Center for Independent Experts (CIE) Independent Peer Review Report

SEDAR 29 Highly Migratory Species Atlantic and Gulf of Mexico Blacktip Shark Review

Desk Review

Dr. Noel Cadigan

St. John's, NL, Canada

July 2012

Executive Summary

I conclude that the Gulf of Mexico Blacktip shark stock is currently not over-fished, but that stock status is uncertain with respect to over-fishing. The latter conclusion is different from the assessment panel report. The difference is because I do not think the stock assessment model was configured properly and according to standard practice for age-structured assessment models. There is some evidence of model mis-specification which causes me to doubt the assessment panel conclusion about over-fishing.

Background

The purpose of the review was to provide an external peer review of the SEDAR 29 stock assessment for the highly migratory species (HMS) Gulf of Mexico Blacktip shark. SEDAR 29 involved a compilation of data, a standard assessment of the stock, and this CIE assessment review. The review was responsible for ensuring that the best possible assessment is provided through the SEDAR process to provide guidance to the Southeast Fisheries Science Center (SEFSC) to aid in their review and determination of best available science, and to HMS when determining if the assessment is useful for management.

The CIE (Center for Independent Experts) reviewer was tasked with conducting an impartial and independent peer review in accordance with the SoW and ToRs (Appendix 2). Each CIE reviewer was required to complete the independent peer review according to the required format and content as described in Annex 1 of Appendix 2. Each CIE reviewer was required to complete the independent peer review addressing each ToR as described in Annex 2 of Appendix 2. Two weeks before the review the NMFS Project Contact provided background information and reports for the peer review.

The specific goals of the review were to:

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. Conduct an independent peer review in accordance with the ToRs (Annex 2).
3. No later than 19 June 2012, 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.

The Review Panel (RP) was composed of two CIE reviewers. The CIE reviewers were independent, and had working knowledge and recent experience in the application of stock assessment, statistics, fisheries science, and marine biology.

Role of reviewer

I reviewed the SEDAR 29 Stock Assessment Report for HMS Gulf of Mexico Blacktip Shark, in accordance with the SoW and ToRs (see Appendix 2). Prior to this I reviewed the background documents I was provided. These are listed in Appendix 1. This report is structured according to my interpretation of the required format and content described in Annex 1 of Appendix 2.

Summary of findings

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

Age and Growth Datasets and Decisions

The assessment review panel (ARP) decided to use the von Bertalanffy growth models presented in SEDAR29-WP-18. Models were estimated using data pooled from two datasets. The rationale for pooling was that differences in growth rates between datasets were not large (but somewhat significant) and pooling provided a broader range of ages and larger samples sizes to fit growth models. I agree with the decision to pool the data, although I cannot conclude the decision is robust because the consequences of not pooling the data were not explored. Alternative growth models that better capture the trends in the data (see Tor1b) should be considered in future assessments.

The ARP decided to infer the maximum observed age of males and females to be 18.5 years, which I conclude was reasonable. However, I did not find a description of how this maximum age was used in the assessment. It was an input in the *demographic gamer* Excel spreadsheet provided where it played a role in determining the age-specific values for natural mortality (M) that were provided in that spreadsheet, but some text description about the spreadsheet should have been provided. An assessment of how sensitive M's are to reasonable deviations in assumptions about t_{max}'s should be provided.

Reproduction Datasets and Decisions

I found no reasons to disagree with the decisions of the ARP. However, see text for ToR1b. Robustness to assumptions/decisions about maturity and fecundity ogives was not considered by the ARP.

Catch Statistics

The ARP demonstrated in the assessment report (AR) and other WP's that total catch statistics were derived after careful consideration of a variety of sources of removals. The associated decisions appeared to be sound – I cannot recommend better alternatives. The robustness of the decisions was addressed by some sensitivity analyses using different "catch streams".

Some length composition information was available for some catch sources (Figures 9-12 in WP-8), but insufficient information (e.g. spatial coverage of samples, etc.) was presented for me to assess the quality of these data. Little age-composition information was available. The assessment model could not utilize length frequencies, so this information was not directly used, although it was indirectly used when inferring selectivity curves for some of the 'fleets'. Although the lack of reliable size and age composition data is a deficiency, the ARP approach to using the available information seemed reasonable.

Indices of Abundance

Abundance indices were standardized using a delta-lognormal model, which is a reasonable approach. Generalized linear mixed models were used for interactions with year factors and other effects. These were treated as random effects and I think this is a good approach. The decisions about which indices to use in the assessment were based on sound rationale.

Each of the many WP's dealing with indices provided a summary of model selection results and model fit diagnostics for the selected or "preferred" model to standardize indices. The approach seemed sound overall, and the description of results in the WP's was above average. However, I am wary of automatic model selection procedures. They are known to lack robustness in some instances and can lead to over-fitting of data and spurious statistical significance of parameter estimates. The approach used for selecting variables in the model did not just rely on small p-values; the authors also looked at the percent magnitude of the change in fit. There is also a problem that annual changes in stock size may be mis-attributed to a change in a covariate. For example, if temperature has changed gradually over time then it is possible that a change in stock size over time could be mis-attributed to temperature effects, and the year effects may be biased. Correlations in covariates can lead to biased estimates of year-effects.

I also am wary of area effects. If 90% of the stock is in area A, and the stock has been increasing in that area (say 10% per year; $N_{2A} = 1.1xN_{1A}$), and 10% of the stock is in area B and the stock has been decreasing there by 20% per year ($N_{2B} = 0.8xN_{1B}$), then overall the stock has been increasing by about 7% per year: $N_2 = N_{2A} + N_{2B} = 1.1xN_{1A} + 0.8xN_{1B} = 1.1x0.9xN_1 + 0.8x0.1N_1 = 1.07x\ N_1$. However, with area effects in a model the differences in average stock size (over time) in both areas are removed, and the model will tend to just average the different trends in both areas, and one would infer that the stock is decreasing by about 5% per year. Area effects should be removed only if they are related to some process affecting catch rates more than stock size in each area. A useful way to think about this is how stratified random bottom trawl survey data are treated. One would not usually base an index on the year effects from a model with year and strata fixed effects and year*strata random effects. The usual approach is essentially to estimate year*strata fixed effects and then get the strata-size weighted average of the strata effects each year. I have used a model with fixed year and strata effects and random year*strata interactions to deal with strata that are not sampled in some years (similar to the rationale in the indices WP's); however, I still base the index on the strata-size weighted average of approximate BLUP's (Best Linear Unbiased Predictions) of strata effects each year.

