Center for Independent Experts (CIE) Reviewer Report for the SEDAR Atlantic Blueline Tilefish Review Workshop (SEDAR 50) held August 29-31, 2017, in Atlantic Beach, North Carolina.

Prepared for:

Center for Independent Experts

By:

Jamie Gibson,
100 Allison Coldwell Road,
Newtonville, Nova Scotia, Canada
B4P 2R1

Executive Summary

This document forms my independent reviewer report of review activities and findings for the 50th Southeast Data, Assessment and Review (SEDAR 50) Review Workshop, held August 29-31, 2017 in Atlantic Beach, North Carolina. Atlantic Blueline Tilefish was assessed as two stocks during SEDAR 50: a stock occurring in waters north of Cape Hatteras, and a stock that occupies waters to the south of Cape Hatteras.

Overall, the Data Workshop Report and supporting working papers clearly presented the information available for assessing these stocks. Data available for the assessment fall into the broad categories of stock delineation, life history parameters, removals (commercial and recreational landings and dead discards), fishery-dependent indices of abundance and length-frequency data. Age-frequency data were expected to be available, but inconsistencies in the age determinations could not be resolved prior to the Assessment Workshop. The material for the Review Workshop was very well-organized in the workshop reports, background information and presentations, which facilitated the review.

The Assessment Workshop Report and supporting working papers also clearly presented the analyses undertaken to assess the status of the two stocks. Two models were used for the stock south of Cape Hatteras: a biomass dynamics model (ASPIC), and a statistical, age-structured model (ASPM). The ASPIC model was fit to combinations of three CPUE abundance indices, none of which provided information about relative abundance after 2007. The ASPM analyses were fit to the same CPUE abundance indices as well as the length-frequency data, and also used information about life history. Life history parameter values are not well known for Blueline Tilefish, so most values were assumed based on what is known about related species. The Analytical Team undertook many sensitivity analyses using both models to evaluate the effects of assumptions made when setting up the models. The key identified sensitivity pertained to a spike in the landings during the early 1980's. The reliability of this spike was questioned during the Review Workshop, and the analyses showed that biomass estimates were sensitive to assumptions made about the magnitude of this spike. Overall, because of the large number of sensitivity analyses undertaken for this stock, I believe the results are reasonably robust given the data inputs, although the range of projected yields from these analyses is wide. Uncertainty in the stock structure and recruitment dynamics, in the life history parameter values, in the removal times series, and in the selectivity of the fishery, as well as the dated CPUE indices and the lack of fishery-independent data, are all considerations when interpreting the assessment results for the southern stock.

The primary analytical method used for the stock north of Cape Hatteras was the Data-Limited Methods Toolkit. The Analytical Team filtered the many available methods within the tool kit and selected three that produced MSY approximations. The catch recommendations from these methods were compared with recent landings. Because fishing occurs in a small portion of the stock's geographic range, and there appears to be no information about abundance in the rest of its range, I believe these results are best interpreted qualitatively. Analyses presented were suggestive that localized depletion could occur. However, because recruitment dynamics are not well understood, this conclusion is uncertain. In addition, the effect of these removals on the entire stock is not known.

Research recommendations to improve future Atlantic Blueline Tilefish assessments include: development of a fishery-independent abundance index covering their full range; estimation of the selectivity of the gear used in the fishery; resolution, if possible, of the age determination issues; research to better understand recruitment processes; and research to better understanding fishing processes and gear selectivity.

Table of Contents

1.0. Background

SEDAR 50, the 50th Southeast Data, Assessment and Review process, was an assessment of the status of Atlantic Blueline Tilefish and its fisheries. Blueline Tilefish, *Caulolatilus microps*, is a deep-living species of fish found in the western Atlantic Ocean from Cape Charles, Virginia, to the Florida Keys; and in the Gulf of Mexico from the Florida Keys to the Yucatan Peninsula, Mexico. It is fished commercially and recreationally throughout much of its range. The unit stock for SEDAR 50 included the entire Atlantic Seaboard, using the boundary between the Gulf of Mexico Fishery Management Council and South Atlantic Fishery Management Council management areas as the southwestern boundary for the stock unit to assess. Atlantic Blueline Tilefish was last assessed in 2013 (SEDAR 32).

This document forms my independent reviewer report of activities, findings and conclusions of the SEDAR 50 Review Workshop, held August 29-31, 2017 in Atlantic Beach, North Carolina. SEDAR 50 consisted of a data workshop, an assessment process consisting of an in-person workshop and four webinars, and the peer-review workshop. Working papers, background material, and the data workshop and assessment reports (Appendix 1) were provided to the Review Panel (Appendix 3) two weeks prior to the Review Workshop. Together these documents provided a comprehensive and clearly presented compilation of the data available for these species, the decisions that were made about how to use these data, the analyses that were undertaken, and the results and conclusions of the assessment. These documents provided a good basis upon which to conduct the peer review. The Analytical Team did a lot of work during and after the Review Workshop; providing the results of multiple sensitivity analyses and population projections as an addendum after the Review Workshop. These results were very helpful for understanding the model behavior, and the consistency of the results from a surplus production model and an age-structured analysis.

2.0. Individual Reviewer Activities

I conducted my activities in accordance with the Statement of Work (SOW) provided for this review (Appendix 2). Prior to the meeting, I reviewed all the assessment and background documents provided for the workshop. I participated in the Review Workshop in Atlantic Beach, North Carolina, August 29-31, 2017. During this meeting, I actively participated as member of the Review Panel, and discussed and questioned several aspects of the data and models. After the Review Workshop, I prepared this individual, independent report and assisted with writing the Review Workshop Report. As outlined in Appendix 2, this independent report is intended to summarize review activities completed during the Review Workshop meeting, including provision of a detailed summary of findings, conclusions, and recommendations for each TOR. The following sections in this document contain my personal perspectives about this assessment.

3.0. Summary of Findings in Accordance with the TOR's

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?***

The Data Workshop Report, Assessment Workshop Report and supporting working papers thoroughly documented the information available for assessing Atlantic Blueline Tilefish. Data types include: information in support of stock identification decisions; information about life history parameters such as age-at-maturity, natural mortality rates, growth rates, length-weight conversion coefficients and the stock-recruitment steepness parameter; commercial landings including discards and discard mortality; recreational fisheries harvests and releases (including mortality rates of released fish); abundance indices based on CPUE indices; length-composition data from the fisheries; and information in support of the decision not to use the age-composition data. In general, I think the data decisions are sound and robust and the data are applied appropriately within the assessment models. Both prior to and during the Review Workshop, the Analytical Team undertook sensitivity analyses to many of the data decisions and found that the assessment results are sensitive to some of the decisions. In my opinion, if the findings are interpreted in the context of these analyses, the input data series are reliable enough and sufficient to support the assessment approach and findings.

Stock Identification

The delineation of stocks for the assessment is a key decision in any assessment. Many assessment approaches, including those used in this assessment are based on the assumption that the stock (or population) is closed. Stocks are often delineated such that this assumption is met. Under this assumption, intrinsic factors such as growth, reproductive rates, carrying capacity, natural mortality and mortality caused by human activities are the significant determinants of a population's abundance and dynamics, whereas extrinsic factors such as immigration and emigration can be ignored.

Stock identification decisions were well described in the Assessment Workshop Report. The SEDAR 50 Stock ID Working Group considered genetics, life history data, adult distributions, oceanographic features and data on drifter movement when delineating stocks of Blueline Tilefish in US waters. Although the limited genetics data indicated that that Blueline Tilefish are genetically homogeneous at large and small spatial scales, the biological processes that determine abundance and dynamics, particularly recruitment, may occur predominantly at smaller scales. Eggs and larvae of Blueline Tilefish are thought to be planktonic, and are expected to travel on the ocean currents. Analyses of drogue drifter data showed that a

significant portion of drifter tracks originating in the Gulf of Mexico moved into the US South Atlantic with very little movement in the opposite direction. A portion of drifter tracks originating in the US South Atlantic moved into the Mid-Atlantic with some movement of drifters in the opposite direction. If eggs and larvae behave like drifters, there is potential for substantial movement from the Gulf of Mexico into the US South-Atlantic with little movement in the opposite direction, and also for substantial exchange between the South and Mid-Atlantic with a net movement north. Movements of adult ABT are thought to be very limited. This information, coupled with a spatial mismatch between the CPUE indices and recent removals, led to a decision that Blueline Tilefish in the Atlantic be modelled as two stocks: one extending from the SAFMC/GMFMC boundary to Cape Hatteras, and the other extending north from Cape Hatteras.

