48th NORTHEAST REGIONAL STOCK ASSESSMENT REVIEW COMMITTEE (SARC-48)

Reviewer Report to the Center for Independent Experts on the Tilefish, Ocean Quahog and Weakfish Benchmark Stock Assessment Review (SARC 48) held June 1-4, 2009 in Woods Hole, Massachusetts.

Prepared for:

Center for Independent Experts

By:

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

Executive Summary

This document contains an independent reviewer report of review activities and findings for the 48th Northeast Regional Stock Assessment Review Committee Meeting (SARC 48), held June 1-4, 2009 at the Northeast Fisheries Science Center in Woods Hole Massachusetts. Benchmark assessments for three stocks were reviewed at the meeting: golden tilefish, ocean quahog and weakfish. I found that all three assessments provided a sound basis for management advice.

The tilefish assessment working group completed the TOR's for this assessment. Neither of the two models presented (a surplus production model and a statistical, age-and-length-structured model fit to the CPUE and length-frequency data) provided both ideal fits to the data as well as plausible parameter estimates. However, the working group thoroughly investigated both models, acknowledged issues with the models and provided practical management advice given the uncertainties identified in the analyses. There was a consensus within the review committee that the stock is not overfished and that overfishing is not occurring. The review committee also agreed with the working group that there is little evidence that the stock has rebuilt to B_{TARGET}.

The ocean quahog assessment working group also completed the TOR's for this assessment. The commercial landings and fishing effort are well characterized for this stock. Analyses were very thorough, at times using multiple independent methods to validate conclusions. As a whole, the stock is slowly being fished down to towards its B_{MSY} reference point (1/2 of the virgin biomass). There was a consensus among the review committee that the stock is not overfished and that overfishing is not occurring. The unique biology of quahog (slow growth, low levels of recruitment and very long-lived) creates time lags that are outside the planning horizons for most managed activities and presents unique challenges for the assessment of this stock. These issues were handled well in the assessment.

The weakfish assessment working group did a considerable amount of work in meeting the TOR's for this assessment. Multiple analyses were presented to provide estimates of abundance, total mortality and fishing mortality, including an ADAPT VPA, an analysis of survey data as abundance indices, and a Steele-Henderson production model including predation effects. While there are technical issues with some of the modeling, taken on the whole the analyses indicate that abundance has declined markedly, total mortality is high, non-fishing mortality has recently increased and that the stock is currently in a depleted state.

Overall, given the high quality of the assessment documents and presentations, as well as the willingness of the presenters to openly discuss their work, it was simply a pleasure to participate in this review. The working groups are to be commended for their efforts in these assessments.

Table of Contents

1.0 Background

This document contains my independent reviewer report of review activities and findings for the 48th Northeast Regional Stock Assessment Review Committee Meeting (SARC 48), held June 1-4, 2009 at the Northeast Fisheries Science Center in Woods Hole Massachusetts. Benchmark assessments for three stocks were reviewed at the meeting: golden tilefish, ocean quahog and weakfish. Prior to the meeting, the review committee (Appendix 1), were provided with a Statement of Work (Appendix 2), including the Terms of Reference (TOR) for each assessment as well as for the review committee (RC). Assessment documents and background material (Appendix 3) were provided via a website during the two weeks before the meeting. During the meeting there was a general consensus among the RC on nearly all of the main discussion points and findings of the committee as outlined in the Summary Report. This document contains a summary of those findings as well as my own views about these assessments.

2.0 Individual Reviewer Activities

Prior to the meeting I reviewed all the assessment and background documents provided for the review. The meeting chair assigned RC members to act as the lead reviewer for each assessment with the primary responsibility of drafting the review text for that stock for the Summary Report, although we were all expected to conduct an independent review of each assessment. I was assigned golden tilefish.