The points above illustrate that standardizing catch rates is complicated in practice. I was pleased to see that the authors routinely compared standardized and nominal indices. When the trends in both indices are substantially different then I suggest that careful consideration be given to whether the standardization is effective. I conclude that the approach used was sound overall. In particular I find that the approach of combining indices from different surveys (e.g. states, universities) using similar gear types is useful. Indices based on restricted spatial coverage are much less useful, especially if the spatial distribution of the stock has changed over time. Authors often looked at indices for different age groups and compared indices from previous analyses, and this addresses some robustness issues. Hence, I conclude the approach to indices was reasonably robust. It would be useful in future assessments to perform retrospective standardizations, including the variable selection procedures, as a way to more formally demonstrate the robustness, or lack thereof, of the standardized indices.

Many of the indices WP's gave a description of the spatial coverage of the index effort. I find annual spatial plots of effort and catch rates like in WP-3 and Appendix A of WP-15 to be quite helpful. However, WP-6 did not provide a good description of the spatial distribution of effort, and it seemed relatively restricted (see text for ToR3a).

Technical notes: (1) The authors used SAS PROC GLIMMIX and there can be issues with the BLUP's of random effect interactions because of the estimation approach this package uses (i.e. penalized quasi-likelihood). In the latest version of SAS (version 7 or higher) GLIMMIX offers another and better approach based on maximization of the marginal likelihood via a Laplace approximation. In many instances the two approaches give similar results, but with Binary data, or Binomial data with low numbers of "trials", the BLUP's can be quite different and the Laplace results are preferred. (2)The abstract for WP-22 indicates the Negative Binomial distribution was used, which is a mistake.

Natural mortality

The approach used to derive age-specific values of M was not described well. A spreadsheet was provided that the ARP used to derive M. However, I could not follow the calculations in the time provided for this review. Some text description for this should be provided.

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

Age and Growth Datasets and Decisions

I found some evidence of lack-of-fit for the von Bertalanffy growth model. In Figure 2.6.3 of the AR the size at age 1 appears to be under-estimated for both males and females, but more so for males. This was not discussed in either the AR or WP-18. It was hard to assess how significant this lack-of-fit was because WP-18 did not provide standardized residual plots. Such diagnostic graphs should have been provided.

Reproduction Datasets and Decisions

No evidence on how well the logistic model fit the maturity data was provided in either the AR or SEDAR29-WP-09. This should be provided for future assessments.

Catch Statistics

Uncertainties in catch statistics were acknowledged in some instances (e.g. recreational discard mortality rates) by the RP, but they were not quantified overall, nor are these uncertainties easy to quantify. A useful approach for this is to review different sources of catches on an annual basis with the intent to provide best estimates and a range of plausible values for each source of catches. Often I find that there is ancillary information available on an annual basis from the catch 'data-providers' that can be used to help derive this range. Unfortunately this ancillary information is often not recorded and is basically lost over time so that it is difficult to decide what are useful ranges for historic catches. I have no basis to conclude anything about whether the uncertainties in the Blacktip sharks landings are within normal or expected levels.

The lack of age composition data was acknowledged, as was the incomplete information on length compositions. I am not an expert in shark stock assessment so I cannot conclude if this is within normal or expected levels.

Indices of Abundance

Uncertainties in indices were acknowledged, reported, and within normal or expected levels. However, calculated standard errors do not reflect uncertainties due to variable selection, and this could be partially addressed using retrospective analyses. There are other sources of uncertainty that the standardization models cannot reflect, such as survey/index coverage, reporting biases, etc. These were partially quantified using a ranking procedure. This is a reasonable approach, especially since there are no 'gold-standards' that I am aware of for weighting different indices in stock assessment models. This has always been an important and influential issue to deal with in stock assessment (see research recommendation in ToR6).

Natural mortality

The ARP conducted a low and high M sensitivity analyses which partially addresses uncertainty. However, some consideration should be given to whether M has changed over time, perhaps as a function of changes in the ecosystem (physical or biological).

c. Are data applied properly within the assessment model?

This ToR overlaps with ToR2. I assume the difference is of a technical nature and that ToR1c relates to whether the data are "read" properly by the assessment model. The code for the assessment model (SSASPM; WP-24) was complicated and I was not able to understand all of it. However, the maturity schedule in the input data file did not appear to be sex-specific and was different from the values in Table 2.5.3 of the AR – for males or females. Fecundities in WP-24 were different from the values in Figure 2.6.5 of the AR. Catches in WP-24 were identical to those presented in Table 3.5.1A of the AR. The 6 survey indices in Table 3.5.3 of the AR were identical to those used in WP-24. The scaling factors for observation variance in WP-24 were the rankings in Table 2.5.10 of the AR. It seems that the standard errors derived from the index standardization models are not used in the stock assessment model. Methods to use this valuable information should be considered for future assessments – although I recognize that this methodological issue requires further investigation (see research recommendation in ToR6).

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

The data for growth rates and maturation rates seemed sufficient to support the assessment approach and findings. The assessment model was based on constant growth and maturation rates and the data support these assumptions overall.

I find it difficult to assess the reliability of the catch data or whether they were sufficient to support the assessment approach and findings. However, I find that the ARP expended considerable effort to account for the major sources of removals for Blacktip sharks in the Gulf of Mexico. There may be other possible sources of removals that have occurred in the past that have not been accounted for because there is simply no data to support any estimation. It is a major criticism of fishery stock assessment that we are so reliant on a type of data that is historically unreliable in many situations. We often base assessments on catch histories whose accuracy is difficult to verify. This may be the case with this assessment, which was configured to fit fleet catch series almost exactly when this data source must have some uncertainties, although the magnitude of the uncertainties was not quantified. I am not confident that the sensitivity analyses using different catch streams adequately reflects the uncertainty in the catch data. Further research is required to quantify the accuracy of catch time-series and account for this in the assessment.