Overall, the decisions about stock delineation appear practical, but based on the information above, may not result in stocks that represent closed populations. For example, if a significant portion of the recruitment in the stock south of Cape Hatteras (southern stock) is the result of egg or larval drift from the Gulf of Mexico, there would not be a meaningful stock-recruitment relationship for the southern stock and productivity could be overestimated if only data for the southern stock is used in its estimation. It would not make sense to model the Gulf of Mexico and the southern stock as a single closed population because the flow of eggs and larvae is mostly in one direction (under the assumption that the drifter data represent movement of eggs and larvae, abundance of Blueline Tilefish in the Gulf of Mexico would not be expected to be largely influenced by the abundance Blueline Tilefish in the US South Atlantic). While modelling Blueline Tilefish in the Gulf of Mexico and US South Atlantic as two stocks with an asymmetrical flow of eggs and larvae between the two areas is theoretically possible, in my opinion, it would be difficult to support this complex model with the data presented in this assessment.

Based on the above, I agree with the decisions made about stock structure for SEDAR 50, however, the recruitment dynamics of Blueline Tilefish appear quite complex, and if driven by currents, are likely highly variable on multiple spatial scales and in different areas. Implications for the selection of an assessment model are discussed under TOR 2.

Age composition

In contrast with SEDAR 32, age composition data were not used in SEDAR 50. This decision was based on low agreement among readers from three laboratories, low within reader consistency and results from an SCDNR bomb radiocarbon study that indicated potential under-aging. These issues are well described in SEDAR50-DW-18 and in the Data Workshop Report. Issues included variation in increment or annulus count depending on the section of the otolith that was read, uncertainty in how to count opaque zones, and consistency in identifying the first annulus.

The decision that the age composition data were not reliable had major implications for this assessment, including the selection of assessment models, estimation of growth parameters and

estimation of natural mortality rates based on maximum age. Notwithstanding these implications, I believe the Data Workshop recommendation not to use the age composition data is the correct recommendation.

Life history data

Particularly given the decision that the age determinations were not reliable, Blueline Tilefish life history parameter values are not well known. Von Bertalanffy growth parameters and the Beverton-Holt steepness parameter values were derived from meta-analyses of data from similar species. Meristic conversion factors were estimated using length-frequency data. Several methods were used to estimate the natural mortality rate for Blueline Tilefish, with the final decision to estimate the natural mortality rate based on an assumed maximum age of 40 years. Age at 50% maturity was estimated from empirical data, although, as pointed out by the Analytical Team, too few immature fish were captured to be sure the estimated value was reliable. Discard mortality rates were derived from fishery observer data from the Gulf of Mexico.

Overall, I believe the decisions about the life history data are well thought out, are well described in the Data Workshop and Assessment Workshop reports, and are defensible. With respect to their application in the assessment models for the stock south of Cape Hatteras, with the exception of the discard mortality rates and meristic conversion factors (which appear reasonably well known), they do not enter into the ASPIC model. With the ASPM analysis, the analytical team undertook many sensitivity runs using different values of natural mortality and steepness that I believe sufficiently address the uncertainty in these parameter values. With the data-limited methods used for the stock north of Cape Hatteras, the life history parameters are input as a range, and the bounds placed on the parameter values appear reasonable. Given the Review Workshop recommendation that the results of these analyses are best interpreted qualitatively, the values appear sufficient to support the findings for the northern stock.

Abundance indices

Abundance indices were not available for the Blueline Tilefish stock north of Cape Hatteras. For the stock south of Cape Hatteras, the Data Workshop Report and background documents provided thorough descriptions of the available abundance indices, methods used for their calculation, their strengths and weaknesses, justification for their inclusion or exclusion as data inputs for the model and presentation of data source variance. Three commercial fishery-dependent CPUE time series, four recreational fishery-dependent CPUE time series and five fishery-independent surveys were considered for the assessment. All but three, the recreational headboat, the commercial logbook longline, and the commercial logbook handline were excluded because the series were either quite short and/or they encountered very few Blueline Tilefish each year. The strengths and weaknesses of these three indices were well described in the Data Workshop Report. Uncertainties are similar to many fishery-dependent indices, including the potential for catchability to vary over time or with abundance, or due to changes in species being targeted.

None of the indices provide information about abundance in the last ten years. The headboat index was truncated in 2005 due to increased targeting of Blueline Tilefish; the longline index after 2007 due to changes in regulations and trends in fishing; and handline index after 2006 because of a shift in fishing effort almost entirely to north of Cape Hatteras. The truncation of the CPUE indices due to changes in regulations and trends in fishing behavior means that there are no data about relative abundance since 2007. All three CPUE series were standardized using the delta-GLM approach in which two separate generalized linear models are combined; one which models the presence/absence of the species, and one that models the catch rates on trips that caught Blueline Tilefish.

Ideally, an abundance index is collected using an experimental design that ensures the resulting index is proportional to abundance. Commercial CPUE information is not typically collected in this way. If fisheries are able to target areas of higher abundance, or if the efficiency of the fishery changes over time, hyperstability in the index may occur. As pointed out by another reviewer, there may be year-area effects in the index that would not have been addressed using the standardization approach applied to these data that could potentially lead to hyperstability. If localized depletion occurs and fisheries are able to target areas subject to less recent fishing, hyperstability would be expected. Additionally, length- and age-frequency data collected from the fishery would not be representative of the stock, which could lead to bias in the assessment. Additionally, in fisheries that harvest more than one species, changes in the targeted species may also lead to changes in the index that are not the result of changes in abundance.

Although there are potential issues with the CPUE indices, I believe the decisions made about these data are generally sound and that they are sufficient for this assessment. However, the results of the assessment should be interpreted taking into consideration the duration of these series as well as the uncertainties in their representativeness of abundance trends.

Length composition

The length-composition data played a relatively small role in the assessment of the stock south of Cape Hatteras. One of the models, a surplus production model, does not use these data; and the other model, an age-structured model, appeared to be weighted towards the abundance indices. As discussed under TOR 3, the signal provided to the age-structured model by the length composition data appeared to differ from that provided by the abundance indices, but the reason for this discrepancy is not known. The length composition data were integral to the data-limited methods applied to the northern stock, because some of the methods used the mean length to estimate total mortality based on the assumption that the fishery selectivity was asymptotic. This assumption does not appear to be testable with the data presented at the workshop, but if the selectivity of the fishing gear is highly dome-shaped, then the application of these data in the model may not be appropriate. This is a topic for further exploration, but given that the assessment results for the stock south of Cape Hatteras are driven primarily by the abundance indices, and because the conclusions for the stock north of Cape Hatteras are mainly qualitative, I do not believe this to be a major issue for this assessment.

Recreational and commercial landings

The derivations of the recreational and commercial landings are well described in the Data Workshop Report and in the supporting documentation. Commercial landings by weight were developed by gear (handline, longline, other) for the time period from 1950 to 2015. Commercial discards were calculated using fisher-reported discard rates. Recreational landings were obtained from the Marine Recreational Fisheries Statistics Survey and the Marine Recreational Information Program (MRIP/MRFSS); and the Southeast Regional Headboat Survey. Discarded live fish are reported by anglers interviewed by MRIP/MRFSS. Landings in the recreational fishery have become an increasingly important component of the catch, particularly for the northern stock.

Tilefish landings were not reported by species prior to 1985. The methods used to determine the portion of Blueline Tilefish in the landings varied from state to state, but involved determining the proportion of the landings in a later time period and applying that proportion to the earlier time period. For the stock south of Cape Hatteras, the method resulted in a large spike in landings just prior to 1985 that lacked credibility with stakeholders present at the Review Workshop. I agree with the Review Workshop opinion that this source of uncertainty was not well documented in the Data Workshop report, but believe it was well addressed at the Review Workshop via sensitivity analyses undertaken by the analytical team. This appeared to be the key source of uncertainty in the assessment for the southern stock.

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?***

South of Cape Hatteras

The Analytical Team originally intended to use the Beaufort Assessment Model, BAM, (Williams and Shertzer 2015), however because the age-composition data were deemed unreliable, the Team switched to using a logistic surplus production model, ASPIC (Prager 2015), as the primary model and used an age-structured production model (ASPM) as a supporting analysis. This latter model was developed by modifying code for BAM.