I participated in the SARC meeting in Woods Hole, MA, from June 1 to June 4, 2009. The main sessions were open to the public, who contributed constructively to the review, particularly on points of clarification and in discussions about fishery activities. Assessment leaders from the stock assessment workshop presented the assessment results. The structure was fairly informal with a lot of discussion during the presentations, an approach that worked well in this case. The tilefish assessment was led by Paul Nitschke, the quahog assessment by Larry Jacobson, and the weakfish assessment by Jeff Brust. Supporting presentations were given by members of the assessment working groups (WGs). The RC met with the WGs two times after the main presentations to answer further questions and to present additional model runs and analyses. A first draft of text for the Summary Report was prepared during the final day of the meeting.

Panel members were required to prepare their individual, independent reports after the meeting. As outlined in Appendix 3, these reports should state in the reviewers own words whether each ToR of the Stock Assessment Workshop was completed successfully, should state whether they accepted or rejected the work that they reviewed, and should include an explanation of their decisions (strengths, weaknesses of the analyses, etc.) and recommendations for each TOR. A key determinant of whether a TOR had been met was the extent to which it provided a scientifically credible basis for developing fishery management advice. The following three sections in this document contain my review reports for each of the three assessments.

3.0 Golden Tilefish Assessment Review

The tilefish assessment working group met the TOR's for this assessment. Two models, used to evaluate fishing mortality and total stock biomass, were presented: a surplus production model (ASPIC) and a statistical, age-and-length-structured model fit to the CPUE and length-frequency data (SCALE). Although neither model provided both ideal fits to these data and plausible parameter estimates, there was a consensus that the ASPIC model provided a reasonable approximation of the long term productivity of the stock and was sufficient for providing management advice at this time. There was also a consensus that the stock is not overfished and that overfishing is not occurring. The RC also agreed with the working group that while there is evidence that the stock is above $B_{THRESHOLD}$, there is little evidence that the stock has rebuilt to B_{TARGET} given the poor ASPIC model fit to the most recent CPUE data. As demonstrated by the working group, tilefish are vulnerable to being classified as "overfished" if the CPUE index continues to decline, a plausible scenario given the apparent dependence on large year classes that are produced periodically. These conclusions are expanded upon in the following text under each TOR.

Meeting the Terms of Reference

> **TOR 1. Characterize the commercial catch including landings, effort and discards. Characterize recreational landings. Evaluate utility of study fleet results as improved measures of CPUE.**

The assessment working group addressed this TOR well. Commercial landings time series were presented spanning the period 1915 to 2008. This series shows that landings have been highly variable. With the exception of 1916, commercial landings were highest during the late 1970's and early 1980's and that recent landings are roughly 1/3 to 1/4 those values. Recent observer data appears to indicate that the discarding is low. Recreational landings also appear to be low.

Effort data were available from three sources (the Turner series, the weighout series and VTR series) spanning the period from 1973 to 2008. Rather than combining the data into a single CPUE series, the WG derived CPUE series separately for these data, an approach that was preferred given the limited overlap in vessels targeting tilefish in the weighout and VTR series. The WG also did a thorough job of evaluating the effects of the assumed error distribution on the CPUE indices.

The study fleet evaluation was a good preliminary evaluation of factors that effect CPUE. The study investigated the effects of several factors (e.g. number and size of hooks, length of main line, soak time, time of day, depth fished and area fished) on CPUE. Given the short time series for these data, they could not be used in the current assessment, but the study shows that these kinds of data could be used in future assessments if the study was extended. With respect to the current assessment, the results show that the current method of estimating effort (days absent - number of trips) does correlate well with effort measured in this study, thereby supporting the use of the current CPUE series in this assessment. However, other factors such as number of hooks and potentially location were also shown to be possible explanatory variables that might improve the CPUE index in the future. As such, multivariate approaches, which were not

undertaken in study fleet evaluation, will likely be necessary to choose a set of variables to include in the CPUE standardization, once a long enough time series is obtained. Set-by-set geo-referencing is not available for analysis in the present CPUE data. Two potential issues with the SCALE model analysis (discussed below) are the shape of the selectivity curve and the whether the changes in the length-frequency distributions are representative of changes in the biological characteristics of the stock or changes in fishing practices. These questions could be addressed using the study fleet data if this program was continued in the future.