The assessment model was age-based but the indices were not, nor were the catch data age dis-aggregated as far as I could tell – although I am unsure about this. Normally one would not contemplate an age-based model for this type of data. However, the indices covered different age groups (aggregated via external selectivity functions) and because there are 6+ indices with different assumed selectivities I think they can support the age based model.

I am unsure if it is possible for this implementation of SSASPM to estimate both fishery and index selectivities. I have used the SURBA assessment model, which is a catch free approach based only on age-based survey indices, and in this situation it is generally understood that fishery selectivity is highly confounded with the age-pattern in survey catchability (Q_a), and fully selected Q (Q_{full}) is unidentified. The latter pattern (i.e. Q_a/Q_{full}) is usually fixed based on best-judgment. Q_{full} is also fixed but this is not important because SURBA only provides trends in stock size and not absolute values. I have also been developing a model similar to SURBA but which also utilizes total fishery landings data (but no age-composition) in addition to age-based survey indices. In this case Q_{full} can be estimated but there still seems to be insufficient information to estimate the age-pattern in Q separately from fishery selectivity. This model seems conceptually similar to the SSASPM model implementation for Blacktip sharks, and I question the ability of SSASPM to estimate both fishery and index selectivities simultaneously. Unless the model is given age composition information for fishery catches then I don't think it is possible to estimate both fishery and index selectivities.

It would be useful to have a understanding of how sensitive model results are to changes in selectivities (maybe low and high); one index at a time. Does the fit statistic change?

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

a. Are methods scientifically sound and robust?

The assessment model is age-based but the important age parameters are determined externally. This is unusual, and I had no experience with the algorithm the AG used to infer selectivities. I think the approach is based on the assumption that recruitment and F_{full} are constant for the cohorts represented in the age-comps. I ran a few checks by generating some age-composition data and implementing the algorithm described in Appendix 2 of the AR. The results seemed OK as long as recruitment and F_{full} are approximately constant. Note that step 7 did not result in normalization for me and I just divided results in step 6 by the maximum over ages.

Results of some of my checks are shown in Figures 1-5. Computer code for Figure 5 is presented in Appendix 3. Note that the age-compositions were generated without error, so the results may be optimistic. Figures 1-3 demonstrate situations where the approach worked fairly well. Figures 4 and 5 demonstrate otherwise. Assumptions are satisfied in Figure 4 but they are not in Figure 5. This figure demonstrates that the approach used to infer selectivity can give substantially biased results when recruitment is substantially different for the cohorts contributing to the age compositions used to infer selectivity.

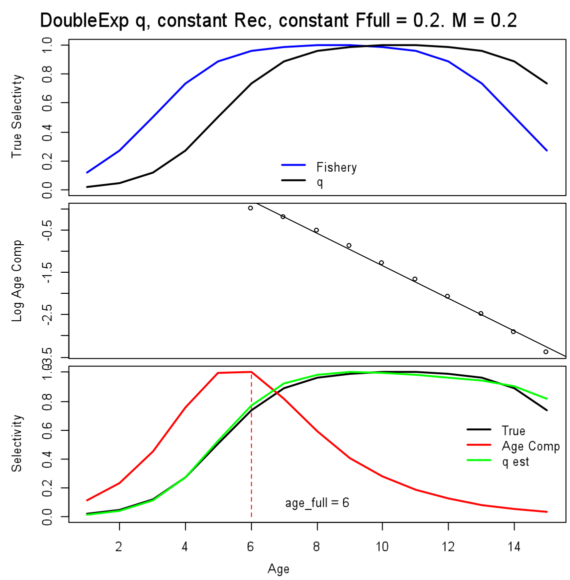

Figure 1. Top panel: Fishery selectivity and index selectivity (i.e. catchability, q) functions used to generate age compositions. Both functions are domed and based on double logistic functions. Middle panel: generated log age composition vs. age. A straight line was fitted via least-squares to ages > a_{full}. Bottom panel: Index age composition scaled to a maximum of one (red), true q age pattern (black), and estimated pattern (green). Cohorts are derived from constant recruitment and experience constant natural mortality (M) and the same fully recruited fishing mortality each year.

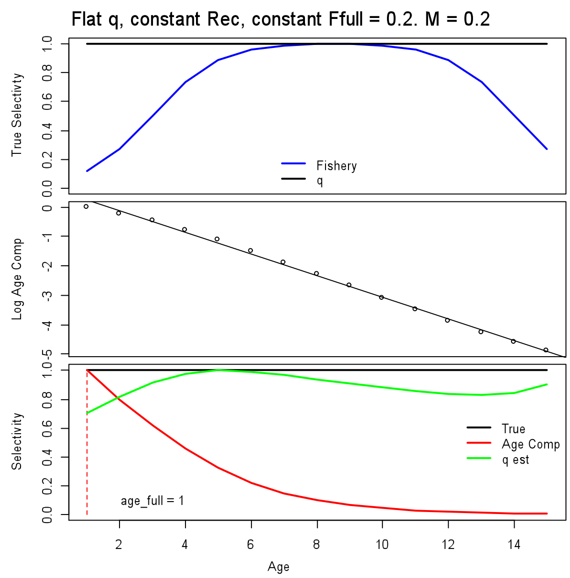

Figure 2. Flat index catchability (q). See Figure 1 for other details.

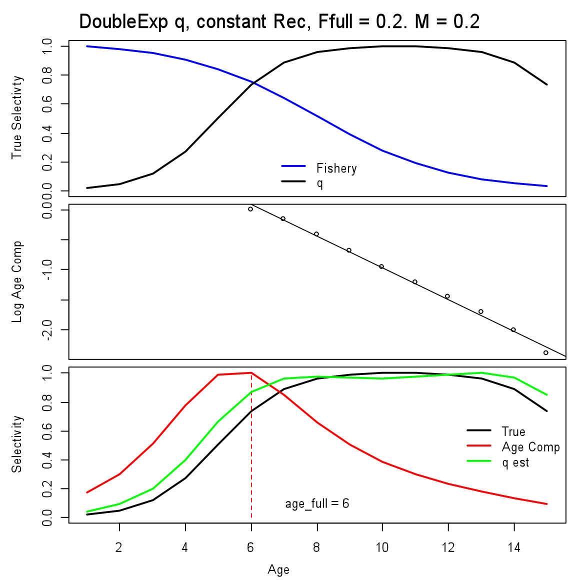

Figure 3. A fishery that targets young sharks. See Figure 1 for other details.

