sources, such as a fishery-independent deep water survey.

10 ToR 7. Provide suggestions on improvements in data or modeling approaches, which should be considered when scheduling the next assessment.

1.1 Data

Clearly, of immediate importance is to develop a consistent ageing protocol for blueline tilefish so that acceptable precision on ages is obtained. This was already recognised by the DW and AW. It will be necessary to do this to ensure the life history parameters can be estimated, as well as be able to use the age data in an age-structured stock assessment.

A deepwater survey would appear to be highly cost effective as it would provide much needed information on a wide range of stocks, including blueline tilefish. Even a single survey could answer some important questions on the spatial distribution of the adult stock, selectivity and the location of juveniles. Given the value of such a survey across a range of stock assessments, it would seem prudent to carry it out.

It may be worth reviewing and correcting catches for all tile fish. In many cases, values are provided to assessments as though they are observations when they are strictly estimates. No information was provided to the RW on how these catches were estimated. It would make sense to see whether a range of realistic catches might be developed for each tilefish species. Although with improved age data, the models should become less sensitive to historical catch, these should still have some influence, particularly on estimated reference points.

1.2 Modeling

The analytical team should explore more formal Bayesian approaches to assessment, particularly where data are limited. The vast majority of fisheries require subjectivity and expert judgement in completing a stock assessment. Bayesian assessments make these decisions explicit and they can be checked and endorsed through the review process. Bayesian statistics also provide a consistent and complete way to deal with missing data, latent variables and prior information (e.g. meta-analyses) as well as subjective information. The main problem with Bayes is having efficient and reliable numerical methods. Techniques are being developed and software, such as Stan (mc-stan.org, Stan 2016), have made significant progress.

Fitting the models using MCMC (as Stan does) would be preferable to the bootstrapping procedure used in these assessments. However, while bootstrapping has theoretical drawbacks, it can be completed relatively easily and within a guaranteed time frame. MCMC is more complex and often getting convergence is difficult, depending on the software used. Nevertheless, I would recommend using MCMC (Stan) where possible, as it is a much more powerful technique for dealing not only with observation error, but process error, latent variables and missing data.

For data poor fisheries, it may be well worth conducting more extensive management strategy evaluations using DLMtools. This would require greater involvement of the fishery management bodies as the aim is a little different to stock assessment. The objective is to develop and test the management procedures that might be applied regardless of the state of nature. Defensible procedures might be developed, which can be shown to be robust across the likely range of status and life history for such stocks, which would make better use of DLMtools. The Delphi method used in estimating blueline tilefish recreational catches illustrates the type of approach that can be used to set ranges on states of nature. However, fishery managers would need to be more involved as the process is assessing decision rules, not the stock.

The theory for the delta lognormal is not strong, although it can account for over-inflated zeros well, and empirically often provides a good fit to many data. However, it may be worth considering alternative models for overdispersion which constrain the trips with zero catch to some theoretical quantity linked to the positive trips and abundance. My preference would always be for alternative models if they can be found (see below).

Another option is to exclude zero catch trips altogether. Since the binomial component of the delta-lognormal model should be aligned with the log-normal component, this should not lead to significant changes in any derived index. The binomial part of the delta lognormal explains zero catch, but if trips with zero catch are over-represented, the results could still be biased. Often it is difficult to tell whether the trip selection criteria are correct. Therefore, if the number of positive trips is sufficient, a safer option is only to use the lognormal part.

An alternative to the lognormal is the Gamma distribution, which may be more robust than the log-normal (McCullagh and Nelder 1989). The Gamma was used in exploratory fits for the current standardisation models. I would recommend fitting the catch rather than CPUE as the response variable, with a log link function and log-effort as the offset. This maintains the error structure in a more appropriate form than fitting CPUE. Having log-effort as an offset with the log-catch response variable makes no difference to fitting log-CPUE, as the variance is scale independent for the logged variable. It does make a difference for the Gamma likelihood. In any case, having effort as a term in the linear predictor also allows the model to fit a parameter to estimate the “diminishing returns” of effort in a trip. This may not be suitable for final standardization model, but could be useful to check whether the measure of effort is appropriate. Given concerns raised by fishermen at the meeting over the effect of bait loss, this may be a useful exercise.