The WG also highlighted some inconsistencies in the dealer, VTR and IVR reported landings, as well as inconsistencies in the market category designation among fishing ports, although the data were considered to be sufficient for the current assessment by both the WG and RC. Addressing these inconsistencies could lead to improved future assessments.

TOR 2. Estimate fishing mortality and total stock biomass for the current year, and for previous years if possible, and characterize the uncertainty of those estimates. Incorporate results of new age and growth studies.

This TOR was completed successfully by the assessment working group. The WG presented the results of two models used to evaluate fishing mortality and total stock biomass for the current year, a surplus production model (ASPIC) and a statistical, age-and-length-structured model fit to the CPUE and length-frequency data (SCALE). Results of new age and growth studies were incorporated into the SCALE model (ASPIC does not use age and growth data). These two models gave conflicting results, in part due to a rapid increase in the CPUE index followed by a rapid decrease combined with a shift in the length-frequency data to smaller sizes in recent years. The was consensus that ASPIC did a better job of characterizing the long term productivity of the stock than did the SCALE model, and the RC endorsed the WG position that the ASPIC model be used as the basis for this assessment. Both the WG and the RC acknowledged uncertainties in the estimation of current year biomass or recent stock trends. The CPUE index shows a recent decline consistent with the passage of a strong cohort through the stock (consistent with the recent length-frequency data). Given the nature of surplus production models, ASPIC is not able to track this kind of dynamic and shows a recent increase in abundance in spite of the declining CPUE trend. As a result, biomass may be overestimated during the last few years leading to uncertainty about whether the stock has rebuilt. Uncertainty in the ASPIC model was evaluated by starting the model at different proportions of K and by bootstrapping to obtain CDF's for model parameter estimates for the preferred model run.

The rapid rise and fall in the CPUE index, combined with the length frequency data, appear to be due to one or two large cohorts presently moving through the fishable stock. An age structured model should be better suited for modeling this kind of dynamic, but the SCALE model did not provide plausible estimates of fishing mortality in the late 1990's (possibly due to a change in fishing practices, productivity or fishery selectivity being interpreted in the model as a change in fishing mortality).

Personally, I was pleased with the SCALE model development, in part due to own experiences (some successful and some not) of fitting age structured models to length frequency data. Given the poor fit of the ASPIC model to the most recent CPUE data, development of an assessment

model better able to handle year class effects is encouraged, and in my opinion, the SCALE model is a good step in this direction and its development should be continued to resolve the issues leading to the implausibly high fishing mortality estimates obtained for some years. Model development is an exploratory process, some things work and some don't. I'm therefore offering the following as considerations (not recommendations) if the model is further developed:

1. A key issue appears to be whether or not the length-frequency data is representative of the stock or is heavily influenced by fishing practices. This is a question perhaps better addressed by the study fleet program. At present, the model appears to be interpreting the lack of larger fish as indicative of their absence and leading to high estimates of fishing mortality. If the length-frequency data are influenced by fishing practices (fishing is occurring in a manner such that larger fish are not being caught), then the use of a dome shaped (e.g. double half Gaussian) would be more appropriate than the logistic selectivity curve currently in the model. My guess is that a dome-shaped curve (particularly if separate curves are used for two time periods) would produce more plausible estimates, however caution is warranted if it is used without evidence that large fish are present but not being captured representatively in the fishery (the evidence could come from the study fleet program).
2. Residual plots would help with diagnosing the model (and also to evaluate its credibility). For example, plotting the residuals for the length-frequency data against length and year separately could help identify whether there are process changes through time and whether the selectivity curve is appropriate.
3. Age-based selectivity could be explored.
4. The recruitment deviates are estimated relative to recruitment in year 1. Depending on how much weight is placed on this component of the objective function, this could force the starting recruitment to be near the mean for the full series. As an alterative, the mean of the recruitment series could be calculated and the residual sum of squares around this mean could be added to the objective function value.
5. The relationship between the standard deviation of the length-at-age could be explored using the new age and growth data, leading to an improved understanding of this relationship (the increased standard deviation with age scenario appears ad-hoc at present).
6. Some of the length-frequency data appear noisy (relatively large differences in frequency for length categories separated only by one centimeter). This could create issues with fitting to these data. Binning the data into larger size groups may help alleviate this potential issue.
7. As currently weighted, the model is fitting to the landings data almost perfectly. An alternative would be to assume the landings are known without error, which would allow the fishing mortalities to become derived parameters, rather than estimated ones. This would reduce the number of estimated parameters by one for each year in the model.
8. Presently, the likelihoods are not fully specified. This means that the weights are not directly comparable as the objective function components are contributing to the overall objective function value at different magnitudes. Specifying the likelihoods as fully as possible would aid in the interpretation of the weighting and would also allow the interpretation of the MCMC output from AD-Model Builder as posterior probability densities.