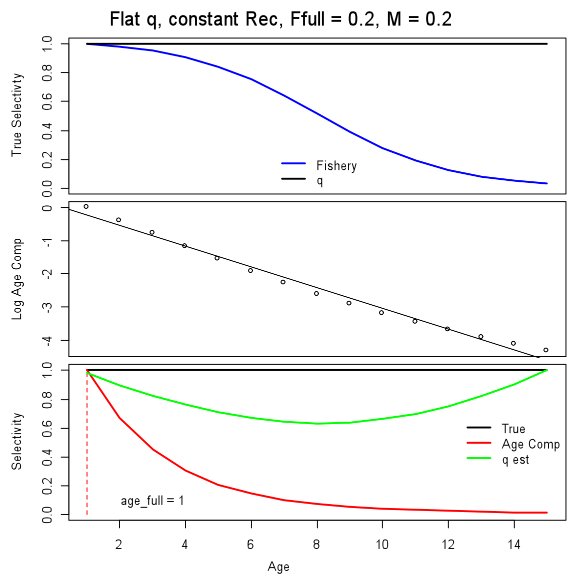

Figure 4. Flat index q and a fishery that targets young sharks. See Figure 1 for other details.

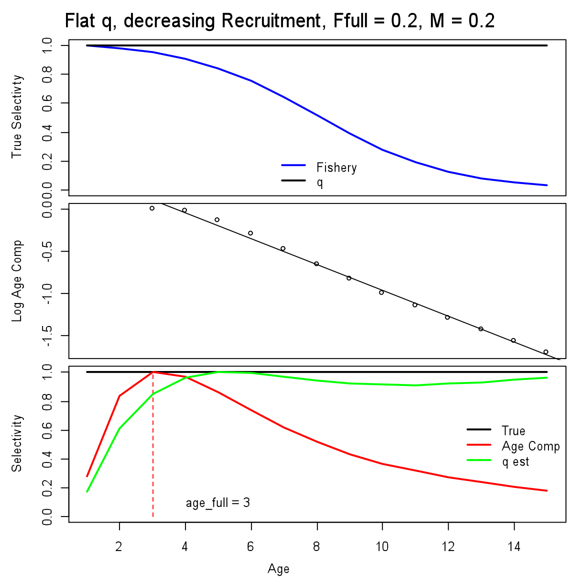

Figure 5. Flat index q, a fishery that targets young sharks, and a linear decrease in recruitment from 50 to 1. See Figure 1 for other details.

Uncertainties related to estimating selectivity functions are not accounted for in the assessment model, as noted in the AR.

Commercial landings were decomposed into four sources, and the fishery selectivity of these catch streams was assumed to be constant over time. The ARP should consider whether the gears, management measures, etc. contributing to these catch streams have changed over time in ways that could change selectivity.

Apart from this, the basic assessment model seemed OK and the ARP reported that it has been used before in other SEDAR assessments and in International Commission on the Conservation of Atlantic Tunas ICCAT assessments. However, I am confused about what age-composition information was used by the model. Equations 6 and 7 in the ARP indicate the model is age-based but the age compositions of catch were not provided, nor did I see them in WP-24. I could see where they were read by the AD Model Builder (ADMB) code, but I could not find the inputs. It was really hard to follow what the code was doing in any event, and I did not follow all of the model description on pg. 58 in the AR. For example, I did not understand the τ_i parameter in Equation 7. Equation (13) suggested that data were weighted by annual CV's but I could not see where these were read in by the ADMB code.

There was insufficient information available for me to reach a conclusion about soundness and robustness of the methods themselves. This cannot really be separated from conclusions about model implementation and the data available, but I have tried to separate my comments under the ToR components.

Little information was presented on the Bayesian Surplus Production (BSP) model,only the output in Figure 3.6.38. It was not possible to evaluate this model application.

Note that in equation (12) in the AR, the whole equation is multiplied by 0.5. This is what the ADMB code does.

b. Are assessment models configured properly and used consistent with standard practices?

I have some concerns with the specific implementation:

1. The population was assumed to be in a virgin state in 1981, even though there was some fishing prior to 1981. I am not sure how much fishing occurred prior to 1981, but Blacktip sharks are a coastal species and I would think they were exploited somewhat even before then. The ARP should have investigated a sensitivity alternative to this assumption. My preference would be to assume the population had a stationary age-distribution in 1981, perhaps with a prior on initial F. The Gulf of Mexico ecosystem was probably not in virgin conditions during the 1970s, which makes me doubt that Blacktip sharks were either.
2. As far as I can tell the authors used the Beverton-Holt model to directly model recruitment (e.g. Figure 3.6.13 in AR), with no process error. The only process error identified in Table 3.5.6 of the AR was fleet-specific deviations in the first-order lognormal autoregressive process for effort. This is not consistent with standard practice in age-based models. For a while the use of stock-recruit models was discouraged in age-based models and the size of each cohort was estimated separately (e.g. ADAPT, XSA). Stock-recruit models are making a come-back in recent stock assessment models, but usually with process error. These process errors (i.e. recruitment deviations) are very important for short-term management of the stock, and they should be included in the SSASPM model configuration. The lack of recruitment process errors is similar to the implicit way recruitment is modeled in a surplus production model, and this is a reason

why the SSASPM assessment model results tend to be similar to results from the BSP model.

c. Are the methods appropriate for the available data?

I suggest that there are sufficient data available to estimate annual recruitment process errors. At least we should be convinced why this is not a good approach. Therefore, the current model configuration is not appropriate for the stock.

ToR 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?

