Report for the Center of Independent Experts on the SEDAR 32 South Atlantic blueline tilefish and Gulf of Mexico menhaden assessment review (August 27 to 30, 2013)

Stephen J. Smith
383 Portland Hills Drive
Dartmouth, Nova Scotia
Canada B2W 6R4

Executive Summary

The Southeast Data, Assessment, and Review (SEDAR) 32 Review Panel met from 27 to 30 August 2013, in Morehead City, NC to review the data and assessments for Atlantic blueline tilefish and Gulf of Mexico menhaden. The panel consisted of a Chair, three CIE reviewers, and two independent reviewers. This was the first assessment for blueline tilefish in the SEDAR process while Gulf of Mexico menhaden was last reviewed at SEDAR 27 in 2011. The results of the age-based and age-aggregated models all indicate that the Atlantic blueline tilefish stock is overfished and overfishing is occurring. The results of the stock assessment were judged to be the best scientific information available, however the reliance on commercial and recreational fishery based indices of abundance limited the assessment team's ability to judge whether the recent increase in landings north of Cape Hatteras was due to a northward shift in distribution or a newly discovered but previously un-fished part of the population. In addition, the lack of a recruitment index made it impossible to verify recruitment estimates from the model that were not supported by the age compositions of the landings.

There was no evidence for menhaden of overfishing or of the stock being overfished given commonly applied benchmarks in the region and based on the results from the age-based and age-aggregated models. Managers are in the process of developing the goals and objectives for the menhaden fishery including biomass and F benchmarks for this fishery. Without established thresholds, it is not possible to provide quantitative estimates of stock status. Landings data for this fishery were of high quality and fishery-independent indices for recruitment and adults were also available for this assessment. The assessment was also of high quality and represents the best scientific information available. More fishery-independent indices may become available for future assessments once a rapid method for resolving species identification has been developed. The lack of older fish in the catch relative to their presence in the Louisiana gillnet index for adult fish was of concern with respect to estimating productivity of the stock.

Background

The review workshop of the 32nd Southeast Data, Assessment, and Review (SEDAR) process was convened in Morehead City, NC from August 27 to 30, 2013. The purpose of the workshop was to review stock assessments for Atlantic blueline tilefish and Gulf of Mexico menhaden. The South Atlantic blueline tilefish stock is within the jurisdiction of the South Atlantic Fisheries Management Council and the state waters of North Carolina, South Carolina, Georgia, and Florida. The Gulf of Mexico menhaden stock is within the jurisdiction of the Gulf States Marine Fisheries Commission and the state waters of Texas, Louisiana, Mississippi, Alabama, and Florida.

The SEDAR 32 Assessment Process was held via a series of webinars from April through July 2013. The pre-assessment webinar was held April 17, 2013. Specific assessment webinar dates were May 8, May 23, June 5, June 19, July 10, and July 24, 2013.

Blueline tilefish had not been assessed in the SEDAR process prior to this assessment while Gulf of Mexico menhaden was last assessed in 2011 at SEDAR 27.

Description of Individual Reviewer's Role in Review Activities

Background information, meeting arrangements and other material were made available to the reviewers on July 29, 2013 either via email or through an ftp site. The menhaden stock assessment report was available as of August 6, while the blueline tilefish stock assessment document was made available on August 9. I reviewed these two main assessment documents accessing the background information from the ftp as necessary to get more detail on the data used or analyses that were carried out. On August 21, I participated in a one-hour conference call with available reviewers and assessment leads hosted by Julia Byrd (South Atlantic Fishery Management Council) and Steve Cadrin (panel chair) to go over arrangements, agenda, etc., and also to go over any questions or clarifications concerning the assessment documents. I identified some issues with the surplus production model results for blueline tilefish that were addressed later in that same week by the assessment team.

The review meeting was held August 27 to 30 at the Crystal Coast Civic Center in Morehead City, NC. The panel review chair assigned me to develop text for the review report sections on the Data term of reference (TOR) (TOR 1 for both species) and Research Recommendations TOR (TOR 6 for blueline and TOR 5 for menhaden) based on my notes and those contributed by other panelists. The other two CIE panelists were given similar assignments. The chair and the two non-CIE panelists were responsible for the compiling all of the text into the final review report.

The first day of the meeting was devoted to the presentation of the material on blueline, while the presentation on menhaden took up most of the second day. On the Thursday, the two assessment teams returned with presentations dealing with their responses to issues and questions that the panel had raised during the original presentations. The panel spent Friday morning drafting the report and reviewing the draft material as a group.

Summary of Findings

SEDAR 32 South Atlantic blueline tilefish assessment review

1. Evaluate the data used in the assessment.

   a) Are data decisions made by the Data Workshop and Assessment Workshop sound and robust?

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

   c) Are data applied properly within the assessment model?

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

The review panel focused attention on the definition of the stock area, accuracy of aging data, the decisions to include age compositions but exclude length compositions from the model, the reliability of the commercial and recreational landings data, and the choice of fishery-dependent abundance indices used in the model.

The management area for blueline tilefish extends from Florida to Rhode Island and all landings of this species in this area were included. Genetic or tagging data are not available for this species and it was assumed that the population would exhibit a Gulf of Mexico/Atlantic split similar to many other species in this same area. However, it was also noted that previous work on the related golden tilefish indicated a split north and south of Hatteras. The fishery for blueline tilefish appears to be focused in a few smaller areas, the locations of which have changed over time and a larger percentage of the recent landings are now coming from North Carolina waters north of Cape Hatteras. This species is also known to burrow in soft bottom habitats and this fine scale structure may result in local depletion. There was some discussion about whether increased landings in the area north of Cape Hatteras represented a previously untapped area for blueline tilefish or were due to a northward change in stock area but there was no information available to decide between the two possibilities.