TOR 3. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}). Comment on the scientific adequacy of existing and redefined BRPs.

The assessment working group successfully completed this TOR. Yield-per-recruit and spawner biomass-per-recruit reference points were provided as proxies for F_{MSY}, as well as estimates of F_{MSY} and B_{MSY} from the ASPIC model. These later values were compared with those from previous assessments; B_{MSY} increased by 22% and F_{MSY} decreased by 24% from the 2005 assessment. Bias in the BRP's from ASPIC was evaluated both via bootstrapping and by re-estimating them after adding a few more years of assumed data. The reference points appear to be sensitive to relatively small (assumed) changes in CPUE during the next few years.

There was a lot of discussion at the review meeting about the use of equilibrium reference points for populations that do not appear to remain near an equilibrium size. Both the data and model results from this assessment appear to indicate that year class effects (intermittently-produced, large cohorts) can have a significant influence on abundance. If abundance is naturally variable, then assessing its status against a stationary biomass reference level could lead to a population being classified as overfished when the change in abundance is due to a natural fluctuation. The RC expressed a preference for the use of F_{MSY} proxies (e.g. $F_{40\%}$) over the biomass-based reference points for tilefish for this reason.

TOR 4. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

This TOR was completed successfully by the assessment working group. As discussed above, neither the ASPIC nor the SCALE modeling results are completely satisfactory. The WG placed greater emphasis on the ASPIC model results and the RC agreed that this was appropriate because ASPIC was thought to capture the long-term productivity of the population. However, there was consensus that ASPIC would not be able to estimate abundance well when abundance was changing as a result of 1 or 2 large cohorts moving through the stock and, given the lack of fit to the most recent CPUE data, that the most recent biomass estimates were likely biased high.

Within a surplus production model, factors such as fishing selectivity are subsumed into the productivity parameter (the productivity is estimated conditional on how the biomass is removed, which is not observed or estimated). At the request of the RC, the WG reran the ASPIC model during the meeting using only the VTR data (the most recent index series) in an attempt to determine whether there was evidence that the productivity parameter had decreased in recent years, which could potentially occur if fishing practices has changed (this could be thought of as a reverse retrospective analysis to determine whether the early part of the time series was largely dictating the productivity estimates). This model run produced implausibly high estimates of the intrinsic rate of population growth and the length of the data series was quite short for fitting this type of model. No strong conclusions were drawn from this analysis, other than it did not provide evidence for decrease in productivity.

The RC agreed with the WG that it was not possible to determine with sufficient certainty whether the population had rebuilt to B_{TARGET}, because, even though the 2008 point biomass estimate from ASPIC is above B_{MSY}, ASPIC was overestimating recent CPUE and a significant portion of the bootstrapped 2008 biomass estimates were below B_{MSY}. Additionally, if current abundance is largely dependent on one or two large cohorts moving the stock, abundance may decline as they pass out of the stock.

TOR 5. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).

-Provide numerical short-term projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (alternate states of nature).