I am concerned about the low F's and constant recruitment implied by the base model run and the various sensitivity runs. My basic conclusion is that catches declined substantially over the last 20 years, but indices have not responded in a consistent manner. The best way to fit the data overall is to split the differences in indices with a basically flat stock trajectory. The only way this can happen is if the catches are not having an effect - i.e. F is low (Figure 3.6.15) and the stock is at a stationary distribution (i.e. Figures 3.6.14 and Figure 3.6.12) with little variation in recruitment (Figure 3.6.13). However, overall the base run model diagnostics indicate the model is mis-specified. I see somewhat consistent patterns in Fig. 3.6.10 that the model does not pick up. The pattern is an increase in indices since the mid 90's to the mid 2000's, then the start of a decline. This pattern holds for 1) PC+MML+MS; 2) BLLOP; 3) NMFS LLSE; 4) MS+MS-LA+AL. The anomalies are: 1) Texas does not decline since the mid 2000's; 2) Everglades National Park (ENP),did not increase. I am not in a good position to evaluate the relative merits of the various indices, but ENP is based on angler interviews at just 2 ports, one of which is at the edge of the stock boundary. My sense from WP-06 was that the spatial coverage of effort for this index was only a small part of the stock range. Although this index had a rank of 4, the model ended up fitting this better than all but one other index (Figure 3.6.11). Although the ENP index covers a long time-span, I suggest the index should be given a much higher rank because of the apparent restricted spatial coverage. My conclusion is that the model has not captured the overall signal in the indices. My concern is that a model with higher fishing mortality and reasonable variations in recruitment process error may fit the data better.

Sensitivity scenario 8 did not include the Texas and ENP indices but did not indicate substantially different results compared to the base model run. I think an issue here is that the recent trends in the NMFS LLSE index cannot be explained by changes in catch alone, and without the ability to vary recruitment much the model chooses low F's as the best option to fit the time-series.

The estimation did not do anything with the modern effort mean parameters (Table 3.5.6). They are identical to priors. I think they are completely confounded with q's (see Equation 8 in AR).

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

Despite my concerns that fishing mortality may be under-estimated, the stock seems to be increasing and the level of indices overall seems good so I do not see any reason to conclude the stock is over-fished. Therefore I agree with the conclusion that the stock is not over-fished.

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

While none of the sensitivity runs indicated overfishing, I am less confident about this conclusion than the conclusion about over-fished status. I am worried that 3 of 6 indices in the base run have been declining more than predicted in the last 5 years. These three indices have broader spatial coverage than the Texas and ENP indices which have not declined recently. Because Fmsy is low (~0.1), I am concerned that a model with reasonable variation in recruitment could indicate an Fcurrent more similar to Fmsy. Of course this is speculation, and I cannot conclude that there is over-fishing.

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

The assessment model assumed a Beverton-Holt stock-recruit relationship. No analyses were presented about the reliability of this relationship. However, if it is reliable then it should be 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?

I am not confident about the specific quantitative estimates of status. Such points estimates usually involve considerable error. I conclude the ARP basic conclusion about over-fished status is reliable. I conclude that the over-fishing status is uncertain.

ToR 4: Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:

a. Are the methods consistent with accepted practices and available data?
b. Are the methods appropriate for the assessment model and outputs?
c. Are the results informative and robust, and useful to support inferences of probable future conditions?
d. Are key uncertainties acknowledged, discussed, and reflected in the projection results?

Since the ARP found that the stock was not overfished and overfishing was not occurring, no projections were presented. Hence, I cannot conclude anything for this ToR.

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

a. ***Comment on the degree to which methods used to evaluate uncertainty reflect and capture 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.***

Process error in the stock-recruit relationship was not included in the assessment model. This is usually included in age-based models and tends to be a substantial component of assessment error. This could seriously affect conclusions about over-fishing, but I do not think it would affect conclusions about over-fished status.

Uncertainty in fishery catches were not quantified, and they are usually difficult to quantify. Their uncertainty in the assessment was quantified using a low and high catch sensitivity analysis. I do not feel that this provides an adequate characterization of the assessment conclusions uncertainty due to uncertainty in catches. A better approach would be to bootstrap the model over a range of catches; however, this would not reflect uncertainty due to bias (i.e. under-reported) landings. There are very few methods available to deal with this, and the ones I am familiar with (e.g. Bousquet et al., 2010) are works in progress.

The index CV's from standardization models (i.e. measurement errors?) may not have been used. This is not uncommon. It seems like a good idea to use the CV's although it is not clear to me how this should be done in practice.

Bousquet, N., Cadigan, N., Duchesne, T., and Rivest, L.-P. 2010. Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Can. J. Fish. Aquat. Sci. 67: 1247-1261.

ToR 6: Consider the research recommendations provided and make any additional recommendations or prioritizations warranted.

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

These are longer-term initiatives that should improve the assessment.

1. Continued conventional tagging and advanced tagging technologies, with a view to producing additional stock assessment information. So far tagging information seems to have been just used to indicate movement patterns; however, tagging data can give more information than this (e.g. telemetry tagging+acoustic arrays or pop-up satellite tagging can be used to estimate M). This may involve more intensive tagging at specific sites.
2. Consider estimating variance components within the assessment model, to produce a more statistically rigorous model. This should involve using the ADMB random effects module to deal with random effects. However,
3. it would be useful for SEDAR or some other group to have a methodological workshop dealing with how to weight different sources of information in a stock assessment model. For example, the total assessment-model CV of a survey index is probably not the product of the within survey CV times the "extra" CV. Also, it is not clear if "self-

weighting" is always a good idea – I know of situations where it can fail badly (i.e. with 2 indices, one long and one short). The ranking approach used by the ARP may be a good idea, but there are issues. For example, what effect does adding 5 to the ranks have? I think there should be no effect, but is this the case?

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

This assessment was too complicated for a desk review. The assessment model is not standard (like BSP, ADAPT, etc), and the standardization of indices was fairly complex. The AR was vague in some respects, which is understandable given the nature of the stock and available data. Although the conclusions of the ARP may not be controversial, a good review of this assessment needs a face-to-face meeting or webinar. ToR's often referred to robustness and reliability, and it is difficult to make conclusions for these types of ToR's when the reviewer does not have the ability to run or evaluate model sensitivities.