Age data were obtained from sampling recent commercial fisheries landings that appeared to target a very narrow range of ages (3-5 for recreational and 5-8 for commercial fisheries). There were no age composition data for landings in the earlier part of the series when it was expected that larger/older fish should have been at a higher proportion in the population given the assumption of maximum age of 43 years. The von Bertalanffy growth curve for the recent data indicated that 98% of total growth had been completed by age 15, and therefore ages 15 and older were adopted as a plus group.

Assumptions about the initial age composition raised issues about the current estimates of natural mortality (M) and fishing mortality (F), as well the assumption of flat-topped selectivity. Natural mortality at age was estimated using the methods of Charnov et al. (2012) which are based on estimates of K and L_∞ from Von Bertalanffy growth curves and therefore highly dependent upon the quality of the recent age data. Considerable uncertainty in age determination for blueline tilefish was documented by Harris et al. (2004). A maximum M of 0.15 and a minimum of 0.05 were used for sensitivity training based upon a CV of 54% from the Hoenig method. While scaling the mean rate over the older ages to 0.1 was reasonable given the Hoenig estimate based on maximum age, the lack of fish of age 15 years and older in the recent landings suggests that either M may be higher because the maximum age of 43 is questionable due to the uncertainty in ageing, or fishing mortality was much higher than assumed. This suggests that the higher M alternative should receive more attention in the sensitivity analysis than the lower M, and perhaps M estimates higher than 0.15 should be considered.

Maturity-at-age was based upon estimates for golden tilefish with 50% mature at age 3 and 100% mature at age 4. While these results indicated a relatively younger maturity than may be expected for such a long-lived fish, similar results have been reported for other long-lived species in the region. However, maturity studies of golden tilefish suggest that functional maturity may occur at ages older than histological maturity because of territoriality, dominance and mate choice (Grimes et al. 1988, McBride et al. 2013). If this is also true for blueline tilefish, then the apparent truncation of age composition due to harvesting may result in a decline in the size of males that gain access to the females for spawning. It is not known what impact this decline in size may have on stock productivity.

The available age composition data representing the recent years do not appear to track year-classes even though high recruitment was estimated to have occurred prior to the period that the bulk of these data were collected. This increased recruitment was not actually observed but was estimated by the model to account for recent increases in the adult handline index and recent catches.

While the age compositions were included in fitting the model, the length compositions were removed from the analysis due to preliminary results indicating lack of fit. In light of the uncertainties associated with the ageing data, it seemed strange that the length composition data would not be better fitted by the model. However, sensitivity runs and estimated length compositions from the base run in which the length compositions were not part of the objective function demonstrated that including length composition data resulted in poorer fits to the age compositions and the abundance indices. Varying sampling coverage in time and space was one of the main reasons suggested for the lack of information in the length composition data. The review panel agreed with the assessment team, noting that the residual patterns from model runs with length compositions were not acceptable.

The landing data were considered to be reliable since 1974 and discarding for the commercial fishery was assumed to be negligible, consistent with there being no regulatory reasons for discarding (e,g., size limits). The recreational catch was sporadic and low relative to the commercial catch until 2006. There was considerable discussion about the reliability of the recreational landings estimate for 2006 to 2008, including the very high discard estimates in 2007. Most of these landings appeared to have occurred in North Carolina waters and there was a suggestion that the development of a "deep-drop" fishery may have driven the increase, with the decrease in 2011 due to the implementation of a deep water closure. A quick look at the MRIP data indicated that CVs for 2006 to 2011 decreased relative to the period before and the number of sample intercepts increased, both indicative of increased fishing activity. However, the magnitude of the landings relative to the commercial landings in those same years still seemed to be unprecedented and industry participants questioned the reliability of the recreational estimates.

The commercial and recreational headboat catch rate information were key data for both the Beaufort Assessment Model (BAM) and ASPIC models. These were the only annual abundance indices available and were developed using the standard approach of fitting delta-GLM models to filter out annual trends from other factors associated with these data. The recreational index used here represented the earlier period when the SSB was being fished down but this index actual represents very low levels of catch. There was no overlap between this index and the two commercial indices. A three-year running smooth of headboat catch rate information including data after 1992 was presented, suggesting somewhat similar trends to the commercial indices in the later years.

While the landings data were taken from the whole area, the catch rate abundance indices were confined to data between 28° and 35° N latitude to more reflect the core stock area. As noted above, the model interpreted recent increases in catch and the handline index to be due to high recent recruitment. The validity of this assumption will be important for forecasting future productivity.

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

The Beaufort Assessment Model (BAM), implemented in AD Model Builder software (Fournier et al., 2012) was used to develop a statistical age-based forward projecting assessment model of the population. In addition, two production type models were also fit to the data. An age-based production model was produced using BAM with the recruitment deviations option turned off. An age-aggregated surplus-production model implemented using the ASPIC package (Prager, 2005) was also used for comparative purposes.

The BAM base case model and rationale for modeling decisions are well described in the assessment report and were further explored during the Review Workshop. The base case run included commercial and recreational landings, age composition data and three indices of abundance (recreational head boats, commercial long line and hand line). Natural mortality varied by age and was assumed constant through time. Steepness was fixed at 0.84 based on meta-analyses (Myers et al., 2002; Shertzer and Conn, 2012). Selectivities and catchabilities were all estimated as constant for the full assessment period (1974–2011).

The model was fit to the data using appropriate methods, consistent with standard practice. Analysis included iterative reweighting using the method of Francis (2011) and exploration of a variety of data configurations and parameterizations. The modeling and decisions made to develop the base case run and the sensitivity testing were all well described in the Assessment Report and supporting working documents, and were further elaborated during the SEDAR 32 Review Workshop where additional diagnostics (likelihood components, weights, likelihood profiles) were made available. The modeling procedures adopted appeared to be robust. Landings and indices were fit using lognormal likelihoods. Age composition data were fit using robust multinomial likelihoods. The treatment of the data and the relative importance given to the various components were well explored and appeared appropriate. The model structure was adequate to capture the main patterns in the data.