-If possible, comment on the relative probability of the alternate states of nature and on which projections seem most realistic.

-For a range of candidate ABCs, compute the probabilities of rebuilding the stock by November 1, 2011.

-Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.

The WG completed this TOR by providing deterministic projections based on both the ASPIC model and SCALE model results. Given the uncertainty in the point estimates of 2008 biomass and uncertainty in the future direction of the CPUE indices, these projections have limited utility for predicting future stock size, but are useful for showing the uncertainty in these projections. While the deterministic projections do not fully characterize the uncertainty in future stock size (which would require the use of stochastic projections including uncertainty in both present stock size and productivity levels, as well as random variability in the future productivity), the deterministic projections under a wide range of scenarios (e.g. trends in future CPUE) were very useful in understanding the sensitivities of the model output to small changes in the input data.

Of particular interest was a scenario in which landings remained at current levels and CPUE continued to decline. Under this scenario, the stock would be estimated to have both lower productivity and a higher estimate of B_{MSY}, and would end up being classified as "overfished". This is a plausible scenario given the trend in CPUE during the last three years, and indicates that the stock is vulnerable to becoming classified as "overfished" (and to not have recovered by 2011).

Although the SCALE model was not adopted as the main model in the assessment, the projection results assuming constant recruitment and fishing mortality rates of 0.13 (F_{MSY}) and 0.19 (F_{2008}) both showed population growth, although landings would be nearer half the current TAC of 905 mt.

TOR 6. Review, evaluate and report on the status of the research recommendations offered in recent SARC reviewed assessments. Identify new research recommendations, including recruitment estimation.

This term of reference was completed by the working group by evaluating and reporting on progress made on research recommendations in recent SARC reviewed assessments, and by proposing new research recommendations. I agree with most of these research recommendations.

From my perspective, management advice would be improved if a model able to deal with strong year class effects was developed to the extent that it could be the primary model used in the assessment. Although the WG considered the development of a forward projecting, catch-at-length model to be complete (the SCALE model), greater emphasis was placed on the ASPIC model results in this assessment. Although not listed by the WG as a research recommendation, I encourage continued development of this model based on the belief that an age-structured model would be better able to track the year-to-year variability inherent in the tilefish data.

Part of having an assessment model capable of dealing with the strong year class effects is trusting that the data being input into the model is representative of stock characteristics (abundance, length-frequency, etc.) and not an artifact of changes in fishing practices or some other factor. For this reason, I strongly endorse the WG recommendation to continue the development of the haul-based fishery dependent CPUE index (the study fleet project), not only as it pertains to improving the CPUE index, but also if developed in a way that assists in the interpretation of size-frequency changes through the collection of data on spatial distribution and population size structure. Development of a fishery-based random stratified survey is one way this recommendation could be addressed.

The WG highlighted that there are differences in the dealer, VTR and IVR reports of landings, as well as in market category designation among ports and that developing protocols to address these inconsistencies would be beneficial. Presently the size category information is not being used in the assessment model, although it could be if these inconsistencies were addressed.

4.0 Ocean Quahog Assessment Review

The ocean quahog assessment working group met the TOR's for this assessment. All ocean quahog in US waters are managed as a single stock. The commercial landings and effort are well characterized for this stock. A delay difference model (KLAMZ), fit to data by region (except Southern Virginia) and for the stock as a whole, was used to estimate fishable biomass and fishing mortality. As a whole, the stock is being fished down to towards its B_{MSY} reference point (1/2 of the virgin biomass). The model output compared favorably with both a cumulative catch virtual population model and landings per unit effort data not used in fitting the model. There was a consensus among the RC that the stock is not overfished and that overfishing is not occurring. The RC also agreed that the assessment was a sound scientific sufficient basis for the management of the stock. Quahogs have a somewhat unique life history: they grow slowly during their first years of life, animals 50-100 years are common, specimens 287 and 405 years of age have been documented, recruitment events are regional and infrequent, and the age of

50% selectivity into the fishery is on the range of 13 to 28 years. This biology creates time lags that are outside the planning horizons for most managed activities, a topic that created considerable discussion at the review committee meeting (without leading to definite conclusions).