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

These are shorter-term initiatives that should improve the assessment. They are prioritized.

1. Include recruitment process error.
2. Continue collecting length and age data.
3. Continue initiatives to combine partial indices to provide for more spatially complete stock size indices.
4. Investigate if fishery and index selectivities are confounded with the available data. To the extent possible, estimate selectivities within the assessment model.
5. Consider alternative growth models and implications on M. The R *scam* function can be easily used to fit a nonparametric growth curve, although you need to be careful with prediction/extrapolations.

ToR 8: Prepare a Peer Review Report summarizing the Reviewer's evaluation of the stock assessment and addressing each Term of Reference.

This is the present document.

Conclusions and recommendations

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

I agree with the decision to pool the growth data. The ARP decided to infer the maximum observed age of males and females to be 18.5 years, which I conclude was reasonable.

The ARP demonstrated in the assessment report (AR) and other WP's that total catch statistics were derived after careful consideration of a variety of sources of removals. The associated decisions appeared to be sound – I cannot recommend better alternatives. The ARP approach to using the available size and age composition data seemed reasonable.

Abundance indices were standardized using a delta-lognormal mixed-effects model, which was a reasonable approach. The variable selection procedure used in the standardizations seemed reasonable overall, although lack of robustness is a concern. The decisions about which indices to use in the assessment were based on sound rationale.

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

There is some evidence of lack-of-fit with the von Bertalanffy growth model. It was hard to assess how significant this lack-of-fit was because standardized residual plots were not provided. Such diagnostic graphs should have been provided. No evidence was provided on how well the logistic model fit the maturity data.

Uncertainties in catch statistics were acknowledged in some instances (i.e. recreational discard mortality rates) by the ARP, but they were not quantified overall, nor are these uncertainties easy to quantify. I have no basis to conclude anything about whether the uncertainties in the Blacktip sharks landings are within normal or expected levels. The lack of age composition data for landings was acknowledged, as was the incomplete information of length compositions. I cannot conclude if this is within normal or expected levels.

Uncertainties in indices were acknowledged, reported, and within normal or expected levels. However, calculated standard errors do not reflect uncertainties due to variable selection, and this could be partially addressed using retrospective analyses. Other sources of uncertainty that the standardization models cannot reflect were partially quantified using a ranking procedure.

c. Are data applied properly within the assessment model?

The maturity schedule in the SSASPM input data file did not appear to be sex-specific and was different from the values in Table 2.5.3 of the AR – for males or females. Fecundities in WP-24 were different from the values in Figure 2.6.5 of the AR. Landings and survey indices were applied correctly. I was unsure how index standard errors or landings age compositions were applied.

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

The data for growth rates and maturation rates seemed sufficient to support the assessment approach and findings. I find it difficult to assess the reliability of the catch data or whether it is sufficient to support the assessment approach and findings. However, I find that the ARP expended considerable effort to account for the major sources of removals for Blacktip sharks in the Gulf of Mexico. Although indices are not age-based, they covered different age groups (aggregated via external selectivity functions) and because there are 6+ indices with different assumed selectivities I think they can support the age based model. I am unsure if it is possible for this implementation of SSASPM to estimate both fishery and index selectivities. This is because I am unsure if independent information on the age-compositions of landings was used.

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

a. Are methods scientifically sound and robust?

The assessment model is age-based but the important age parameters are determined externally. The algorithm used to estimate selectivities seems OK as long as recruitment and fully-selected fishing mortality are constant for the cohorts represented in the age composition data. Otherwise the approach can produce substantially biased and misleading estimates. Uncertainties related to estimating selectivity functions are not accounted for in the assessment model.

There was insufficient information available for me to reach a conclusion about soundness and robustness of the methods themselves. This cannot really be separated from conclusions about model implementation and the data available (see ToR2b,c).

Little information was presented on the Bayesian Surplus Production (BSP) model. It was not possible to evaluate this model application.

b. Are assessment models configured properly and used consistent with standard practices?

c. Are the methods appropriate for the available data?

No. Standard practice in age-based models is to include process error in the stock-recruit model used in the stock-assessment model. This was not done in this case. There is sufficient data available to estimate annual recruitment process errors. Therefore, the current model configuration is not appropriate for the stock.

Tor 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?

I suspect the results are not as reliable as the report indicates. I am concerned about the low F's and constant recruitment implied by the base model run and the various sensitivity runs. The model has not captured the overall signal in the stock size indices. A model with higher fishing mortality and reasonable variations in recruitment process error may fit the data better.

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

The stock is not over-fished. I reach this conclusion because the level of indices overall seems good.

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

I cannot conclude anything about overfishing.

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

No analyses were presented about the reliability of this stock recruitment relationship.

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?

I am not confident about the specific quantitative estimates of status. Such point estimates usually involve considerable error. I conclude that the ARP basic conclusion about over-fished status is reliable. I conclude that the over-fishing status is uncertain.

Tor 4: Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:

a. Are the methods consistent with accepted practices and available data?
b. Are the methods appropriate for the assessment model and outputs?
c. Are the results informative and robust, and useful to support inferences of probable future conditions?
d. Are key uncertainties acknowledged, discussed, and reflected in the projection results?

Since the ARP found that the stock was not overfished and overfishing was not occurring, no projections are presented. Hence, I cannot conclude anything for this ToR.

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

a. Comment on the degree to which methods used to evaluate uncertainty reflect and capture 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.

Process error in the stock-recruit relationship was not included in the assessment model. This is usually included in age-based models and tends to be a substantial component of assessment error. This could seriously affect conclusions about over-fishing, but I do not think it would affect conclusions about over-fished status.

Uncertainty in fishery catches was quantified using a low and high catch sensitivity analysis. I do not feel that this provides an adequate characterization of the assessment conclusions uncertainty due to uncertainty in catches

Tor 6: Consider the research recommendations provided and make any additional recommendations or prioritizations warranted.

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

These are longer-term initiatives that should improve the assessment.