The production models provided useful comparisons with the base case catch-at-age model results. The main point of difference between these models and the base case was that the production models did not estimate an increase in recruitment in the most recent years, and instead estimated a higher fishing mortality. Despite this difference, the results of the production models did suggest similar stock status to the base case in terms of MSY benchmarks. The production models were appropriately configured and implemented here, and are standard tools for stock assessment.

The Monte Carlo Bootstrapping (MCB) method was used to characterize the uncertainty around the estimates and stock status outputs from the base case model. This method simulates replications of the data using parametric bootstrapping of the landings and indices data, conditional on the distributional assumptions used in the model. The length composition replicate data were resampled from the original data. In addition, values for M and steepness were drawn from probability distribution functions representing possible ranges of likely values for these parameters. Uncertainties were presented as quantiles of the frequency distributions of the various outputs from the model fits to 3043 accepted replicate draws of the above data and parameters. Each individual model fit used the weights developed for the base case run.

There was some discussion about whether all combinations of M and steepness values based on random draws would be biologically appropriate. This is a subject that needs further study for the benefit of this and other assessments that use this technique. It was also noted that the introduction of random variation to M and steepness was essentially adding process error to what was an observation error model fitting approach. Estimates of the management quantities (MSY, B_{MSY}, F_{MSY}) in the base run were estimated using estimates based on a deterministic model structure. The impact of having a stochastic model structure with process error on estimating these management quantities has been investigated for surplus-production models by Bousquet et al. (2008) who showed that F_{MSY} from the stochastic model will be less than the deterministic estimates, estimates of MSY will be higher, and those for B_{MSY} lower. The size of the differences will be a function of the amount of stochastic error in the model. Means of management quantities from the MCB runs do not equal estimates from the base run and differed in the same direction as predicted above for the surplus-production models. While these differences may not be always apparent when comparing ratio benchmarks, for consistency sake, the MCB median estimates of the benchmarks should be used in the ratios for evaluating stock status from the MCB model results.

3. Evaluate the assessment findings with respect to the following:
   a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b) Is the stock overfished? What information helps you reach this conclusion?
   c) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

All of the reviewers agreed that the BAM base run provided the best representation of stock status. The model was evaluated through a series of sensitivity runs that explored a number of issues with the data, model structure and assumptions. The two production models arrived at the same stock status despite interpreting recent changes in stock size differently. The median status results from the MCB run also resulted in the same stock status.

Spawning biomass in 2011 was estimated as 445 thousand pounds., which was less than the estimate of Minimum Stock Size Threshold (489 thousand pounds), so the stock is overfished.

SSB has been below SSB_{MSY} for the past two years (2010–2011). The majority of viable sensitivities runs indicate that the SSB2011 was < SSB_{MSY}. The only exception was the increase in M run which indicated the SSB was greater than the SSB_{MSY}. This was considered unlikely based on additional sensitivity runs requested by the Review Panel. Production model outputs of population status generally agree with the catch-at-age model and indicate a B/B_{MSY} of less than 1 in 2011.

Based on the BAM base run fishing mortality (F) estimates, overfishing is occurring for the South Atlantic Blueline tilefish. The ratio of the geometric mean F over the past 3 years to F_{MSY} was greater (2.37) than 1.0 and has been for the past several years. The dramatic decrease in F_{2011} was primarily the result of a fishery closure. Production model outputs all indicate an average F/F_{MSY} well in excess of 1.0.

The stock/recruitment relationship does not appear to be very informative. There is no information on steepness in the data and there are large positive deviations in the early to mid-2000s to accommodate the increased catches and handline index estimates in the mid to late 2000s. In the terminal three years of the assessment, estimated recruitment did not deviate from the spawner-recruit curve. The recruitment used for the projections was taken from the curve and represents the mean recruitment, not including the high years in the early to mid-2000s. This approach was considered to be reasonable given the data.

The quantitative estimates of stock status appear to be reliable given the agreement on stock status amongst the different models used and the results of the sensitivity runs.

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?

The projection methods used in this stock assessment were consistent with accepted practices in the region and elsewhere, and the available data. Initially the review panel had several concerns regarding the use of MCB approach as a measure of uncertainty. The MCB analysis is considered an approximation of uncertainty for the base run. A number of the limitations were identified in the assessment report. In addition, there was the point raised above about using the median estimates of the management quantities rather than those from the base run to evaluate stock status from the MCB results.

Projection results were informative and robust within the range of observations and inputs from the MCB. Currently F is estimated as the geometric mean of the three previous years. Given the observed rapid changes in F and the preliminary landings estimates for 2012 and 2013, consideration might be given to using actual landings for future projections or drop the 2011 from the estimate of F for 2013 and 2014.

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

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

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

Uncertainty was explored in the assessment modeling using extensive sensitivity runs and likelihood profiling, retrospective analyses and MCB. All of the methods used are standard stock assessment methods. Issues considered in sensitivity runs include variations in M and steepness, alternative maturity vector, adjustment of model weights and exclusion of each series of indices, allowing catchability to vary, inclusion of ageing error, and allowing recreational selectivity to be dome shaped. The sensitivity runs of the base case explored variants of the current model structure but cannot include the impact of other processes such as environmental or geographic effects that are not part of the current structure. However, very useful information was presented on the various sensitivity runs and the panel was satisfied that there had been sufficient exploration of the assessment uncertainties.

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 for blueline tilefish were provided in the data and assessment working group documents. The Panel noted that many of these recommendations reflected concerns across a range of deep-water species and therefore confined their attention to those specific to the stock assessment of blueline tilefish.

While the panel supports work on stock structure, we recommend starting with the available information on describing the differences in demographics/life history characteristics over the range of the management area. Additionally, the available information on habitat in the areas listed should be evaluated before initiating any new studies.