The data sources for the ASPIC model were the removals time series (commercial and recreational landings and dead discards) and three CPUE time series (the commercial longline CPUE index, the commercial handline CPUE index, and the recreational headboat CPUE time series). The model does not use life history parameter information, age-composition data or length composition data. Production is modelled using a two parameter logistic growth curve: the intrinsic rate of increase at low abundance in the absence of density dependence (r), and the

unfished biomass or carrying capacity (K). As implemented in SEDAR 50, the logistic growth curve is symmetrical. As such, B_{MSY} equals $0.5K$, and F_{MSY} equals $0.5r$. The population was assumed to be in an unfished state at the start of the removal time series.

The ASPM model is a statistical age-structured model that provides a very flexible modelling framework. Landings were modelled for the three fleets separately and can be entered either as numbers of fish or by weight. Life history parameters can be estimated externally, or, if sufficient data exists, estimated within the model. In SEDAR 50, the model was fit to the removals time series, the three CPUE time series and to the annual length composition data for the removals. A Beverton-Holt model was used as the stock-recruitment relationship, the fishery selectivity was assumed to be flat-topped for all fisheries, and the Baranov catch equation was used to model the combined landings and discards.

In the absence of information about life history processes and fishery selectivity, the surplus production model data requirements do match the data available for the Blueline Tilefish stock north of Cape Hatteras, but the biological processes are highly simplified. For example, the biomass is modelled as "observed" by the fishery and if the selectivity of the fishery changes (e.g., changing fleets), then the biomass as modelled also changes. The annual time step may not be appropriate if the fishery selects primarily older fish, and the assumption of a symmetrical logistic growth curve is typically not consistent with the results from age-structured models.

The ASPM model explicitly includes information about life history and fishery characteristics, and biological and fishing processes are modelled much more explicitly. However, they are more data-intense and information about these processes needs to be provided in some way. In the absence of sufficient data to support parameter estimation, they can be assumed from other stocks or related species, but introduce uncertainty as a result.

In summary, one of the models, ASPIC, matches the data available for this stock, but models biological processes in a restrictive way that may not reflect the true dynamics of this stock. In contrast, the ASPM model better matches the underlying life history and fishing processes, but requires more assumed information than is required for ASPIC. Because the underlying processes are better modelled using ASPM, and because the assumptions can be evaluated using sensitivity analyses, I agree with the Review Workshop consensus preference for the ASPM as the primary model for this stock.

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

Notwithstanding their relative strengths and weaknesses, I consider the two modelling approaches used for the Blueline Tilefish stock south of Cape Hatteras to be sound. Both models have been used extensively in previous assessments, and are based on accepted scientific principles. In my opinion, the application of both models followed accepted scientific practices. I particularly liked the use of a large number of sensitivity analyses to evaluate the robustness of the conclusions to decisions made when setting up the models.

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

In my opinion, the models were, for the most part, configured appropriately. However, I also believe that some of the Review Workshop recommended modifications to the base models did improve the models. The Analytical Team originally averaged the results of two models fit to the longline and handline CPUE indices separately. This was done to avoid the problem that the handline CPUE index was dominating the model fit because it had the lower coefficient of variation, although both indices were considered equally plausible. I consider the approach recommended by the Review Workshop to put both indices into the same model with equal CV's to be a better way of including two equally plausible indices. The primary recommended modification to the reference ASPM model was to move the age-at-maturity from age-2 to age-6. The Review Workshop judged this to be more conservative (because the lower age at maturity confers more resilience to exploitation) and possibly more plausible than an age-of-maturity of age-2. Sensitivity analyses were undertaken around this value.

If, as suggested with high uncertainty in the Assessment Workshop Report, the stocks are actually open populations and dispersal occurs via egg and/or larval drift, then a model with an underlying stock-recruitment relationship may not be the most appropriate. An alternative would be to estimate annual year-class strength independent of spawner biomass. This is not the same as assuming that spawner biomass is not an important consideration, rather year-class strength is simply estimated without the constraint that it is a function of biomass (dynamics can then be explored external to the assessment model). Analyses undertaken at the Review Workshop indicated, not surprisingly, that the data were not sufficient to estimate recruitment deviates and therefore might not be sufficient to support this approach. While I consider the models as used sufficient for this assessment, there is an inconsistency between the assumptions of the models and the (somewhat speculative) information about dispersal provided in the Assessment Workshop Report.

Although it appeared possible to estimate von Bertalanffy growth parameters from the length composition data using the ASPM, the extent to which these result from other model assumptions such as the natural mortality rate and the assumed flat-topped selectivity curve is not clear. Although I don't think this potential issue undermines the assessment, further work is needed prior to assuming these are representative of blueline tilefish.

c) Are the methods appropriate for the available data?

I consider the methods used in the assessment of the Blueline Tilefish stock south of Cape Hatteras appropriate for the available data. The ASPIC model analysis was appropriate given the lack of age data, but is more restrictive (e.g., assuming that biomass can be modelled using a logistic growth curve) than the ASPM model. The ASPM model is much more flexible, takes advantage of life history information and length-composition data, and can be used for model exploration, but the plausibility of the many scenarios that are produced through exploration is

difficult to assess given the data limitations. In my opinion, the results of this assessment are best applied considering the suite of results from both models.

When the ASPM model was fit primarily to the length-composition data, abundance was estimated to be higher than when the model was weighted towards the abundance indices, with comparatively little change in abundance in response to removals by the fishery. Although this issue was not fully explored at the Review Workshop, this result suggests that signal in the length-composition data may not be very consistent with the decline in the early 1980's followed by gradually increasing abundance as indicated by the model runs weighting the abundance indices more heavily. Although the reason for this discrepancy is not known, possible explanations include increasing fishing efficiency (the abundance indices are not representative of abundance), targeting of areas that have been subjected to less fishing pressure (neither the length-composition data nor the indices are representative of changes in population-level abundance), selectivity is not appropriately modeled, or there is an issue with the derivation of the length-composition data. In my opinion, this is a topic for further exploration which I think is important, particularly given that some of the same questions may arise with the age-composition data if the aging issues are addressed. I do not believe this question undermines the results of this assessment.

North of Cape Hatteras

The primary analytical tool used for the Blueline Tilefish stock North of Cape Hatteras was the Data-Limited Methods Toolkit (DLMtool; Carruthers et al. 2014, Carruthers et al. 2015, and Carruthers and Hordyk 2016). This program undertakes an evaluation of potential assessment methods, develops catch recommendations based on the methods, and, although not used in SEDAR 50, has the capability to evaluate the performance of the data-limited approaches via management strategy evaluation.

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

Many of the methods in the DLMtool use standard equations that have been used extensively in the past. This tool has been used in other assessments and has been peer reviewed in the primary literature. I consider these methods to be sound. Because there were few data available for this stock (only the removals time series and a small amount of length-composition data) the use of this tool use is appropriate for this stock.

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

The Analytical Team filtered the many methods available in the DLMtool in order to select methods that would provide TAC recommendations. They selected three methods that used MSY approximations and compared the results from these methods with scenarios based on average catch. Two of these methods use a mean length-based mortality rate estimator (Gedamke and Hoenig 2006) to estimate the total mortality rate for the stock. The length-based mortality

estimator is an assumption rich method for estimating the total mortality rate from the von Bertalanffy growth model parameters, an estimate of the length at full selection and the mean length of fish greater than the length at full selection. Two issues pertaining to the use of these methods are: 1) whether the length-composition data are representative of the entire stock, and 2) whether the selectivity of fisheries matches the assumed selectivity for using a mean length-based mortality rate estimator. Most of the landings for the stock North of Cape Hatteras are from the southern part of its range and there is no information about relative abundance further north. Length composition from the fishery may potentially reflect the characteristics of the portion of the stock in the area where the fishery is taking place rather than the entire stock. Second, an assumption of the length-based method is that all fish larger that the length at which they are fully selected are, in fact, fully selected. The use of a length-based mortality rate estimator would be less appropriate if the fishery selectivity was dome-shaped, an assumption at times made for gears such as hooks. The effect of both of these potential issues would be to overestimate the total mortality rate at the scale of the entire stock.

c) Are the methods appropriate for the available data?