1. Continued conventional tagging and advanced tagging technologies, with a view to producing additional stock assessment information.
2. Consider estimating variance components within the assessment model, to produce a more statistically rigorous model.
3. It would be useful for SEDAR or some other group to have a methodological workshop dealing with how to weight different sources of information in a stock assessment model.

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

This assessment was too complicated for a desk review.

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

These are shorter-term initiatives that should improve the assessment. They are prioritized.

1. Include recruitment process error in the assessment model.
2. Continue collecting length and age data.
3. Continue initiatives to combine partial indices to provide for more spatially complete stock size indices.
4. Investigate if fishery and index selectivities are confounded with the available data. To the extent possible, estimate selectivities within the assessment model.
5. Consider alternative growth models and implications on M.

Tor 8: Prepare a Peer Review Report summarizing the Reviewer's evaluation of the stock assessment and addressing each Term of Reference.

This is the present document.

Appendix 1: Bibliography of materials for review

Workshop Document List

Documents Prepared for the Assessment Process

SEDAR29-WP-01	Relative abundance of blacktip shark, *Carcharhinus limbatus*, from the eastern Gulf of Mexico	John Carlson, Dana Bethea, John Tyminski, and Robert Hueter
SEDAR29-WP-02	Standardized catch rates of blacktip sharks (*Carcharhinus limbatus*) in the U.S. Gulf of Mexico from the Shark Bottom Longline Observer Program, 1994-2010	John K. Carlson, Loraine Hale, Alexia Morgan, and George Burgess
SEDAR29-WP-03	Indices of Blacktip Shark Based on NMFS Bottom Longline Surveys (1995 ‐ 2011)	Walter Ingram
SEDAR29-WP-04	Commercial Bottom Longline Vessel Standardized Catch Rates of Blacktip Sharks in the Gulf of Mexico and US South Atlantic, 1996-2010	Kevin McCarthy
SEDAR29-WP-05	Standardized catch rates for Gulf of Mexico Blacktip Sharks from the U.S. Pelagic longline logbook using generalized linear mixed models	Enric Cortés and Ivy Baremore
SEDAR29-WP-06	Standardized catch rates of blacktip sharks from the Everglades National Park Creel Survey	John K. Carlson and Jason Osborne
SEDAR29-WP-07	Tag and recapture data for blacktip shark, *Carcharhinus limbatus*, in the Gulf of Mexico: 1999-2010	Dana M. Bethea, John K. Carlson, and Mark A. Grace
SEDAR29-WP-08	Updated catches of Gulf of Mexico blacktip sharks	Enric Cortés and Ivy Baremore
SEDAR29-WP-09	Reproduction of the blacktip shark *Carcharhinus limbatus* in the Gulf of Mexico	Ivy E. Baremore and Michelle S. Passerotti
SEDAR29-WP-10	A standardized CPUE index of abundance for Gulf of Mexico blacktip sharks from the Marine Recreational Statistics Survey (MRFSS)	Elizabeth A. Babcock
SEDAR29-WP-11	Catch rates and size distribution of blacktip shark *Carcharhinus limbatus* in	J. Marcus Drymon and Sean P. Powers

	the northern Gulf of Mexico, 2006-2010		
SEDAR29-WP-12	Relative abundance of blacktip shark based on a fishery-independent gillnet survey off Texas	Walter Bubley, John K. Carlson,	
SEDAR29-WP-13	Standardized catch rates of blacktip sharks (*Carcharhinus limbatus*) collected during a gillnet survey in Mississippi coastal waters, 1998-2011	Eric Hoffmayer, Glenn Parsons, Jill Hendon, and Adam Pollack	
SEDAR29-WP-14	Standardized catch rates of blacktip sharks (*Carcharhinus limbatus*) collected during a bottom longline survey in Mississippi coastal waters, 2004-2011	Eric Hoffmayer, Jill Hendon, and Adam Pollack	
SEDAR29-WP-15	Standardized catch rates of blacktip sharks (*Carcharhinus limbatus*) collected during a SEAMAP bottom longline survey in Mississippi/Louisiana coastal waters from 2008 to	Jill M. Hendon, Eric R. Hoffmayer and Adam G. Pollack	2011.
SEDAR29-WP-16	Mark/Recapture Data for the Blacktip Shark, *Carcharhinus limbatus*, in the Gulf of Mexico from the NEFSC Cooperative Shark Tagging Program	William Swinsburg, Nancy E. Kohler, Patricia A. Turner, and Camilla T. McCandless	
SEDAR29-WP-17	A Preliminary Review of Post-release Live-discard Mortality Estimates for Sharks	Dean Courtney	
SEDAR29-WP-18	Updates to age and growth parameters for blacktip shark, *Carcharhinus limbatus*, in the Gulf of Mexico	Michelle S. Passerotti and Ivy E. Baremore	
SEDAR29-WP-19	Commercial Bottom Longline Vessel Standardized Catch Rates of Blacktip Sharks in the United States Gulf of Mexico, 1996-2010, with targeting determined using logistic regression	Kevin McCarthy	
SEDAR29-WP-20	Dead discards of blacktip sharks in the Gulf of Mexico shark bottom longline fishery	Kevin McCarthy and John Carlson	
SEDAR29-WP-21	A combined fishery independent gillnet series for juvenile blacktip sharks in the eastern Gulf of Mexico	John Carlson, Robert Hueter, Eric Hoffmayer, and Walter Ingram	
SEDAR29-WP-22	Standardized catch rates of blacktip sharks (Carcharhinus limbatus) collected during bottom longline surveys in Mississippi, Louisiana, and Alabama coastal waters from 2004 to 2010	Eric Hoffmayer, Jill Hendon, Marcus Drymon, Sean Powers, Adam Pollack, and John Carlson	