Given that this is an age-based assessment, the comparison and calibration studies for the age determination should receive high priority, along with the marginal increment analysis to determine if the opaque zone is formed annually. Many species would probably benefit from expanding the MRIP program to include age sampling.

The collection of information to better describe spawning season and spawning periodicity could probably start with fishery-dependent sources but will need data from fishery-independent programs to cover the range of the species. The latter program would probably have to be tailored to provide samples across the deep-water snapper/grouper complex.

Studies of discard mortality should be low priority given the current negligible discard rate in the commercial fishery. The collection of additional information on catch (e.g., lengths, ageing material) is important, especially for the areas north of Hatteras, but would likely require an observer program developed for all fisheries focusing on the deep-water snapper/grouper complex.

The BAM model is reliant on historical information, and any data on size compositions, maximum size, etc., obtained from historical recreational fishing photos could be quite useful. One of the main issues raised about the recreational fishery concerned the high landings in the mid-late 2000s, especially the high landing and discard estimates for 2007. Closer scrutiny of these estimates requires data at higher resolution than was apparently available for this stock assessment.

With respect to developing a fishery-independent survey, sampling of deep-water habitats may elucidate habitat characteristics, and spatial distributions of blueline tilefish and other deep-water reef fishes. If a sufficient time series is developed, then a useful fishery-independent index may be available for the stock assessment.

Overall, the material provided to the panel and the presentations made at the SEDAR 32 meeting were of excellent quality. The assessment team members were responsive to all requests made for additional work and provided complete responses to all requests. The amount of material provided for both blueline tilefish and Gulf of Mexico menhaden was extensive and a three and one half day meeting may not have been long enough to consider all of the material to the same level of detail.

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

   The reliance of the assessment on fishery-dependent abundance indices and the lack of a recruitment index were identified as weaknesses of the current approach that could be improved upon. Having an area-wide fishery-independent survey could provide information on geographic changes in distribution and on validation of recruitment trends, both identified as issues with this assessment. While the size of this fishery may not by itself warrant the cost of implementing such a survey, there may be broader advantages in designing a survey for the complex of deep-water species.

SEDAR 32A Gulf of Mexico menhaden assessment review

1. Evaluate the data used in the assessment, addressing the following:
   a) Are data decisions made by the Assessment Workshop sound and robust?
   b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c) Are data applied properly within the assessment model?
   d) Are input data series reliable and sufficient to support the assessment approach and findings?

   The landings were judged to be accurate as the largest portion had been due to the reduction fishery and there has been a log system in place including daily catch records since 1964. Cooperation by industry with supplying information to NMFS is impressive (weekly electronic reporting, 100% participation in the voluntary program, access for port sampling and provision of freezer space for samples). The decision to start the series in 1977 was quite reasonable given the concerns about the data quality for age composition data prior to 1977, inexplicable truncated age distribution in the early 1970s and other issues with these early data as noted in past stock assessments. Sensitivity analyses were used to evaluate the impact of including the longer time series of age compositions.

More information on the spatial distribution of the fishery was requested. The analysis presented on fishery hotspots for 2008, 2009 and 2011 was quite informative and a longer time series would provide information on the spatial overlap between the fishery and the fishery-independent indices used in the assessment. These data may provide insight into the potential distribution of older fish off of western Louisiana and to the east of Alabama/Mississippi; areas not covered by the seine or gillnet survey indices used in the assessment.

The use of population fecundity as a proxy for spawning stock biomass was discussed. A relationship produced in the early 1980s relating numbers of eggs to female length was used in this assessment to estimate length-specific fecundity in the model, thus larger, older fish are estimated to produce more eggs per capita than younger fish. This fish has a protracted spawning season and is assumed to be an indeterminate batch spawner. If older fish produce more batches or higher quality eggs, then their contribution to stock-specific fecundity would be underestimated using the current approach. Ovarian egg number could be a reliable index to SSB if all the ovary samples were at the same stage of reproductive development, but that would seem unlikely to be the case. However, accounting for the relationship between size and fecundity was recognized as a step in the right direction.

There was also discussion about the lack of older fish in the catches being due to the potential for older fish being less vulnerable to the fishery as a function of age-specific spatial distributions. The major grounds for the fishery are within 10 miles of the coast, but the resource distribution is out to 60 miles. Although the fishery may be constrained by spotter planes pilots being reluctant to go offshore, the majority of the stock was considered to be inshore during the warmer months. Based on early-season catches that are further offshore (10–20 miles), there do not appear to be older fish offshore during the fishery. However, the fact that the fishery may target more abundant schools of smaller and younger fish could provide another explanation of the lack of older fish in the catch.

Results from a 2012 study with alternative sampling protocols suggest that sampling the top of hold only did not accurately represent catch, particularly with respect to the presence of older fish in the catch at age estimated from the samples. The study had limited sample size and poor coverage of the port-week strata, and the results suggested that older fish were less than 5% of the catch in the alternative-design samples. However, the lack of older fish in the commercial catch was of concern given that older fish do appear in the gillnet survey used in the assessment.

Several issues were identified with the age data. Multiple age-readers aged fish in the 1960s–early 1970s, but only a single age-reader has aged fish since the 1970s. No formal protocol for ageing quality control appears to exist. Three informal analyses of ageing accuracy or repeatability produced questionable results (e.g., 71% agreement between otolith and scale derived age estimates; 82% agreement between age estimates from scales aged in 2005 and again in 2012; and, substantial disagreement in age estimates from the 1970s versus contemporary re-ageing of those samples). Given the short-lived nature of the fish, reader error of even one year can cause substantial bias in an age-based assessment.