Commercial longline in particular should be able to provide reasonable indices of abundance. However, all indices of abundance were unreliable or truncated. This presents a real problem for

future stock assessments for this species. Without an abundance index going forward, the stock assessments will become increasing unreliable. Either the fishery management bodies will need to ensure the commercial fisheries provide data for valid abundance indices or a scientific survey will need to be conducted.

A deep water longline survey could provide a much better abundance index than commercial data because much more information could be recorded. The number and weight of fish caught on a longline (and perhaps less clearly on handline) is related to the number of hooks, fish density, hook fishing time and the density of other species. Constant catch rates and constant bait loss rates leads to a simple exponential model. The proportion of each species and empty hooks making up the catch for the longline set is proportional to these rates. Fishermen at the RW reported that bait loss was very rapid and an important factor in determining catch rates. Catches will show overdispersion because fish aggregate. An appropriate likelihood model to try with these data would be the Dirichlet multinomial, which includes an additional parameter for the variance. If this works for the survey, it may provide the basis for developing a better model for the commercial longliners.

1.3 SEDAR Process

SEDAR is a rigorous review process and is very demanding for all participants. Overall, the assessments were well presented and explained. Presentation of information was good, particularly the summary information. The overall result was increased confidence in the stock assessment and scientific advice.

Stakeholders were invited to provide information to the review meeting and representative fishermen from the US South Atlantic were present. This was particularly helpful in this case as they provided valuable information to the Review Panel (RP) and the RP was able to respond and act on the information given. Although one of the fishermen at the RW said they had suggested a sensitivity run with an alternative catch history, somehow this suggestion had been lost during the DW and AW, probably because of significant changes to the planned stock assessment. There was an opportunity to raise this issue again, and it turned out the assessment was very sensitive to this issue. This was a direct result from the stakeholder involvement in the process and should been seen as a SEDAR success.

My suggestions for improvements would be:

- If the process is to make good use of DLMtools, the assessment team will need a set of management procedures that can be tested. This suggests that the Management Councils will need to identify an acceptable parameterised set of procedures for each fishery taking into account mixed gear and species fisheries, and acceptable controls.
- SEDAR should consider developing a process that fully documents assessments using a scripting markup language in RStudio (see Grolemund and Wickham 2016). This would make the assessment fully reproducible. This could make reviewing easier, and also give the assessment teams a little more flexibility with their approach when working in R. However, this may also require considering producing key tools, such as ASPIC, as R packages or providing an R interface to the software. This approach could be tested to start with on the CPUE standardisation, where it should work well.
- Only add independent CIE sections to the end of the consensus RP report, rather than completing entirely separate reports. These sections would allow CIE reviewers to identify differences and personal opinion in important areas, avoiding repetition and perhaps minor unimportant observations.

11 ToR 8. Prepare a Peer Review Summary of the Panel's evaluation of the stock assessment, addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.

A separate Review Panel summary report was produced.

12 References

Stan 2016. Stan Development Team 2016. Stan Modeling Language: User's Guide and Reference Manual. Version 2.11.0.

McCullagh, P. and Nelder, J.A. 1989. Generalized linear models, Second Edition. Chapman and Hall, New York.

Grolemund, G. Wickham, H. 2016. R for Data Science. O'Reilly Media, Sebastopol, California.