Final Stock Assessment Reports

SEDAR21-SAR	Gulf of Mexico Blacktip Shark

Reference Documents

SEDAR29-RD01	SEDAR 11 (LCS) Final Stock Assessment Report	SEDAR 11 Panels
SEDAR29-RD02	Distributions of Sharks across a Continental Shelf in the Northern Gulf of Mexico	J. Marcus Drymon, Sean P. Powers, John Dindo, Brian Dzwonkowski, and Terry Henwood
SEDAR29-RD03	Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea	D.B. Keeney, M.R. Heupel, R.E. Hueter, and E.J. Heist
SEDAR29-RD04	Estimation of catches of sandbar (*Carcharhinus plumbeus*) and blacktip (*C. limbatus*) sharks in the Mexican fisheries of Gulf of Mexico (SEDAR 11-DW-06)	R. Bonfil and E. Babcock
SEDAR29-RD05	Abundance Indices Workshop: Developing protocols for submission of abundance indices to the SEDAR process	SEDAR Procedural Workshop I
SEDAR29-RD06	Do differences in life history exist for blacktip sharks, *Carcharhinus limbatus*, from the United States South Atlantic Bight and Eastern Gulf of Mexico?	John K. Carlson, James R. Sulikowski, Ivy E. Baremore
SEDAR29-RD07	Hierarchical analysis of blacknose, sandbar, and dusky shark CPUE indices (SEDAR21-AW-01)	Paul Conn

Appendix 2: A copy of the CIE Statement of Work

Attachment A

Statement of Work for Dr. Noel Cadigan

External Independent Peer Review by the Center for Independent Experts

SEDAR 29 Highly Migratory Species Atlantic and Gulf of Mexico Blacktip Shark Review

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 Technical Representative (COTR), 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.

Project Description: SEDAR 29 will be a compilation of data, a standard assessment of the stock, and CIE assessment review conducted for HMS Gulf of Mexico Blacktip shark. The desk review provides an independent peer review of SEDAR stock assessments. The review is responsible for ensuring that the best possible assessment is provided through the SEDAR process and will provide guidance to the SEFSC to aid in their review and determination of best available science, and to HMS when determining if the assessment is useful for management. The stocks assessed through SEDAR 29 are within the jurisdiction of the Highly Migratory Species Division of NOAA Fisheries and the states of Texas, Louisiana, Mississippi, Alabama, and Florida. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Two CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of reviewing the technical details of the methods used for the assessment. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

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

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, and other pertinent information. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements.

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

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) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 19 June 2012, 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.

21 May 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact.
4 June 2012	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers. Background documents may be sent to the CIE reviewers one week earlier.
4-15 June 2012	Each reviewer conducts an independent peer review as a desk review.
19 June 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator.
3 July 2012	CIE submits the CIE independent peer review reports to the COTR.
10 July 2012	The COTR 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 COTR within 10 working days after receipt of all required information of the decision on changes. The COTR 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 COTR 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 COTR (William Michaels, via William.Michaels@noaa.gov).

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 COTR within 10 working days after receipt of all required information of the decision on changes. The COTR 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 COTR 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 COTR (William Michaels, via William.Michaels@noaa.gov).

Support Personnel:
William Michaels, Program Manager, COTR
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:

Julie A Neer, SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
Email: Julie.neer@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.

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 2: Terms of Reference for the Peer Review

SEDAR 29 HMS Gulf of Mexico Blacktip Shark

1. Evaluate the data used in the assessment, addressing the following:
   a) Are data decisions made by the assessment panel sound and robust?
   b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c) Are data applied properly within the assessment model?
   d) Are input data series reliable and sufficient to support the assessment approach and findings?
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?
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?
4. Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:
   a) Are the methods consistent with accepted practices and available data?
   b) Are the methods appropriate for the assessment model and outputs?
   c) Are the results informative and robust, and useful to support inferences of probable future conditions?
   d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
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 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 Report summarizing the Reviewer's evaluation of the stock assessment and addressing each Term of Reference.

Appendix 3: R code for check of external selectivity estimation

```
age=1:15
n=length(age)

logit_sel1 = 0.5*(8 - age)
logit_sel2 = -3 + age
sel1 = exp(logit_sel1)/(1+exp(logit_sel1))
sel2 = exp(logit_sel2)/(1+exp(logit_sel2))
sel=sel1*sel2;
sel=sel1;
sel=sel/max(sel)

F_full = 0.2

M = 0.2
Z = M + F_full*sel
cZ = c(0,cumsum(Z)[1:(n-1)])

logit_sel1 = 16 - age
logit_sel2 = -5 + age
q1 = exp(logit_sel1)/(1+exp(logit_sel1))
q2 = exp(logit_sel2)/(1+exp(logit_sel2))
q=q1*q2;
q=q/max(q)
q=q/q

par(mfrow=c(3,1),oma=c(4,0,2,0.5),mar=c(0,5,0.5,0))

plot(age,sel,xlab='age',ylab='True Selectivty',lwd=2,type='l',col='blue',ylim=c(0,1),xaxt='n')
lines(age,q,lwd=2)

legend("bottom",c("Fishery","q"),col=c('blue','black'),
lty=1,lwd=2,bty='n')

Nage= exp(-cZ)
Nage = seq(1,50,length=n)*exp(-cZ)

Iage = q*Nage
age.comp = Iage/max(Iage)

p1 = which.max(age.comp)
full.age = age[p1]
```

```
ind = p1:n
temp = lsfit(age[ind],log(age.comp[ind]))

plot(age[ind],log(age.comp[ind]),xlab='age',ylab='Log Age Comp',xaxt='n',xlim=range(age))
abline(temp)

Eage.comp = exp(temp$coef[1] + temp$coef[2]*age)

ratio = age.comp/Eage.comp
est.q1 = ratio/ratio[p1]
est.q2 = ratio/max(ratio)

plot(age,q,type='l',lwd=2,xlab='Age',ylab='Selectivity',ylim=c(0,1))
lines(age,age.comp,lwd=2,col='red')
segments(full.age,0,full.age,1,lty=2,col='red')
text(full.age+2,0.1,paste('age_full = ',full.age,sep=''))
mtext(side=1,line=2.5,outer=T,'Age',cex=0.7)

#lines(age,est.q1,lwd=2,col='blue')
lines(age,est.q2,lwd=2,col='green')

legend("right",c("True","Age Comp","q est"),col=c('black','red','green'),
lty=1,lwd=2,bty='n')

mtext(side=3,line=0.2,outer=T,'Flat q, decreasing Recruitment, Ffull = 0.2, M = 0.2')
```