Meeting the Terms of Reference

TOR 1. Characterize commercial catch including landings, effort, and discards.

This ToR was completed well by the assessment working group. All of the landings are from commercial fisheries and the quality of both the landings and effort appears high due to good compliance with a mandatory logbook program, which include landings, fishing location and fishing effort. There are no significant discards, but a 5% incidental mortality correction factor is applied to the landings. Landings in the US EEZ were high during 1987 to 1996 with a peak of 22,500 mt. Recent landings (since 2006) have been about 15,500 mt.

Spatial patterns for the fishery are also well documented, showing a shift in the fishery to offshore and northern grounds. During the 1980's, nearly all of the landings were from the Delmarva and New Jersey regions, while in 2008, the bulk of the effort and landings (excluding Maine waters) were from the Long Island region.

The commercial selectivity of the fishery is assumed to be similar to that of commercial dredges in Iceland based on similarities in dredge design and towing speeds. Commercial length composition data is collected by port agents from landings samples.

TOR 2. Estimate fishing mortality, spawning stock biomass, and stock biomass for the current and previous years. Characterize uncertainty of the estimates.

This TOR was completed very well by the assessment working group. The analyses were well thought out and thorough. Several data sources were available for the including 12 NEFSC surveys from 1982-2008 (a triennial survey), depletion fishing experiments used to estimate the survey dredge efficiency, and the removals from the commercial fishery. Landings per unit effort were also available, and were used to assess model results rather than being incorporated directly into the model.

The NEFSC survey is a stratified survey and strata are sometimes missed. During this assessment, missing values were "borrowed" from the same strata previous or following survey. As pointed out by the WG, a model-based approach for filling in these cells should be developed. Considerable effort has gone into the analysis of gear effects on the dredge efficiency, particularly effective tow length, to ensure that the surveys are as comparable as possible. Survey dredge efficiency is estimated using depletion estimates. The average survey dredge efficiency in 2008 was about 0.320, although it was found to variable (range: 0.207 to 0.467). This uncertainty could be carried forward through the assessment, although given the very low estimated of fishing mortality, would be unlikely to alter the overall conclusions about status. These results are used to calculate efficiency-corrected, swept-area biomass estimates (ESB) that are an estimate of stock biomass that is independent of the fishery assessment model used to estimate

fishing mortality. Fishing mortality rates were estimated from the ratio of the catch to ESB for each region and year. These estimates were used as a check for model based fishing mortality estimates. A simple virtual population analysis was also used as a check on the model-based estimates.

A forward-projecting assessment model (KLAMZ) is used in this assessment. The model was used for the Delmarva, New Jersey, Long Island and Southern New England regions as well as for the exploitable stock (entire stock less Georges Bank) during 1977-2008. In the 2008 assessment, a step function for the recruitment pattern was used for the Long Island, Southern New England and the exploitable stock, in contrast with the constant recruitment used for other regions. Likelihood profiles were used to choose the break year for the step function. Model output was compared to the VPA and ratio method of estimating fishing mortality, and the trends were compared to trends in landings per unit effort. Results are not sensitive to the modeling approach. Uncertainty in the model estimates was evaluated using the delta method and by bootstrapping.

Overall, I found these approaches and cross checks to be very thorough and believe they resulted in a sound basis for management advice. Taken together, the results indicate a decline in fishable biomass by 2008 in the Delmarva region to about 30% of the 1978 biomass level, to about 40% in the New Jersey region, to about 89% in the Long Island region, to about 99% in the Southern New England region and an increase of about 13% on Georges Bank. This pattern follows the shift northward and offshore in the fishery.

TOR 3. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}). Comment on the scientific adequacy of existing and redefined BRPs.