Appendix 1: Bibliography of materials provided for review

Document #	Title	Authors
Documents Prepared for the Data Workshop (DW)		
SEDAR50-DW01	Brief Summary – Habitat and Developing Spatial Species Information for Blueline Tilefish in the South Atlantic Region	Pugliese 2016
SEDAR50-DW02	Summary of the 2015 blueline tilefish cooperative-with-industry data collection project	Kellison 2016
SEDAR50-DW03	A Preliminary Assessment of Reproductive Parameters for Blueline Tilefish in Atlantic Waters from Virginia to Florida	Kolmos et al. 2016
SEDAR50-DW04	Distribution of scientifically collected blueline tilefish (*Caulolatilus microps*) in the Atlantic, and associated habitat	Klibansky 2016
SEDAR50-DW05	Summary of the results of a genetic-based investigation of blueline tilefish (*Caulolatilus microps*)	McDowell 2016
SEDAR50-DW06	Preliminary Genetic Population Structure of Blueline Tilefish *Caulolatilus microps* along the East Coast of the United States	O'Donnell and Darden 2016
SEDAR50-DW07	Description of age and growth for blueline tilefish, *Caulolatilus microps*, caught north and south of Cape Hatteras, NC	Schmidtke and Jones 2016
SEDAR50-DW08	Standard Operative Procedure for Embedding and Sectioning Blueline Tilefish (*Caulolatilus microps*)	Ostrowski 2016
SEDAR50-DW09	Summary of Northeast Fisheries Science Center Blueline Tilefish Survey Data	Nitschke and Miller 2016
SEDAR50-DW10	Summary of Mid-Atlantic Commercial Blueline Tilefish Data	Nitschke and Miller 2016
SEDAR50-DW11	Distribution of blueline tilefish (Caulolatilus microps) in the U.S. EEZ from fishery-dependent and fishery-independent data collections	Farmer and Klibansky 2016
SEDAR50-DW12	**Recommendations from the SEDAR 50 (Blueline Tilefish) Stock ID Work Group Meeting**	**SEDAR 50 Stock ID Work Group 2016**
SEDAR50-DW13	Comparison of Blueline Tilefish Otolith Derived Ages: Comparing Increment Counts Derived by Readers from NMFS SEFSC-Beaufort and SCDNR Age Laboratories	Ballenger 2017
SEDAR50-DW14	TBD	TBD
~~SEDAR50-DW15~~	SEDAR 50 Public Comments – visit the following link to view public comments submitted for SEDAR 50	

	https://safmc.wufoo.com/reports/sedar-50-public-comments/	
SEDAR 50-DW16	**SEDAR 50 Stock Identification Joint SSC Review Webinar Consensus Statements**	**Joint SSC Sub-Panel 2016 (Includes MAFMC, SAFMC, GMFMC representatives)**
SEDAR 50-DW17	**SEDAR 50 Stock Identification – Management/Science Call Recommendations**	**Council, Science Center, and Regional Office Leadership**
SEDAR50-DW18	Blueline Tilefish Age Workshop II	Potts et al. 2016
SEDAR50-DW19	Reproductive parameters for Blueline Tilefish in Atlantic Waters from Virginia to Florida	Kolmos et al. 2017
SEDAR50-DW20	Virginia Blueline Tilefish Data Collection Summary	Cimino 2017
SEDAR50-DW21	Summary of the Blueline Tilefish meristic conversions using data from the entire US Atlantic and Gulf of Mexico	Ballew and Potts 2016
SEDAR50-DW22	SEDAR 50 Discard Mortality Ad-hoc Group Working Paper	Discard mortality ad-hoc group
SEDAR50-DW23	Estimating dispersal of blueline tilefish (*Caulolatilus microps*) eggs and larvae from drifter data	Klibansky 2017
SEDAR50-DW24	ToR #7 Ad Hoc Work Group Working Paper	ToR #7 Ad-Hoc Work Group
SEDAR50-DW25	Standardized catch rates of blueline tilefish (*Caulolatilus microps*) in the South Atlantic and Gulf of Mexico waters of the U.S. from recreational headboat logbook data	SFB-NMFS 2017
SEDAR50-DW26	Standardized catch rates of blueline tilefish (*Caulolatilus microps*) in the South Atlantic and Gulf of Mexico waters of the U.S. from commercial logbook handline data	SFB-NMFS 2017
SEDAR50-DW27	Standardized catch rates of blueline tilefish (*Caulolatilus microps*) in the South Atlantic and Gulf of Mexico waters of the U.S. from commercial logbook longline data	SFB-NMFS 2017
SEDAR50-DW28	SEDAR 50 additional management actions provided by R. Hudson	Hudson 2017
Documents Prepared for the Assessment Workshop		
SEDAR50-AW01	South Atlantic U.S. Blueline Tilefish (*Caulolatilus microps*) length composition from	SFB-NMFS 2017