I believe that the methods used are 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?***

<u>South of Cape Hatteras</u>

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?

Overall, I consider the population estimates (model output – e.g., abundance, exploitation, biomass) reliable, consistent with input data and population biological characteristics, and, when interpreted in the context of their uncertainty, useful to support status inferences. However, the uncertainty in the population estimates is high. Sources of this uncertainty include: population structure and recruitment dynamics, data inputs (dated CPUE indices, life history parameters,

and the removals time series), as well as model selection, configuration and parameter estimation. The Analytical Team undertook many sensitivity analyses both before and during the Review Workshop to appropriately characterize the uncertainty.

Uncertainty in the model selection appeared well addressed via analyses undertaken at the Review Workshop. The biomass estimates from the ASPIC model are not directly comparable to those from ASPM; nor are status determinations directly comparable because the reference points against which biomass and fishing mortality rate estimates are compared also differ between the models. However, the reference model runs from both ASPIC and the ASPM show similar abundance trends, and projected yields at $F_{current}$ differ by about 23%, indicating that the two models are producing estimates on roughly the same scale.

Uncertainty in the abundance indices was primarily addressed using the ASPIC model by fitting the model to various combinations of the commercial handline, longline, and the headboat indices. The longline and handline indices provide a relatively consistent signal and status determinations from models fit to these time series are reasonably consistent. Models including the headboat index provide status determinations suggesting F/F_{MSY} ratios that are lower, and $\boldsymbol{B}/B_{\mathrm{MSY}}$ ratios that are higher, than model runs without this index. Of the models considered, I agree with the Review Panel recommendation that the model run including the handline and longline indices with equal weighting is the most appropriate ASPIC base model run. This model indicates a decline in biomass in the early 1980's followed by increasing biomass in the 2000's to about 50% of the unfished biomass.

Sensitivity analyses undertaken to address uncertainty in life history parameter estimates were undertaken using the ASPM. Relative to the reference model, these included different ages-at-maturity, different steepness values and different instantaneous natural mortality rates. These analyses produced similar biomass trends, although as expected, the magnitude of the biomass differed among runs. However, I believe these analyses generally support the ASPIC model results. Sensitivity to the removals time series, specifically to the magnitude of the catch peak in the early 1980s, was identified as an uncertainty at the Review Workshop. If this spike is reduced to 10% of its value, biomass at the start of the fishery did not need to be so high as to be able to support the high landings and biomass is further depleted. Although this reduction may be extreme (although we don't know), I agree with the Review Workshop conclusion that this is the key source of uncertainty in this assessment to consider when applying the assessment results. Yield projections from this model run are lower than from other model runs.

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

Overall, I believe the assessment results provide evidence that the stock south of Cape Hatteras is not overfished. The ASPIC base model run and the sensitivity analyses pertaining to which CPUE series are included all indicate *B*/*MSST* ratios above 1.3. The ASPM model results also support this conclusion: *B*/*MSST* ratios from the ASPM reference model and sensitivity analyses are all above 2.2.

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

The results of this assessment also provide evidence that the stock south of Cape Hatteras is not undergoing overfishing, although this conclusion has greater uncertainty than whether the stock is in an overfished state. Estimated F/F_{MSY} ratios are less than one from the Review Workshop base ASPIC model run, and also from the ASPM analyses. However, two of the ASPIC model runs that included the commercial handline CPUE index resulted in estimates of F/F_{MSY} ratios slightly greater than one. Additionally, for the ASPIC base model run, bootstrap results indicate a significant portion of the probability density for F/F_{MSY} is greater than one.

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

Both the ASPIC and ASPM models used in this assessment assume a stock-recruitment or biomass-production relationship. The Review Workshop was presented with considerable information suggestive that egg and larval drift could result in dispersal among the three putative stocks, resulting in stocks for which recruitment is partially driven by immigration and/or emigration. If recruitment is largely dependent on abundance of the other stocks, then assuming stock-recruitment relationships for the individual stocks would be inappropriate. However, at present, recruitment dynamics for Atlantic Blueline Tilefish are not well understood and the magnitude of dispersal among stocks is not known. Given this uncertainty as well as the data limitations, assuming that a stock-recruitment or biomass-production relationship exists is a practical decision without which this assessment likely could not have been completed. Surplus production models require the assumption and exploratory analyses undertaken at the review workshop indicated that the data could not support time-varying estimates of year-class strength from the ASPM without the stock-recruitment relationship.

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?

Overall, I believe that the quantitative estimates of status determination criteria are appropriate for management use for this stock. The status determinations described above were generally robust to many sensitivity analyses undertaken using both ASPIC and ASPM. However, the strength of the conclusions about whether overfishing is occurring, and whether the stock is in an overfished state, differs among model runs, not all of which are equally plausible. As discussed above, there is uncertainty as to whether or not the stock can be considered a closed population. With respect to status determination criteria for determining whether the overfishing is occurring, I prefer SPR-based reference points ($F_{x\%}$) over MSY-based reference points (F_{MSY}), because the stock-recruitment relationship does not need to be known or assumed to estimate SPR. If a significant portion of the recruitment originates outside the stock boundaries, determining the appropriate percent SPR is problematic, but the recommendation of the Review Workshop to use $F_{40\%}$ does not seem inconsistent with what is known about the life history of Blueline Tilefish.

North of Cape Hatteras

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?

In my opinion, the DLM analyses undertaken for the northern Blueline Tilefish stock are consistent with the input data and, to the extent that they are known, with the biological characteristics of the stock. However, nearly all of the data available for the northern Blueline Tilefish stock are from the fishery in an area just north of Cape Hatteras where the majority of the landings are taken. As such, the extent to which the data are representative of the entire stock is not known. Additionally, abundance, biomass and exploitation rates were not provided via the analyses undertaken for this stock, which I consider appropriate given the high uncertainty for the inputs for the analyses. I think the results of the analyses are highly uncertain, are not informative about whether the stock is in an overfished state, but may be considered to provide a qualitative evaluation of whether overfishing is occurring, at least on a local scale.

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

The methods used to assess the Blueline Tilefish stock north of Cape Hatteras do not provide information about whether the stock is overfished. Given the lack of information about abundance, particularly in the northern part of this stock's range, as well as its data-poor condition in general, I agree with the Analytical Team's decision that a determination about whether the stock is overfished cannot be made at this time.

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

A key uncertainty in the evaluation of whether the stock south of Cape Hatteras is undergoing overfishing is the lack for information from the northern part of its range. However, the analyses provided in the Assessment Workshop report do provide an indication that localized depletion could occur in the southern part of its range from which most of the landings are presently being taken. Although there is high uncertainty in this evaluation, the average catch for the late time period (AvC.late: 2006-2015) had a median of 474,000 lbs. In contrast, the three methods that provide catch recommendations based on MSY approximations, Fdem_ML, SPMSY, and YPR_ML, provided catch recommendations with medians ranging from 110,000 to 310,000 lbs. The Analytic Team also provided an analysis comparing the annual removals per unit habitat area between the area just north of Cape Hatteras and those for the southern stock. This analysis also indicated that the removals in the area just north of Cape Hatteras are high relative to removals experienced by the southern stock. Although I agree that these two lines of evidence are suggestive that the high removals over a small area could lead to localized depletion, given the uncertainty in the data inputs, recruitment dynamics and dispersal processes, and the proportion of the northern stock that inhabits this area, I agree with Assessment Workshop conclusion that a determination of the effects on the entire stock north of Cape Hatteras cannot be made with the available information.

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

I agree with the Review Workshop consensus that an informative stock recruitment relationship is not available for the stock south of Cape Hatteras. An estimate of the steepness of the stock recruitment relationship, obtained from a meta-analysis of species with similar life histories, was provided to the model. For the DLM methods that use mean length to estimate total mortality, the unfished equilibrium recruitment can be inferred, but given the uncertainty in the selectivity of the fishery as well as in the stock-level representativeness of the sampling upon which the length estimates are based, it is not clear that this relationship accurately reflects the productivity of this stock. This does not hamper the qualitative conclusions made about the status of this stock, but I do not believe the available information is sufficient to make reliable predictions of 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?

Quantitative estimates of status determination criteria for the North of Cape Hatteras stock were not provided by the DLM analyses. Given the limited data available for this stock, I believe these results are best interpreted qualitatively.