An ageing error matrix was included in the BAM model but was based upon comparisons of age estimates from scales and otoliths, with the otolith ages being assumed to be the true ages. However, there was also error in the age estimates from otoliths, perhaps just as much as in the scales of short-lived species like menhaden. The ageing error matrix also assumed that the error is symmetric about age. In most fishes older fish tend to be under aged with scales as annuli pack at the scale margin and become difficult to discern.

There was evidence of a shift in age composition in the landings from mostly age-1 in the1960–1980s to mostly age-2 in the most recent years. Several hypotheses for the shift were discussed in the assessment report (habitat alteration affecting recruitment of juvenile fish into estuaries, decreased fishing mortality, recent contractions in the spatial distribution of the fishery, changing spatial distribution of age-1 menhaden, influence of hypoxic habitats on spatial distribution) but ageing drift was ruled out based on age determinations from re-reading archived scale samples.

A number of available abundance indices were excluded from being used in the model. A juvenile trawl index, which was highly correlated with the seine index, was included in the SEDAR 27 assessment model, but dismissed here because it was judged that trawls are not efficient for pelagic fish, the spatial extent of the survey was not appropriate for the resource, and the western portion of the survey has species identification problems. A research recommendation was included in the assessment report for genetic sampling by size to solve the species identification problem. The gillnet index used in the assessment was limited to the Louisiana series. Data from the western and eastern portions of the resource area were excluded because of mixed species catches and species identification problems. A larval survey was not used because of poor winter coverage, complex recruitment dynamics from larvae to fishery recruitment, and problems with species identification. Members of the Review Panel questioned why some of these indices were excluded prior to assessing their impact on model fit, such as through likelihood profiling.

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

The Beaufort Assessment Model (BAM), implemented in AD Model Builder software (Fournier et al., 2012) was used to develop a statistical age-based forward projecting assessment model of the population. In addition, an age-aggregated surplus-production model, implemented with the ASPIC package (Prager, 2005), was also used for comparative purposes. The base case model and rationale for modeling decisions were well described in the assessment report and were further explored during the Review Workshop. The base case run included commercial and recreational landings, age and length composition data and two indices of abundance, one representing recruits and the other adult fish. Natural mortality was assumed constant through time but age-specific based on the method of Lorenzen (1996) and scaled based on tagging studies. Steepness was fixed at 0.75. Selectivities and catchabilities were all estimated as constant for the full assessment period (1977–2011).

The model was fit to the data using appropriate methods, consistent with standard practice. Analysis included iterative reweighting using the method of Francis (2011) and exploration of a variety of data configurations and parameterizations. The modeling processes and decision making that resulted in a proposed base case run and sensitivity testing are well described in the Assessment Report and supporting working documents and were further elaborated during the SEDAR 32 Review Workshop where additional diagnostics (Likelihood components, weights, likelihood profiles) were made available. The modeling procedures adopted appear to be robust. Landings and discards were fit closely. Landings and indices were fit using lognormal likelihoods. Age composition data were fit using robust multinomial likelihoods. The treatment of the data and the relative importance given to the various components were well explored and appeared appropriate. The model structure was adequate to capture the main patterns in the data.

The production model provided useful comparisons with the base case catch-at-age model results. The main point of difference between this model and the base case was that the production model did not fit the higher gillnet index estimates in 2008, 2009 and 2011. The higher recruitment deviations estimated by the BAM model to support those years could not be accommodated in the production model with a constant intrinsic rate of growth over the time period. Despite this difference, the results of the production models did suggest similar stock status to the base case in terms of MSY benchmarks. The production model was appropriately configured and implemented here, and is a standard tool for stock assessment.

The Monte Carlo Bootstrapping (MCB) method was used to characterize the uncertainty around the estimates and stock status outputs from the base case model. This method simulates replications of the data using parametric bootstrapping of the landings and indices data conditional on the distributional assumptions used in the model. Replications of the length composition data were resampled from the original data. In addition, values for M and steepness were drawn from probability distributions functions representing possible ranges of likely values for these parameters. Uncertainties were presented as quantiles of the frequency distributions of the various outputs from the model fits to 4068 accepted replicate draws of the data and parameters. Each individual model fit used the weights developed for the base case run.

There was some discussion about whether all combinations of M and steepness values based on random draws would be biologically appropriate. This is a subject that needs further study for the benefit of this and other assessments that use this technique. It was also noted that the introduction of random variation to M and steepness was essentially adding process error to what was an observation error model fitting approach. Estimates of the management quantities (MSY, B_{MSY}, F_{MSY}) in the base run were estimated using estimates based on a deterministic model structure. The impact of having a stochastic model structure with process error on estimating these management quantities has been investigated for surplus-production models by Bousquet et al. (2008) who showed that F_{MSY} from the stochastic model will be less than the deterministic estimates, estimates of MSY will be higher, and those for B_{MSY} lower. The size of the differences will be a function of the amount of stochastic error in the model. Means of management quantities from the MCB runs do not equal estimates from the base run and differed in the same direction as predicted above for the surplus-production models. While these differences may not be always apparent when comparing ratio benchmarks, for consistency sake, the MCB median estimates of the benchmarks should be used in the ratios for evaluating stock status from the MCB model results.

3. Evaluate the assessment findings with respect to the following:
   a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b) Is the stock overfished? What information helps you reach this conclusion?
   c) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

The review panel examined the consistency of the input data and population biological characteristics with the abundance estimates, exploitation, and biomass estimates. Panelists agreed that the base BAM parameterization chosen by the assessment team provided the best representation of stock status and also felt the usage of MCB for projection estimates was appropriate.

Fishery landings were dominated by age-2 fish with fishing occurring after this age group has spawned at least once. However, the selectivity pattern for the reduction fishery was flat topped, and there is uncertainty about the presence of older fish (age-3 and older) in the reduction fishery landings given that they have been observed in fishery-independent gillnet catches.