	the recreational fisheries	
SEDAR50-AW02	Commercial length composition weighting for U.S. Blueline Tilefish (*Caulolatilus microps*)	SFB-NMFS 2017
SEDAR50-AW03	Additional Commercial Fishery Statistics: Landings in Weight and Number, Mean Weights, Update to Uncertainty, and Catch and Effort Maps	SEDAR 50 Commercial WG
Documents Prepared for the Review Workshop		
SEDAR50-RW01	Information to help interpret results from the data limited toolkit for Atlantic Blueline Tilefish north and south of Cape Hatteras	Ahrens 2017
Final Assessment Reports		
SEDAR50-SAR1	Assessment of Atlantic Blueline Tilefish	To be prepared by SEDAR 50
Reference Documents		
SEDAR50-RD01	SEDAR 32 South Atlantic Blueline Tilefish Stock Assessment Report	SEDAR 32
SEDAR50-RD02	List of documents and working papers for SEDAR 32 (South Atlantic Blueline Tilefish and Gray Triggerfish) – all documents available on the SEDAR website.	SEDAR 32
SEDAR50-RD03	Managing A Marine Stock Portfolio: Stock Identification, Structure, and Management of 25 Fishery Species along the Atlantic Coast of the United States	McBride 2014
SEDAR50-RD04	Workshop to Determine Optimal Approaches for Surveying the Deep-Water Species Complex Off the Southeastern U.S. Atlantic Coast	Carmichael et al. 2015
SEDAR50-RD05	Report to Virginia Marine Resources Commission: Grant F-132-R-2 The Population Dynamics of Blueline and Golden Tilefish, Snowy and Warsaw Grouper and Wreckfish	Schmidtke et al. 2015
SEDAR50-RD06	Estimated Catch of Blueline Tilefish in the Mid-Atlantic Region: Application of the Delphi Survey Process	Allen et al. 2016
SEDAR50-RD07	MAFMC Memo: Blueline Tilefish Catch Series – Feb 23, 2016	Didden 2016
SEDAR50-RD08	Reproductive Biology of the Blueline Tilefish, Caulolatilus microps, off North Carolina and South Carolina	Ross and Merriner 1983
SEDAR50-RD09	Fish species associated with shipwreck and	Ross et al. 2016