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?***

<u>South of Cape Hatteras</u>

Projections for the stock south of Cape Hatteras were undertaken using the ASPIC model output, and using the age-structured model output for comparison with the ASPIC model results. Using the ASPIC model output, projections were run at three levels of fishing mortality (F_{MSY}, $F_{current}$ and F_{target}) and uncertainty in the projections was quantified via stochastic projections that extended the bootstrap fits from the assessment model. Results were summarized using central tendencies, 5th and 95th percentiles from replicate projections (from the bootstrap results), and as probabilities of the stock being overfished or of overfishing occurring. These methods are consistent with accepted practices and the available data. Projections based on the age-structured model included $F_{current}$, F_{MSY}, $F_{30\%}$ and $F_{40\%}$. The latter two reference fishing mortality rates were added because a specification of a stock-recruitment relationship is not required to calculate

these rates, although one is used in the projections. This approach is also consistent with best practices and the available data.

Both the ASPIC and age-structured model projections were carried out using the same underlying dynamical equations as were used in the assessment model. As such, they are appropriate for the assessment model and output.

When interpreted in the context of their uncertainties and the broad range of sensitivity analyses, I believe the results of the projections are informative, robust and useful to support inferences of probable future conditions. Uncertainty in the projections may result from estimation error, from uncertainty in the data inputs, or from model structure. Sensitivity analyses undertaken as part of the Review Workshop indicated that uncertainty about model structure and about the landings spike in the 1980's had the greatest effect on abundance, status and projected landings. When projected using the reference age-structured model and a fishing mortality rate equating to $F_{40\%}$, point estimates of 2017-2020 yields are roughly 2.5 times higher than those projected using the base ASPIC model run when fishing at F_{target}. In contrast, the age-structured model run with the catch spike removed produced projected yields at $F_{40\%}$ about 2.5 times lower than the base ASPIC model run for the same time period. This level of uncertainty is large relative to the uncertainty associated with parameter estimation in the ASPIC base model run. Additionally, the sensitivity analyses projections also have associated parameter estimation uncertainty that was not fully explored during the Review Workshop, so the factor of 5 differences in the sensitivity projections does not capture the full uncertainty in future conditions.

North of Cape Hatteras

Projections were not undertaken for the stock North of Cape Hatteras. I believe this decision was appropriate because the data limited tools used in the assessment of this stock were not implemented to determine status, but rather to place recent landings in the context of MSY proxies estimated using some of the DLMtools. Although long-term projections can be undertaken within DLMtoolbox, they are conditional on specific management procedures, none of which have been adopted for this stock.

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?***

In my opinion, there are five main sources of uncertainty in this assessment: stock structure, recruitment processes, data inputs (the landings, life history parameters, and indices), model selection, and parameter estimation uncertainty. For the stock south of Cape Hatteras, the first

four of these sources were explored using sensitivity analyses during either the Assessment Workshop or the Review Workshop, and the fifth using the assessment models.

As discussed under TOR 1, there may not be a way to delineate Blueline Tilefish into discrete stocks that can be considered closed populations. The assessment team did explore models in which landings north of Cape Hatteras were included in the model for the southern stock, but could not obtain plausible model fits. While this does not fully address the uncertainty associated with stock structure in the assessment, given the limited information on the movement and dispersal available for Blueline Tilefish, I am not aware of a better approach.

South of Cape Hatteras

The primary issue with modeling stocks that are not closed populations is that recruitment (in the case of Blueline Tilefish) can depend on the abundance and productivity of other stocks. One way to approach this issue is to use a model that estimates annual year-class strength independent of an underlying stock-recruitment relationship. At the request of the Review Workshop, the Analytical Team did attempt to fit the ASPM model while estimating annual recruitment deviates, but, not surprisingly, the data did not appear to be sufficient to support their estimation. While this attempt did not address the uncertainty in the recruitment dynamics, the suggestion by the Review Panel to use $F_{40\%}$ as a reference point for the ASPM model does provide a status determination criteria that would be robust to the overestimation of the steepness parameter that would be expected to occur if a significant portion of the recruitment occurs via larval drift from other stocks.

As discussed under TOR 2, the Analytical Team choose to use ASPIC as the primary assessment model, because of the limited amount of information about life history parameter values for Blueline Tilefish and the lack of age data that could be used in the model. Uncertainty associated with the use of this model was explored by using ASPM as a second model. Although the results from the two models are not directly comparable, the base ASPIC model run and the reference ASPM model run did give results that appear roughly similar with respect to abundance trends and scale. In my opinion, this approach addresses the uncertainty associated with model selection.

Uncertainty in the parameter estimates from the ASPIC model was evaluated using 1,000 bootstrap runs for each model. This is a standard approach with this model.

During both the Assessment and Review workshops, the Analytical Team undertook a large number of sensitivity runs to address uncertainty associated with the data inputs, including selection and weighting of the abundance indices using ASPIC; and uncertainty in the landings data (alternative catch histories) and life history parameter estimates using ASPM. The Analytical Team originally fit separate models to the commercial handline and longline CPUE time series, and combined the bootstrap results from the two separate models. The Review Panel suggested that if both indices were considered equally plausible, they could be included in the same model with equal weighting. I agree that the bootstrap analysis on the single model with

both indices is better than combining the bootstrap results from two model runs. An alternate approach would be to carry the results of models fit to each index separately forward as separate possible states of reality.

The Analytical Team also undertook a large number of sensitivity runs using the ASPM during the Review Workshop. These included sensitivity to different life history parameter values (lower steepness, lower natural mortality, higher age at maturity), different catch histories (reducing the peak landings in the early 1980's to 0.5 and 0.1 times their value), an attempt to estimate recruitment deviates (discussed above), fitting the model primarily to the length-frequency data by down-weighting the abundance indices, and removing the last three years of the abundance indices. The results of these analyses showed that the assessment results are most sensitive to the assumptions about catch history. As discussed under TOR 3, reducing the magnitude of the peak landings leads to a change in estimated abundance that lowers catch recommendations.

Overall, I believe the methods used to evaluate uncertainty in the assessment of the stock south of Cape Hatteras reflect and capture the majority of the sources of uncertainty in the population, data sources, and assessment methods. I believe the implications of the uncertainty in the technical conclusions are clearly illustrated throughout the report.

North of Cape Hatteras

Because the Blueline Tilefish stock north of Cape Hatteras is data-limited, uncertainty in the data sources and assessment results is relatively high. The majority of the landings come from the southern part of this stock's geographic range and little information is available about its abundance in the northern part of its range. As is the case with the stock south of Cape Hatteras, recruitment dynamics are not well understood. Uncertainty in the life history parameter values is partially addressed via Monte Carlo simulation. For each population simulation, life history parameter values are drawn from uniform distributions intended to be representative of uncertainty in the parameter values. This approach is reasonable, but does not address the issue that all combinations of parameter values are not equally plausible due to correlation among parameter values (e.g., among growth parameters; between steepness and natural mortality).

The primary technical conclusion for this stock is that recent landings in the southern part of this stock's range are potentially high. I agree with the Review Workshop conclusion that this conclusion is highly uncertain and that the results are best interpreted qualitatively, particularly in terms of the effect on the entire stock. I believe the uncertainties in these conclusions and results of the assessment for this stock are clearly documented in the Assessment Workshop and Review Workshop reports.

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.***

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

Research recommendations were provided by the Data and Assessments workshops and were reviewed at the Review workshop. In my opinion, the four main areas where further research would help improve the assessment for both stocks: development of fishery independent indices, resolving the age determination issues, research on recruitment dynamics, and research towards better understanding the fishing process.

The models for the stock south of Cape Hatteras were fit to fishery-dependent indices, and, as discussed above, there is uncertainty about the potential for hyperstability to be an issue with these indices. Additionally, all indices were truncated because of changes in fishing practices and regulations such that none of the indices provided information about recent relative abundance. Given these changes, it appears unlikely that a reliable and current fishery-dependent index covering the range of each stock will become available in the near term. For this reason, I strongly endorse the Index Working Group's (IWG) unanimous consensus that a coast-wide fishery-independent survey is needed for Blueline Tilefish. To me, this is the top research priority. However, the IWG also recommended that, in the absence of a fishery-independent index, additional information on the targeting behavior of fishermen, in particular the depth or geographic locations fished within a given trip as well as more refined information on fishing effort, is needed to improve the assessment. It is less clear to me that this information would sufficiently replace a fishery-independent index unless the same areas are consistently fished in the long term. However, it would help to determine whether the fishery-dependent indices do track changes in relative abundance. Additionally, if there are areas of higher abundance due to the limited movement of adult Blueline Tilefish and less recent fishing effort in the area, and if fishermen are able to target these areas, the age- and length-composition data from the fishery would be unlikely to be representative of the stocks. Research on gear selectivity as part of the fishery-independent surveys would greatly benefit the assessments for both stocks.