Very high F estimates were estimated during time series considered, especially during the 1980s. Fishing mortality has subsequently declined to range between 1.0 and 3.5 y^{-1}. The 2011 full F was 2.36 y^{-1}, with much lower F estimates for the older ages.

Currently there are no formal benchmarks established for Gulf menhaden to evaluate stock status. The assessment team presented a suite of potential options for the Review Panel to evaluate. Values of SSB_{2011}/SSB_{MED}, $SSB_{2011}/SSB_{30\%SPR}$, $SSB_{2011}/SSB_{35\%SPR}$, $SSB_{2011}/SSB_{40\%SPR}$ from the BAM base run exceeded 1.0. Results from the surplus production model also estimated SSB_{2011}/SSB_{MSY} to be much greater than 1.0. Therefore, it is unlikely the Gulf menhaden stock would be evaluated to be overfished given commonly applied benchmarks in the region.

F_{MSY} was undefined because all of the fish mature and spawn at least once before being harvested. The surplus production model produced results relative to estimates of MSY with no indication of exceeding the criteria typically used to evaluate overfishing. The review panel agrees with the assessment that it is unlikely the Gulf menhaden stock is experiencing overfishing given commonly applied benchmarks in the region.

Managers are currently defining the goals and objectives for the Gulf menhaden fishery, as well as establishing biomass and F benchmarks. Without established thresholds, it is not possible to provide quantitative estimates of stock status.

4. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
    a) Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
    b) Ensure that the implications of uncertainty in technical conclusions are clearly stated.

Uncertainty was explored in the assessment modeling using extensive sensitivity runs and likelihood profiling, retrospective analyses and MCB. All of the methods used are standard stock assessment methods. Issues considered in sensitivity runs included scaling and the form of M, S-R steepness and form, adjustment of model weights and exclusion of each series of indices, alternative selectivity assumptions for the commercial reduction fishery, change in the start year, alternative weightings and alternative growth specification. The sensitivity runs of the base case explored variants of the current model structure but cannot include the impact of other processes such as environmental or geographic effects that are not part of the current model structure. However, very useful information was presented on the various sensitivity runs and the panel was satisfied that there had been sufficient exploration of the assessment uncertainties.

5. Consider the research recommendations provided by the Assessment workshop and make any additional recommendations or prioritizations warranted.
    a) Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
    b) Provide recommendations on possible ways to improve the SEDAR process.

The panel provided the following comments on the research recommendations that given in the assessment documents.

Several issues were identified with ageing for menhaden including the lack of formal protocols for inter-reader comparisons and calibration/reference data sets. Given the short-lived nature of the fish, reader error of even one year can cause substantial bias in an age-based assessment. Given the pending retirement of the single ager, assessment of the accuracy of ageing and the establishment of formal protocols should be done as soon as possible.

It was not apparent to the panel that stock structure was an issue in the stock assessment and the panel did not see value in undertaking genetic studies on stock structure. However, the panel did see considerable benefit in using simpler genetic techniques such as DNA barcoding to aid species identification, which is currently problematic in peripheral range areas as sampled in the Texas, Alabama, and Florida surveys. Resolution of species identification and any other measures to ensure more consistency across the many state surveys that were excluded from the assessment could provide a more representative basis for monitoring abundance.

The recommendation to consider an aerial survey should be pursued, although the turbid waters close to the Mississippi may limit detectability of fish schools. This kind of survey offers an opportunity to form a partnership between the states, federal government and the fishing industry in a monitoring program to ensure sustainability.

The panel recommended that addressing the sampling of the catch throughout the holds of the reduction fishery vessels be rated as very high priority given concerns about the selectivity of larger fish to the catch. The 2012 study indicated that sampling only the top of the hold may underestimate the proportion of older fish in the catch and given the use of fecundity for spawning stock biomass result in an underestimate of productivity (see below).

While the studies proposed to update knowledge about the reproductive biology of Gulf menhaden would be nice to do, the panel felt that the current approach is adequate for now and more priority should be given to resolving the selectivity pattern of older fish to the fishery so that their reproductive contribution to the population can be better accounted for.

Overall, the material provided to the panel and the presentations made at the SEDAR 32 meeting were of excellent quality. The assessment team members were responsive to all requests made for additional work and provided complete responses to all requests. The amount of material provided for both blueline tilefish and Gulf of Mexico menhaden was extensive and a three and one half day meeting may not have been long enough to consider all of the material to the same level of detail.

6. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

   The Louisiana gillnet survey used in the menhaden assessment has a number of different mesh sizes and concern was expressed about developing a single index over these different mesh sizes, especially given the length frequencies presented in the assessment (Figure 5.44, menhaden assessment). The panel recommends evaluating the efficacy of developing separate indices by mesh size or accounting for the different mesh sizes within the same index.

   The Louisiana seine survey was used as a recruitment index for the menhaden in this assessment. Starting in late 2010, the state has reduced the sampling for this survey to a core set of stations on a quarterly basis due to budgetary reasons and to accommodate other priorities. Given the importance of this survey index to the assessment, the panel recommended that the survey return to the former sampling frequency and geographic coverage.

Conclusions and Recommendations

The results of the age-based and age-aggregated models all indicate that the Atlantic blueline tilefish stock is overfished and overfishing is occurring. The assessment and data were judged to be the best scientific data available for the evaluation of stock status. The stock assessment is completely reliant on commercial and recreational indices for abundance and as a result does not have any information on the stock in areas that are not being fished. The recent increase in landings in the areas north of Cape Hatteras are a case in point where it is uncertain whether this area contains a newly discovered biomass of blueline tilefish that had not been previously exploited or if there has been a general northward movement in the stock due to changing climate conditions. The lack of a recruitment index makes it difficult to verify if the increased recruitment in the mid-2000s estimated by the model was real, especially given the lack of evidence for this recruitment in the age compositions of the landings. Studies on reproductive biology including information on spawning season and spawning periodicity was recommended. The establishment of a fishery-independent survey for the deep-water reef fish complex could in time provide useful information on habitat and distribution of blueline tile fish and help resolve questions about year-class strength.