	natural hard-bottom habitats from the middle to outer continental shelf of the Middle Atlantic Night near Norfolk Canyon	
SEDAR50-RD10	Systematics and Biology of the Tilefishes (Perciformes: Branchiostegidae and Malacanthidae), with Descriptions of Two New Species	Dooley 1978
SEDAR50-RD11	Integrating DNA barcoding of fish eggs into ichthyoplankton monitoring programs	Lewis et al. 2015
SEDAR50-RD12	Age, growth, and reproductive biology of blueline tilefish along the southeastern coast of the United States, 1982-1999	Harris et al. 2004
SEDAR50-RD13	Description of the Circulation on the Continental Shelf	Bumpus 1973
SEDAR50-RD14	Spawning Locations for Atlantic Reef Fishes off the Southeastern U.S.	Sedberry et al. 2006
SEDAR50-RD15	Observations and a Model of the Mean Circulation over the Middle Atlantic Bight Continental Shelf	Lentz 2008
SEDAR50-RD16	Modeling larval connectivity of the Atlantic surfclams within the Middle Atlantic Bight: Model development, larval dispersal and metapopulation connectivity	Zhang et al. 2015
SEDAR50-RD17	Tilefishes of the Genus *Caulolatilus* Construct Burrows in the Sea Floor	Able et al. 1987
SEDAR50-RD18	Delineation of Tilefish, *Lopholatilus chamaeleonticeps*, Stocks Along the United States East Coast and in the Gulf of Mexico	Katz et al. 1983
SEDAR50-RD19	Chapter 22: Interdisciplinary Evaluation of Spatial Population Structure for Definition of Fishery Management Units (excerpt from Stock Identification Methods – Second Edition)	Cadrin et al. 2014
SEDAR50-RD20	Overview of sampling gears and standard protocols used by the Southeast Reef Fish Survey and its partners	Smart et al. 2015
SEDAR50-RD21	Age, Growth, and Mortality of Blueline Tilefish from North Carolina and South Carolina	Ross and Huntsman 1982
SEDAR50-RD22	Radiocarbon from nuclear testing applied to age validation of black drum, *Pogonias cromis*	Campana and Jones 1998
SEDAR50-RD23	A long- lived life history for a tropical, deepwater snapper (*Pristipomoides filamentosus*): bomb radiocarbon and lead-radium dating as extensions of daily increment analyses in otoliths	Andrews et al. 2012
SEDAR50-RD24	Age and growth of bluespine unicornfish (*Naso unicornis*): a half-century life-span for a	Andrews et al. 2016

	keystone browser, with a novel approach to bomb radiocarbon dating in the Hawaiian Islands	
SEDAR50-RD25	Age, growth and reproduction of the barrelfish *Hyperoglyphe perciformis* (Mitchill) in the western North Atlantic	Filer and Sedberry 2008
SEDAR50-RD26	Age, growth, and spawning season of red bream (Beryx decadactylus) off the southeastern United States	Friess and Sedberry 2011
SEDAR50-RD27	Great longevity of speckled hind (*Epinephelus drummondhayi*), a deep-water grouper, with novel use of postbomb radiocarbon dating in the Gulf of Mexico	Andrews et al. 2013
SEDAR50-RD28	Refined bomb radiocarbon dating of two iconic fishes of the Great Barrier Reef	Andrews et al. 2015
SEDAR50-RD29	Age validation of the North Atlantic stock of wreckfish (*Polyprion americanus*), based on bomb radiocarbon (^{14}C), and new estimates of life history parameters	Lytton et al. 2016
SEDAR50-RD30	Stock Complexes for Fisheries Management in the Gulf of Mexico	Farmer et al. 2016
SEDAR50-RD31	Modelling community structure and species co-occurrence using fishery observer data	Pulver et al. 2016
SEDAR50-RD32	Descriptions of the U.S. Gulf of Mexico Reef Fish Bottom Longline and Vertical Line Fisheries Based on Observer Data	Scott-Denton et al. 2011
SEDAR50-RD33	Natural mortality estimators for information-limited fisheries	Kenchington 2014
SEDAR50-RD34	The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural systems and aquaculture	Lorenzen 1996
SEDAR50-RD35	Mortality Rate of Fishes in the Pelagic Ecosystem	Peterson and Wroblewski 1984
SEDAR50-RD36	A Mathematical Model of Some Aspects of Fish Growth, Respiration, and Mortality	Ursin 1967
SEDAR50-RD37	MAFMC Memo: Blueline Tilefish Catch Series – Mar 14, 2016	Didden 2016
SEDAR50-RD38	Mid-Atlantic Fishery Management Council SSC Memo: Proposed BLT Subcommittee Report – March 22, 2016	Miller 2016
SEDAR50-RD39	Hierarchical analysis of multiple noisy abundance indices	Conn 2010
SEDAR50-RD40	Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding	McAllister et al. 2001

SEDAR50-RD41	Evaluating methods for setting catch limits in data-limited fisheries	Carruthers et al. 2014
SEDAR50-RD42	Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act	Restrepo et al. 1998
SEDAR50-RD43	A simple method for estimating MSY from catch and resilience	Martell and Froese 2012
SEDAR50-RD44	Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish	Gedamke and Hoenig 2006