Reliable age-composition data are like gold for a stock assessment, and although I agree with the decision not to use the age data in this assessment, the decision limited the options available for modelling these populations. For this reason, I also strongly endorse the research recommendations about resolving the aging issues, if they can be resolved. As described, issues with the age determinations include identification of annuli in opaque zones in the otolith, identification of the first annulus, and variation in the increment count depending on the area of the otolith that is read. If these issues can be resolved, a lot of information about the age-

composition of the landings, as well as information about growth would become available. However, given the within and among lab variation in age determinations, as well as the variation within individual readers, it is not clear to me that these issues will be fully resolved. Additionally, if a reliable method for age determination is developed, the application of the age-composition data in the assessment model will warrant careful consideration. The data may be representative of the age-composition of the landings, but if not necessarily of the age structure of the stock. Towards the objective of resolving the aging issues, research recommendations pertaining to age validation methods (e.g., further investigate the potential shift in the Radio Bomb Carbon data and reference curve for Blueline Tilefish age validation) are, in my opinion, the highest priority. Standardized aging methods following best practices are also necessary.

Although supporting information is limited, adult Blueline Tilefish are thought to be relatively immobile, dispersal is thought to occur in the egg and larval life stages via drift, and dispersal among areas may be highly asymmetrical depending on ocean currents. If so, it is unlikely that stocks can be defined spatially such that the ideal condition that the stock is closed is met. If so, delineation of stocks (or sub-stocks if Blueline Tilefish is considered a single stock throughout its range) will likely be based more on practical aspects, such as management boundaries or the spatial scale on which fisheries occur, as was done in SEDAR 50. In this context, research that helps identify the flow of eggs and larvae among the putative stocks (connectivity) would be expected to lead to a better understanding of the effects of harvesting on abundance and productivity of the stocks. Therefore, I endorse the research recommendations from the Data and Assessment workshops pertaining to connectivity and recruitment dynamics, albeit with a preference for field studies rather than those based solely on the analysis of existing genetic material. These include: research to evaluate spawning season duration and pelagic egg/larval stage duration; research to describe movements/migration of adult Blueline Tilefish; taxonomic work on the identification of Blueline Tilefish eggs and larvae; research on the location, duration, and dispersal mechanisms of the egg and larval stages; particle modeling to investigate hypotheses about movement of eggs and larvae; design and implementation of ichthyoplankton surveys to investigate larval transport; mining of existing ichthyoplankton collections for presence of Blueline Tilefish larvae; and collection of information/data on reproductive and larval behavior for use in modelling larval dispersal.

Most of the research recommendations from the Assessment Workshop pertain to addressing data issues, rather than analytical problems, correctly reflecting the uncertainties and limits placed on the assessment by the limited information about the ecology of Blueline Tilefish and limited amount of data. If dispersal among stocks is both high and asymmetrical, development of models and stock status determination criteria that explicitly incorporate connectivity is recommended.

Selectivity curves assumed for SEDAR 50 were asymptotic, although dome-shaped selectivity curves may be more appropriate for hooks. If aging issues remain unresolved and abundance indices indicative of recent abundance are not developed, the assessment will become more data

limited. In this case, research about the selectivity of the fisheries would be expected to aid in the application of methods based on mean length, or changes in mean length.

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

Overall, I think that the SEDAR process provides for a thorough review and evaluation of the available data, provides for thorough consideration and review of analytical approaches and modeling results, provides very good guidance on the information expected to result from the process, and provides very good documentation of the process including decisions made throughout the assessment. I particularly like that all documents remain unchanged between workshops and that changes are documented via updates in the next workshop report or via an addendum. This makes for a highly transparent process. As a minor recommendation, I think that scheduling time for (at least) two rounds of review of the Review Workshop report by the review panelists would help ensure that additions and edits suggested by individual panelists appropriately reflects the consensus opinion of all panelists.

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

In my opinion, the knowledge base pertaining to the ecology of Blueline Tilefish as well as the available data and its associated issues all increased after SEDAR 32, but this resulted in a more data-limited assessment. Abundance indices are now nearly a decade out of date and have uncertainty typical of many fishery-dependent indices; age data are not considered reliable; and the assessment is a model fit to data to 2007 with the results projected forward to 2015 using the landings data. Without a new data source that is indicative of the stock's response to fishing, the uncertainty in the assessment will become greater with the further passing of time. The scheduling of the next assessment will depend in part on the source of new information. If the stocks become more data limited, such that only data limited methods are applicable (or the same methods are used with data-limited methods used as supporting information), then completing research on gear selectivity and on targeting by the fisheries would aid in determining whether methods based on mean length are applicable for these stocks. If the aging issues are resolved, then quite a large amount of older data will become available, enabling the use of models fit to age-composition data that could be developed in the near-term. If a new fishery-independent abundance index becomes the main data source for the assessment, then several years of data collection may be required prior to it becoming informative about the stock.

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.***

Preparation of the Peer Review Summary Report was nearly completed at the time of submission of this report. At the meeting, a list of tasks to be completed following the workshop was developed. This list included final analyses to be carried out by the AT, and writing tasks for the Summary Report which were assigned to the Review Panel members. A draft Review Workshop Summary Report was completed and had undergone one round of review at the time of the submission of this individual report.

4.0. Conclusions and Recommendations in Accordance with the TOR's

SEDAR 50 was an assessment of Atlantic Blueline Tilefish. The unit stock for SEDAR 50 included Blueline Tilefish along the entire Atlantic Seaboard, using the boundary between the GMFMC and SAFMC management area as the southwestern boundary for the stock unit. Atlantic Blueline Tilefish were assessed as two putative stocks, the stock to the south of Cape Hatteras, and the stock to the north. The main conclusions and recommendations from my review are summarized in this section of this report.

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?***

The Data Workshop Report, Assessment Workshop Report and supporting working papers thoroughly documented the information available for assessing Atlantic Blueline Tilefish. Overall, I consider the data decisions made by the Data Workshop and Assessment Workshop to be sound and robust, and uncertainties are acknowledged and within expected levels. The data are applied appropriately within the assessment model and the input data series are sufficient to support the assessment approach and findings. However, due to the high uncertainty in some of the inputs, the sensitivity analyses produce a wide range of biomass estimates.

Stock Identification

The SEDAR 50 Stock ID Working Group considered genetics, life history data, adult distributions, oceanographic features and data on drifter movement when delineating stocks of Blueline Tilefish in US waters. These are well described in the Assessment Workshop Report.

Limited information about the stock structure, coupled with a spatial mismatch between the CPUE indices and recent removals, led to a decision that Blueline Tilefish in the Atlantic be modelled as two stocks, one extending from the SAFMC/GMFMC boundary to Cape Hatteras, and the other extending north from Cape Hatteras. I consider this a practical decision that allowed the assessment to proceed logically, although it may not meet the assumption of the assessment models that the stock is not influenced by dispersal from other stocks**. Research about recruitment dynamics, specifically egg and larval drift, is recommended.**

Age composition

Age composition data were not considered sufficiently reliable for use in SEDAR 50. This decision, despite being what I consider the correct decision, had major implications for this assessment. Resolution of the issues about age determination would allow a lot of older information to be used in the assessment of this stock. **Resolution of the age determination issues, if possible, is one of the top two research recommendations I have for this stock.**

Life history data

Overall, I believe the decisions about the life history data were well thought out, are well described in the Data Workshop and Assessment Workshop reports, and are defensible. Particularly given the decision that the age determinations were not reliable, Blueline Tilefish life history parameter values are not well known. With respect to their application in the assessment models, they are not used in the ASPIC model, and sensitivity to assumed values was explored using ASPM. **In models that draw life history parameter values from assumed distributions, such as the DLMtool, choosing parameterizations in which the values are not dependent on other life history parameters (e.g., using the slope at the origin of a Beverton-Holt stock recruitment model rather than steepness) is recommended to reduce the effects of parameter correlation that may not be otherwise addressed.**

Abundance indices

For the stock south of Cape Hatteras, the Data Workshop Report and background documents provided thorough descriptions of the available abundance indices. Indices used in the assessment were all fishery-dependent CPUE indices. None of the indices provide information about relative abundance since 2007 and could be considered outdated. Ideally, an abundance index is collected using an experimental design that ensures the resulting index is proportional to abundance. Commercial CPUE data are not typically collected in this way and if fisheries are able to target areas of higher abundance, or if the efficiency of the fishery increases over time, hyperstability in the index may occur. Although there are potential issues with the CPUE indices, overall I believe the decisions made by the Assessment Team about these data are sound, and that they are sufficient for this assessment.