According to the results of age-based and age-aggregated models the stock status of Gulf of Mexico menhaden was determined to be not overfished and overfishing was not occurring given commonly applied benchmarks in the region. However, the goals and objectives for the Gulf menhaden fishery, including biomass and F benchmarks, are still being defined for this fishery. Without established thresholds, it is not possible to provide quantitative estimates of stock status. The assessment and data were judged to be the best scientific data available for the evaluation of stock status. There were a number of fishery-independent abundance indices considered for this stock assessment but all but two were rejected due to one or more issues of species identification, spatial coverage or seasonal coverage. The lack of older fish (3+ years) in the catch was a concern given the presence of older fish in the gillnet abundance index. Sampling only the top part of the hold of fishing vessels for age and size composition was suggested as a possible reason for the lack of estimates of older fish in the catch. The lack of older fish could also be due to the fishery targeting on the more abundant schools of one and two year olds. Resolving the species identification issue could result in more fishery-independent indices being used in the stock assessment model.

References

Bousquet, N., T. Duchesne, and L. Rivest. 2008. Redefining the maximum sustainable yield for the Schaefer population model including multiplicative environmental noise. Journal of Theoretical Biology 254:65–75,

Charnov. E.L., Gislason, H. and Pope, J.G. 2012. Evolutionary assembly rules for fish life histories. Fish and Fisheries. 14: 213–224.

Fournier, D. A., Skaug, H. J., Ancheta, J, Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27:233–249.

Francis, R. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124–1138.

Grimes, C. B., Idelberger, C. F., Able, K. W., and Turner, S. C. 1988. The reproductive biology of tilefish, *Lopholatilus chamaeleonticeps* Goode and Bean, from the United States Mid-Atlantic Bight, and the effects of fishing on the breeding system. Fisheries Bulletin. 86:745–762.

Harris, P.J., Wyanski, D.M., and Powers Mikell, P.T. 2004. Age, growth and reproduction of blueline tilefish along the southeastern coast of the United States, 1982-1999. Transactions of the American Fisheries Society. 133: 1190–1204.

Lorenzen, K.1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture**.** Journal of Fish Biology. 49: 627–647.

McBride, R.S., Vidal, T.E., and Cadrin, S.X. 2013. Changes in size and age at maturity of the northern stock of Tilefish (*Lopholatilus chamaeleonticeps*) after a period of overfishing. Fisheries Bulletin. 111:161–174.

Myers, R., Barrowman, N., Hilborn, R., and Kehler, D. 2002. Inferring Bayesian priors with limited direct data: applications to risk analysis. North American Journal of Fisheries Management. 22: 351–364.

Prager, M. H., 2005. User's Manual for ASPIC: A Stock-Production Model Incorporating Covariates (ver. 5) And Auxiliary Programs. National Marine Fishery Service, Beaufort Laboratory Document BL-2004-01, Beaufort, NC.

Shertzer, K.W., and Conn, P.B. 2012. Spawner–recruit relationships of demersal marine fishes: prior distribution of steepness. Bulletin of Marine Science. 88: 39–50.

Appendix 1: Bibliography of materials provided for the review

Background documents available on both species available at the ftp site for this review were given listed in the following documents.

SEDAR 32 South Atlantic Gray Triggerfish and Blueline Tilefish Workshop Document (ListS32_BLT_DocumentList_6.4.2013.pdf)

SEDAR 32A Gulf of Mexico Menhaden Document List (S32A_MenhadenDocumentList_8.6.2013.pdf)

Main documents provided for review:

SEDAR 32. South Atlantic Blueline Tilefish. SECTION II: Data Workshop Report. April 18, 2013. 146 p.

SEDAR 32. South Atlantic Blueline Tilefish. SECTION III: Assessment Workshop Report. August 9, 2013 (Revised August 22, 2013). 140 p.

SEDAR 32A Gulf of Mexico Menhaden. Draft Assessment Report. August 2013. A. Schueller, J. Smith, and S. VanderKooy (eds) xxvii + 372 p.

Appendix 2: CIE Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 32 South Atlantic blueline tilefish and Gulf of Mexico menhaden assessment review

BACKGROUND

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

SCOPE

Project Description SEDAR 32 will be a compilation of data, an assessment of the stock, and an assessment review conducted for South Atlantic blueline tilefish and Gulf of Mexico menhaden. The CIE peer review is ultimately responsible for ensuring that the best possible assessment has been provided through the SEDAR process. The South Atlantic blueline tilefish stock is within the jurisdiction of the South Atlantic Fisheries Management Council and the state waters of North Carolina, South Carolina, Georgia, and Florida. The Gulf of Mexico menhaden stock is within the jurisdiction of the Gulf States Marine Fisheries Commission and the state waters of Texas, Louisiana, Mississippi, Alabama, and Florida. . The Terms of Reference (ToRs) of the peer review are attached in **Annex 2a and 2b**.

OBJECTIVES

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the tasks and ToRs described in the SoW herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the scientific peer-review described herein. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall participate during a panel review meeting to conduct the independent peer review in Morehead City, North Carolina, from 27-30 August 2013.

Statement of Tasks: Each CIE reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Tasks prior to the meeting: The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review in accordance with the tasks and ToRs within the SoW. Upon completion of the independent reviewer selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, and FAX number) to the contractor officer's representative (COR), who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment ToRs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COR prior to the commencement of the peer review.

Foreign National Security Clearance: Foreign National Security Clearance will not be necessary for this review because the panel review meeting will be conducted at a non-governmental facility.