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 50 Atlantic Blueline Tilefish 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

SEDAR 50 will be a compilation of data, an assessment of the stock, and CIE assessment review conducted for Atlantic Blueline Tilefish. The review workshop provides an independent peer review of SEDAR stock assessments. The term review is applied broadly, as the review panel may request additional analyses, error corrections and sensitivity runs of the assessment models provided by the assessment panel. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stock assessed through SEDAR 50 are within the jurisdiction of the South Atlantic Fishery Management Council, Mid-Atlantic Fishery Management Council, and the

states of Florida, Georgia, South Carolina, North Carolina, Virginia, Pennsylvania, New York, New Jersey, Maryland, and Delaware. The Terms of Reference (ToRs) of the peer review and the tentative agenda of the panel review meeting are below.

Requirements

NMFS requires three (3) CIE reviewers to conduct an impartial and independent peer review in accordance with the SoW, OMB guidelines, and the ToRs below. CIE reviewers shall have a working knowledge in the application of fisheries stock assessment processes and results, statistics, fisheries science, and marine biology sufficient to complete the primary task of providing peer-review advice in compliance with the workshop Terms of Reference. Additionally, it will be helpful if the reviewers have a working knowledge of data limited stock assessment approaches.

Tasks for reviewers

1) Review the following background materials and reports prior to the review meeting:

SEDAR 50 Workshop Reports and Working Papers

- Data Workshop Report and Working Papers will be available at the following link: http://sedarweb.org/sedar-50-data-workshop
- Assessment Workshop Report and Working Papers will be available at the following link: http://sedarweb.org/sedar-50-assessment-process
- Review Workshop Working Papers will be available at the following link: http://sedarweb.org/sedar-50-review-workshop

2) Attend and participate in the panel review meeting. The meeting will consist of presentations by NOAA scientists, stock assessment authors and others to facilitate the review, to answer any questions from the reviewers, and to provide any additional information required by the reviewers.

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

4) Each reviewer should assist the Chair of the meeting with contributions to the summary report.

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

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide

requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be at the contractor's facilities, and at Atlantic Beach, NC.

Period of Performance

The period of performance shall be from the time of award through October 27, 2017. The CIE reviewers' duties shall not exceed 14 days to complete all required tasks.

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

Within two weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
August 29 - 31, 2017	Panel review meeting
Approximately 3 weeks later	Contractor receives draft reports
Within 2 of receiving draft reports	Contractor submits final reports to the Government

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each ToR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this

contract. Travel is not to exceed $10,000.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact:

Julia Byrd
SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
(843) 571-4366
julia.byrd@safmc.net

Peer Review Report Requirements

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

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each ToR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the ToRs.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

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

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

3. The report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review

   Appendix 2: A copy of this Statement of Work

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

Terms of Reference for the Peer Review

SEDAR 50 Atlantic Blueline Tilefish Assessment Review

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

b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
c) Are data applied appropriately within the assessment model?
d) Are input data series reliable and sufficient to support the assessment approach and findings?

2. Evaluate the methods used to assess the stock, taking into account the available data.
   a) Are methods scientifically sound and robust? Do the methods follow accepted scientific practices?
   b) Are assessment models configured appropriately and applied consistent with accepted scientific practices?
   c) Are the methods appropriate for the available data?

3. Evaluate the assessment findings with respect to the following:
   a) Are population estimates (model output – e.g. abundance, exploitation, biomass) 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 appropriate for management use? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

4. Evaluate the stock projections, 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?

5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture all sources of uncertainty in the population, data sources, and assessment methods
   - Are the implications of uncertainty in technical conclusions clearly stated?

6. Consider the research recommendations provided by the Data and Assessment workshops 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.
7. Provide suggestions on improvements in data or modeling approaches, which should be considered when scheduling the next assessment.
8. Prepare a Peer Review Summary of the Panel's evaluation of the stock assessment, addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.