Abundance indices were not available for the Blueline Tilefish stock north of Cape Hatteras.

I consider the development of fishery-independent abundance indices as the top recommendation for improving the assessment of both stocks.

Length composition

The length composition data played a relatively small role in the assessment of the stock south of Cape Hatteras. ASPIC does not use these data, and the ASPM model output appeared to be weighted towards the abundance indices. Although this was not fully explored at the Review Workshop**, the apparent discrepancy between the signal provided to the ASPM by the length composition data and that provided by the abundance indices warrants further exploration.** Particularly if the length-composition data are influenced by targeting of areas subject to lower past exploitation, how is the data are used in an assessment model warrants careful consideration, a consideration that may also apply to the age composition data if the aging methods are worked out.

Recreational and commercial landings

The derivations of the recreational and commercial landings are well described in the Data Workshop Report and in the supporting documentation. Prior to 1985, tilefish landings were not reported by species. The methods used to determine the portion of Blueline Tilefish in the landings involved determining the proportion of the landings in a later time period and applying that proportion to the earlier time period. For the stock south of Cape Hatteras, the method resulted in a large spike in landings just prior to 1985, which lacked credibility with stakeholders present at the Review Workshop. This appeared to be the key source of uncertainty in the assessment for the southern stock. **If possible, exploration of alternative methods for partitioning historic landings by species is recommended.**

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?***

South of Cape Hatteras

The Analytical Team used two models for the stock south of Cape Hatteras: a biomass dynamics model (ASPIC), and a statistical, age-structured model (ASPM). The ASPIC model was fit to combinations of three CPUE abundance indices. The ASPM analyses were fit to the same CPUE abundance indices, as well as the length-frequency data, and also used information about life history. The ASPIC model better matches the data available for this stock, but models biological processes in a restrictive way that my not reflect the true dynamics of this stock. The ASPM model better matches the underlying life history and fishing processes, but requires that more

assumed information than is required for ASPIC. I agree with the Review Workshop consensus preference for the ASPM as the better model for this stock.

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

I consider the two modelling approaches used for the Blueline Tilefish stock south of Cape Hatteras to be sound. Both models have been used extensively in previous assessments, are based on accepted scientific principles, and the application of both models followed accepted scientific practices. The large number of sensitivity analyses to evaluate the robustness of the conclusions is one of the strengths of this assessment.

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

In my opinion, the models were, for the most part, configured appropriately. The Review Workshop recommendation for the ASPIC model to weight the longline and handline CPUE indices equally in a single model, rather than to average results from two models fit to the indices separately is more consistent with accepted practices. The recommended modification to the reference ASPM model to move the age-at-maturity from age-2 to age-6 is a practical recommendation to ensure that the stocks resilience is not over-estimated. Neither model would be appropriate if recruitment is significantly influenced by egg and larval drift from other stocks. However, the analyses presented are appropriate because of the high uncertainty about the connectivity of the stocks, and because data are not sufficient to estimate year-class strength without the assumed relationship. **Because the ASPM model can be configured to better match the biological and fishery processes that affect the productivity and status of the stock, the use of an age-structured model is recommended for future assessments, even in the absence of age composition data.**

Research about the selectivity of the fishing gear is recommended.

c) Are the methods appropriate for the available data?

I consider the methods used in the assessment of the Blueline Tilefish stock south of Cape Hatteras appropriate for the available data. In my opinion, the results of this assessment are best applied considering the suite of results from both models. **Recommendations to improve these models pertain primarily to the data inputs rather than the models themselves.**

<u>North of Cape Hatteras</u>

The primary analytical tool used for the Blueline Tilefish stock North of Cape Hatteras was the Data-Limited Methods Toolkit. This program was used for an evaluation of potential assessment methods, to develop catch recommendations based on the selected methods, and to compare these recommendations with recent landings.

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

Because there were few data available for this stock (only the removals time series and a small amount of length-composition data), and because the methods have been previously reviewed, I consider the use of this tool use is appropriate for this stock.

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

The Analytical Team selected three methods that used MSY approximations and compared the results from these methods with scenarios based on average catch. The appropriateness of the length-based mortality estimator used in two of the methods may be questionable because whether the length-composition data are representative of the entire stock, and whether the selectivity of the fisheries matches the assumed selectivity, is not known. However, if the results are interpreted qualitatively and in the context of this uncertainty, then this application is appropriate. **Research about the selectivity of the fishing gear is recommended to aid in the selection of appropriate data-limited methods.**

c) Are the methods appropriate for the available data?

I believe that the methods used are 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?***

<u>South of Cape Hatteras</u>

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?

In my opinion, the population estimates provided for the stock south of Cape Hatteras are reliable, consistent with input data and population biological characteristics and useful to support status inferences. However, the uncertainty in the population estimates and status inferences is high, and the results should be interpreted in the context of the full suite of analyses undertaken for this stock.

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

Overall, I believe the assessment results provide evidence that the stock south of Cape Hatteras is not overfished.

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

In my opinion, the results of this assessment provide evidence that the stock south of Cape Hatteras is not undergoing overfishing, although this conclusion has greater uncertainty than whether the stock is in an overfished state.

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

Both the ASPIC and ASPM models used in this assessment assume a stock-recruitment or biomass-production relationship. The Review Workshop was presented with considerable information suggestive that recruitment is partially driven by immigration and emigration. If recruitment is largely dependent on abundance of the other stocks, then assuming stock-recruitment relationships for the individual stocks would be inappropriate. However, recruitment dynamics for Atlantic Blueline Tilefish are poorly understood and the magnitude of dispersal among stocks is not known. I consider the use of these relationships appropriate in this assessment for this reason. **Research about recruitment dynamics is recommended to determine the appropriate spatial scale on which the assessment should be undertaken and whether these putative stocks can be modelled as closed populations.**

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?

Overall, I believe that the quantitative estimates of status determination criteria are appropriate for management use for this stock. The status determinations described above were generally robust to many sensitivity analyses undertaken using both ASPIC and ASPM, although the strength of the conclusions about whether overfishing is occurring and whether the stock is in an overfished state differs among model runs. **With respect to status determination criteria for determining whether the overfishing is occurring, the use of SPR-based reference points ($F_{x\%}$) is recommended** because the stock-recruitment relationship does not need to be known or assumed to estimate SPR.

North of Cape Hatteras

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?

Nearly all the data available for the northern Blueline Tilefish stock are from the fishery in an area just north of Cape Hatteras where the majority of the landings are taken. As such, the extent to which the data are representative of the entire stock is not known. Additionally, abundance,

biomass and exploitation rates were not provided via the analyses undertaken for this stock, which I consider appropriate given the high uncertainty for the inputs for the analyses. Analyses may be sufficient for qualitative interpretations about whether there is potential for localized depletion.

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

The methods used to assess the Blueline Tilefish stock north of Cape Hatteras do not provide information about whether the stock is overfished. I do not believe that a determination about whether the stock is overfished can be made with the available information.

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

A key uncertainty in the evaluation of whether the stock south of Cape Hatteras is undergoing overfishing is the lack for information from the northern part of its range. The Assessment Workshop Report provided two lines of reasoning indicating that localized depletion could occur in the southern part of its range. Although I agree that these two lines of evidence are suggestive that the high removals over a small area could lead to localized depletion, given the uncertainty in the data inputs, recruitment dynamics and dispersal processes, and the proportion of the northern stock that inhabits this area, I agree with the Assessment Workshop conclusion that the effects on the entire stock north of Cape Hatteras cannot be determined with the available information.