Pre-review Background Documents: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the COR the necessary background information and reports (i.e., working papers) for the reviewers to conduct the peer review, and the COR will forward these to the contractor. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

Tasks during the panel review meeting: Each reviewer shall conduct the independent peer review in accordance with the SoW and stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and contractor.** Each reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the stock assessment ToRs as specified herein. The NMFS Project Contact will be responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact will also be responsible for ensuring that the Chair understands the contractual role of the reviewers as specified herein. The contractor can contact the COR and NMFS Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Tasks after the panel review meeting: Each reviewer shall prepare an independent peer review report, and the report shall be formatted as described in **Annex 1**. This report should explain whether each stock assessment ToR was or was not completed successfully during the SEDAR meeting. If any existing BRP or their proxies are considered inappropriate, each independent report shall include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report shall indicate that the existing BRPs are the best available at this time. Additional questions and pertinent information related to the assessment review addressed during the meetings that were not in the ToRs may be included in a separate section at the end of an independent peer review report.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in Morehead City, North Carolina, from 27-30 August 2013.
3) Conduct an independent peer review in accordance with the ToRs (**Annex 2a and 2b**).
4) No later than September 13, 2013, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

22 July 2013	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
12 August 2013	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers.
27-30 August 2013	Each reviewer participates during panel review meeting and conducts an independent peer review
13 September 2013	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
27 September 2013	CIE submits CIE independent peer review reports to the COR
4 October 2013	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COR. The COR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

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

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall 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 main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed. The CIE independent report shall be an independent peer review of each ToRs.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work

Annex 2a: Terms of Reference for the Peer Review

SEDAR 32 South Atlantic blueline tilefish assessment review

1. Evaluate the data used in the assessment, addressing the following:
   e) Are data decisions made by the Data Workshop and Assessment Workshop sound and robust?

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

   g) Are data applied properly within the assessment model?

   h) 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.
   α) Are methods scientifically sound and robust?

   β) Are assessment models configured properly and used consistent with standard practices?

   χ) Are the methods appropriate for the available data?
3. Evaluate the assessment findings with respect to the following:
   f) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

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

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

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

   j) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, addressing the following:
   e) Are the methods consistent with accepted practices and available data?

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

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

   h) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
   - Ensure that the implications of uncertainty in technical conclusions are clearly stated.
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 guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.
8. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. The CIE reviewers are contracted to conduct an independent peer review, therefore the contractual responsibilities of the CIE reviewers do not include the preparation of the Peer Review Summary.
   - Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review.
   - Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

The panel shall ensure that corrected estimates are provided by addenda to the assessment report in the event corrections are made in the assessment, alternative model configurations are recommended, or additional analyses are prepared as a result of review panel findings regarding the TORs above.

Annex 2b: Terms of Reference for the Peer Review

SEDAR 32A Gulf of Mexico menhaden assessment review

1. Evaluate the data used in the assessment, addressing the following:
   e) Are data decisions made by the Assessment Workshop sound and robust?
   f) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   g) Are data applied properly within the assessment model?
   h) 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.
   d) Are methods scientifically sound and robust?
   e) Are assessment models configured properly and used consistent with standard practices?
   f) Are the methods appropriate for the available data?
3. Evaluate the assessment findings with respect to the following:
   f) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   g) Is the stock overfished? What information helps you reach this conclusion?
   h) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   i) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   j) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
   - Ensure that the implications of uncertainty in technical conclusions are clearly stated.
5. Consider the research recommendations provided by the Assessment workshop 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.
6. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.
7. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. The CIE reviewers are contracted to conduct an independent peer review, therefore the contractual responsibilities of the CIE reviewers do not include the preparation of the Peer Review Summary.

- Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review.
- Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

The panel shall ensure that corrected estimates are provided by addenda to the assessment report in the event corrections are made in the assessment, alternative model configurations are recommended, or additional analyses are prepared as a result of review panel findings regarding the TORs above.

Annex 3: Tentative Agenda

SEDAR 32/32A South Atlantic Blueline Tilefish and Gulf of Mexico Menhaden Review Workshop

Morehead City, NC August 27-30, 2013

Tuesday

9:00 a.m. Convene

9:00a.m. – 9:30a.m. Introductions and Opening Remarks **Coordinator**

- Agenda Review, TOR, Task Assignments

9:30a.m. – 12:00p.m. Assessment Presentation and Discussion (BLT*) TBD

12:00p.m. – 1:30p.m. Lunch Break

1:30 p.m. - 3:30 p.m. Panel Discussion **Chair**

- *Assessment Data & Methods*

- Identify additional analyses, sensitivities, corrections

3:30p.m. – 3:45 p.m. Break

3:30 p.m. - 5:00 p.m. Panel Discussion **Chair**

-Continue deliberations

5:00p.m. – 6:00p.m. Panel Work Session **Chair**

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

Wednesday

8:30 a.m. – 12:00 p.m. Assessment Presentation and Discussion (GM) TBD**

12:00 a.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

3:30p.m. – 3:45 p.m. Break

3:30 p.m. - 5:00 p.m. Panel Discussion **Chair**

-Continue deliberations

5:00p.m. – 6:00p.m. Panel Work Session **Chair**

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

Thursday

8:30 a.m. – 12:00 p.m. Panel Discussion **Chair**

- Review additional analyses, sensitivities

12:00 a.m. – 1:30 p.m. Lunch Break

1:30 p.m. – 3:30 p.m. Panel Discussion **Chair**

-Continue deliberations

3:30 p.m. – 3:45 p.m. Break

3:45 p.m. – 5:00 p.m. Panel Discussion **Chair**

-Consensus recommendations and comments

5:00 p.m. - 6:00 p.m. Panel Work Session **Chair**

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

Friday

8:00 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. – 1:00 p.m. Panel Discussion or Work Session **Chair**

- Review Consensus Reports

1:00 p.m. ADJOURN

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

* BLT = South Atlantic blueline tilefish **GM = Gulf of Mexico menhaden