Tentative AGENDA

SEDAR 50 Atlantic Blueline Tilefish Review Workshop

Atlantic Beach, North Carolina
August 29 - 31, 2017

Tuesday

9:00 a.m.	**Convene**	
9:00 – 9:30 a.m. /	**Introductions and Opening Remarks**	**Coordinator**
	- Agenda Review, ToR, Task Assignments	**Chair**
9:30 a.m. – 12 p.m.	**Assessment Presentation**	**TBD**
12:00 p.m. – 1:30 p.m.	**Lunch Break**	
1:30 p.m. – 3.30 p.m.	**Panel Discussion**	**Chair**
	- *Assessment Data & Methods*	
	- Identify additional analyses, sensitivities, corrections	
3:30 p.m. – 4:00 p.m.	**Break**	
4:00 p.m. – 5:30 p.m.	**Continue Panel Discussion**	**Chair**
5:30 p.m. – 6:00 p.m.	**Public Comment**	

***Tuesday Goals*:** Initial presentation completed, sensitivities and modifications identified.

Wednesday

8:30 a.m. – 11:30 a.m.	**Panel Discussion**	**Chair**
	- Continue deliberations	
	- Review additional analyses	
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:30 p.m. – 3:30 p.m.	**Panel Discussion**	**Chair**
	- Continue deliberations	
	- Review additional analyses	
3:30 p.m. – 4:00 p.m.	**Break**	
4:00 p.m. – 5:30 p.m.	**Panel Discussion or Work Session**	**Chair**
	- Recommendations and comments	
5:30 p.m. – 6:00 p.m.	**Public Comment**	

***Wednesday Goals*:** Preferred models selected, projection approaches approved, Report drafts begun

Thursday

8:30 a.m. – 10:30 a.m.	**Panel Discussion**	**Chair**
	- Review additional analyses, final sensitivities	
	- Projections reviewed.	
10:30 a.m. – 10:45 p.m.	**Break**	
10:45 a.m. – 12:30 p.m.	**Panel Discussion or Work Session**	**Chair**

	- Review Consensus Reports	
12:30 p.m. – 1:00 p.m.	**Public Comment**	**Chair**
1:00 p.m.	**ADJOURN**	

Thursday Goals: Complete assessment work and discussions. Final results available. Draft Summary Report reviewed.

DoubleTree by Hilton Hotel Atlantic Beach Oceanfront
2717 West Fort Macon Road, Atlantic Beach, NC 28512
T: 1 252-240-1155

Aug. 28, 2017 – Sep. 01, 2017
Confirmation Number: 87566229

Appendix 3: Review Workshop Panel Membership

REVIEW PANEL

Scott Crosson	Review Panel Chair	SAFMC SSC
Churchill Grimes	Reviewer	SAFMC SSC
Yan Jiao	Reviewer	MAFMC SSC
Patrick Cordue	CIE Reviewer	CIE
Jamie Gibson	CIE Reviewer	CIE
Paul Medley	CIE Reviewer	CIE

ANALYTICAL REPRESENTATIVES

Nikolai Klibansky	Lead analyst	SEFSC Beaufort
Kevin Craig	Assessment Team	SEFSC Beaufort
Kyle Shertzer	Assessment Team	SEFSC Beaufort
Erik Williams	Assessment Team	SEFSC Beaufort

APPOINTED OBSERVERS

Rusty Hudson	Recreational/Commercial	FL / SFA
Andy Piland	For-hire	NC

COUNCIL REPRESENTATIVES

Dewey Hemilright	Council member	MAFMC
Anna Beckwith	Council member	SAFMC

COUNCIL AND AGENCY STAFF

Julia Byrd	Coordinator	SEDAR
Kimberly Cole	Admin	SEDAR/SAFMC
Jason Didden/Matt Seeley	MAFMC lead	MAFMC
Mike Errigo	Fishery Biologist	SAFMC
Jeff Pulver	Fishery Biologist	SERO

*Did not participate in Review Workshop