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

I agree with the Review Workshop consensus that an informative stock recruitment relationship is not available for the stock south of Cape Hatteras. This does not hamper the qualitative conclusions made about the status of this stock, but I do not believe the available information is sufficient to make reliable predictions of 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?

Quantitative estimates of status determination criteria for the stock north of Cape Hatteras were not provided by the DLM analyses. I believe the results are best interpreted qualitatively.

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?***

South of Cape Hatteras

In my opinion, the projection methods used for the stock south of Cape Hatteras are consistent with accepted practices, and are appropriate for the assessment model and the available data. Both the ASPIC and age-structured model projections were carried out using the same underlying dynamical equations as were used in the assessment model. Using the ASPM, projections were also undertaken using SPR-based reference points which are appropriate given the uncertainty in the recruitment processes. **If the ASPM model implemented in ADMB becomes the standard assessment model for this stock, the use of the MCMC capabilities within ADMB to characterize uncertainty, followed by running projects using the MCMC output to ensure the preservation of covariance between models parameter as well as other model output is recommended.**

In my opinion, key uncertainties in the projections are acknowledged, discussed and reflected in the projection results. This uncertainty is high, but not overly so given the uncertainty in the data inputs. When interpreted in the context of this uncertainty, I believe the projection results are informative and robust, and useful to support inferences of probable future conditions.

North of Cape Hatteras

Projections were not undertaken for the stock north of Cape Hatteras, a decision I consider appropriate given the use of the model**. If the DLMtool is to be the primary assessment method for this stock, then evaluation and selection of an appropriate management procedure via MSE projections is recommended.**

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?***

South of Cape Hatteras

In my opinion, there are five main sources of uncertainty in this assessment: stock structure, recruitment processes, data inputs, model selection, and parameter estimation uncertainty. For The first four of these sources were explored using sensitivity analyses during either the Assessment Workshop or the Review Workshop, and the fifth using the assessment model.

The primary issue with modeling stocks that are not closed populations is that recruitment (in the case of Blueline Tilefish) can depend on the abundance and productivity of other stocks. **Exploration of models that estimate annual year-class strength independent of an underlying stock-recruitment relationship, coupled with the use of SPR-based reference points, is recommended** to address this uncertainty.

Sensitivity to the CPUE indices was addressed by fitting the ASPIC model to various combinations of the three available indices. This is appropriate for this model. Uncertainty associated with the use of this model was explored by using ASPM as a second model. This approach addresses the uncertainty associated with model selection. Uncertainty in the life history parameter values, as well as the early 1980's landings spike, was explored using the ASPM. I agree that uncertainty in the landings time series is the key uncertainty in the assessment. If possible, **exploration of alternative methods of partitioning the landings data by species is recommended**.

Uncertainty in the parameter estimates from the ASPIC model was evaluated using 1,000 bootstrap runs for each model**. If the ASPM model implemented in ADMB becomes the standard assessment model for this stock, the use of the MCMC capabilities within ADMB to characterize uncertainty is recommended.**

In my opinion, key uncertainties are acknowledged, discussed and reflected in the assessment results. This uncertainty is high, but is not in consideration of the uncertainty in the data inputs.

North of Cape Hatteras

The results of the assessment for the stock north of Cape Hatteras have high uncertainty. These are appropriately documented in the Assessment Workshop and Review Workshop reports.

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.***

Research recommendations were provided by the Data and Assessments workshops and were reviewed at the Review workshop**. In my opinion, the four main areas where further research would help improve the assessment for both stocks: development of fishery independent indices, resolving the age determination issues, research on recruitment dynamics, and research towards better understanding the fishing process.**

Research towards meetings these objectives include:

- Development of a coast-wide fishery independent survey
- Studies, or collecting data pertaining to the targeting behavior of fishers
- Studies of the selectivity of the fishing gear (this could be done as part of a survey)
- Research pertaining to age validation methods
- Standardization of aging methods following best practices
- Research that helps identify the flow of eggs and larvae among the putative stocks (connectivity and recruitment dynamics)

- Evaluation of spawning season and pelagic egg/larval stage duration
- Tagging studies to identify movements/migration of adult Blueline Tilefish;
- Taxonomic work on the identification of Blueline Tilefish eggs and larvae;
- Research about the location, duration, and dispersal mechanisms of the egg and larval stages,
  and
- Particle modeling to investigate hypotheses about movement of eggs and larvae.

Overall, I believe that SEDAR is a very thorough and transparent process. A very minor recommendation that I don't believe is an issue for this process is to ensure adequate time for more than one round of review of the Review Workshop Report in the event that consensus is not reached after one round.

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

In my opinion, SEDAR 50 provided a comprehensive presentation of the information available for Atlantic Blueline Tilefish, but the assessment is still data-limited for both stocks. Abundance indices are now nearly a decade out of date and the assessment is primarily a model fit to CPUE indices to 2007 with the results projected forward to 2015 using the landings data. Age determinations are presently considered unreliable. **Without a new data source that is indicative of the stocks' response to fishing, the uncertainty in the assessment will become greater with the further passing of time.** In the absence of this information, research about completing research on gear selectivity and on targeting by the fisheries would aid in determining which data-limited methods are applicable for these stocks. If a fishery-independent survey is developed, sufficient time will need to pass to allow the survey to become informative before it can be included in the assessment. If the aging issues are resolved, then a large amount of older data will become available and the next assessment could occur more rapidly.

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 draft of the Review Workshop Summary Report was completed and had undergone one round of review at the time of the submission of this individual report.

5.0. References

Carruthers, T. R., and A. Hordyk. 2016. Package 'DLMtool', version 3.2. 132 pp. Available from: https://cran.r-project.org/web/packages/DLMtool/

Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick, and J. Cope. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153:48-68.

Carruthers, T., L. Kell, D. Butterworth, M. Maunder, H. Geromont, C. Walters, M. McAllister, R. Hillary, P. Levontin, T. Kitakado, and C. Davies. 2015. Performance review of simple management procedures. ICES Journal of Marine Science 73:464-482.

Gedamke, T., and J. M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Transactions of the American Fisheries Society 135:476-487.

Prager, M. H. 2015. User's Guide for ASPIC Suite, version 7. A Stock-Production Model Incorporating Covariates and Auxiliary Programs. Prager Consulting, Portland, Oregon, USA.

Williams, E. H., and K. W. Shertzer. 2015. Technical Documentation for the Beaufort Assessment Model (BAM). NOAA Technical Memoranddum-NMFS-SEFSC-671.

6.0. Appendices

Appendix 1: Bibliography of Materials Provided for Review
Appendix 2: CIE Statement of Work
Appendix 3: Panel Membership

Appendix 1: Bibliography of Materials Provided for Review.

SEDAR 50
Atlantic Blueline Tilefish
Workshop Document List

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 ****SEE SEDAR50-DW19 FOR FINAL REPRODUCTIVE ANALYSES**	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

Appendix 1: Bibliography of Materials Provided for Review.

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	SFB-NMFS 2017

Appendix 1: Bibliography of Materials Provided for Review.

	Gulf of Mexico waters of the U.S. from recreational headboat logbook data	
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 the recreational fisheries	SFB-NMFS 2017
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	SEDAR 32

Appendix 1: Bibliography of Materials Provided for Review.

	Gray Triggerfish) – all documents available on the SEDAR website.	
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 natural hard-bottom habitats from the middle to outer continental shelf of the Middle Atlantic Night near Norfolk Canyon	Ross et al. 2016
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

Appendix 1: Bibliography of Materials Provided for Review.

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 keystone browser, with a novel approach to bomb radiocarbon dating in the Hawaiian Islands	Andrews et al. 2016
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	Andrews et al. 2013

Appendix 1: Bibliography of Materials Provided for Review.

	novel use of postbomb radiocarbon dating in the Gulf of Mexico	
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	Restrepo et al. 1998

Appendix 1: Bibliography of Materials Provided for Review.

	of the Magnuson-Stevens Fishery Conservation and Management Act	
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: 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.

Appendix 2: CIE Statement of Work.

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

Appendix 2: CIE Statement of Work.

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.

Appendix 2: CIE Statement of Work.

NMFS Project Contact:

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

Appendix 2: CIE Statement of Work.

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.

Appendix 2: CIE Statement of Work.

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?

Appendix 2: CIE Statement of Work.

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.

Appendix 2: CIE Statement of Work.

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.

Appendix 3: Panel Membership

Review Panel Membership

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