This TOR was completed by the assessment working group. Given the slow dynamics of the quahog stock, the population response to fishing cannot be determined, hence it is not possible to relate the reference points directly to MSY. Given this issue, there was a shift away from reference to MSY in defining the reference points in this assessment. Based on a review of F_{MSY} reference points for relatively long-lived west coast groundfish species, a new $F_{threshold}$, F45% (0.0219) was recommended. The new recommended biomass target of 1.837 million mt is one-half of the 1978 pre-fishery biomass. The new recommended $B_{threshold}$ is 40% of the 1978 biomass (1.470 million mt). This latter value is ad hoc, but given that it is not known how the stock will respond to fishing, is probably better than the 25% value currently used.

At present, these reference points are applied to the exploitable stock. Although it was not one of the TOR's there was discussion at the SARC about whether these reference points should be applied to the entire stock (including Georges Bank). While it is not known at present (and genetic information appears to indicate otherwise based on discussion at the SARC), the possibility also exists that there is population sub-structuring such that management on a smaller spatial scale could be required to maintain local abundance. If so, the biomass based reference points would need to be derived at this smaller scale.

4. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

This TOR was completed successfully by the assessment working group. The ocean quahog stock is not currently overfished and overfishing is not occurring. Fishable stock biomass during 2008 was 2.905 million mt, well above the target of 1.837 million mt (one-half of the 1978 pre-fishery biomass). During 2008, F= 0.0100 for the exploitable portion of the EEZ (excluding Georges Bank) and 0.0056 for the whole EEZ, well below the $F_{threshold}$ of 0.0219.

TOR 5. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).

- **a. Provide numerical short-term projections (3-4 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (alternate states of nature).**
- **b. If possible, comment on the relative probability of the alternate states of nature and on which projections seem most realistic.**
- **c. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.**

This TOR was completed successfully by the assessment working group. The WG used both deterministic and stochastic stock projections based on the KLAMZ model output. KLAMZ model projections were run with varying "states of nature": a range of possible values for natural mortality (M=0.015, 0.02 and 0.025) and biomass levels. It is not possible to comment on the probability of these alternate states. The projections were also run with four landings policies based on the current management plan. These analyses appear technically correct. These analyses indicate that low biomass conditions are not likely to occur by 2015 at current catch levels under any of the states of nature. As stated by the WG, ocean quahogs are an unproductive stock that is vulnerable to becoming overfished. The projections indicate that biomass is expected to decline slowly during 2010 to 2015 under most of the harvest rates considered in the projections.

TOR 6. Review, evaluate and report on the status of SARC/Working Group research recommendations listed in recent SARC reviewed assessments. Identify new research recommendations.

This TOR was completed by the assessment working group and I accept most of their recommendations. There are two questions that re-occurred throughout this assessment. The first is the issue of how the stock will respond to exploitation. The currently used survey gear does not adequately capture small quahogs, so there is limited information about whether spawning, settling and juvenile growth is occurring in areas where fishing has occurred, and whether it is at a level indicative of a compensatory response. Incorporation of a sampling protocol (e.g. the use of a grab sampler, box core or some other kind of gear) into the surveys could allow some evaluation of the recruitment process and how it is changing as a result of fishing without waiting for the quahogs to grow to a size where they would be representatively sampled using the survey dredge. The second question relates to the spatial scale of management (during

discussions this question was more directly linked to whether the reference points should be applied to the whole stock or the exploited stock rather than the recommendation made here). The population dynamics theory upon which fishery management is based (that of surplus production), typically assumes a single, closed population. At present, the population structure in the quahog stock is not known. If fishing is occurring on two or more closed populations with differ productivities being assessed as a single stock, the lower productivity populations may be placed at-risk as a result of the combined assessment and management. Alternately, if the population is not closed, such that a significant portion of the recruitment is from another area (as alluded to in the Maine Quahog assessment – working paper B1-A Appendix 2), then protection of spawner biomass in that area may be less important allowing for higher exploitation rates in these areas. The spatial scale of recruitment processes and population structuring could potentially be examined modeling larval transport or further genetic research.

5.0 Weakfish Assessment Review

The weakfish assessment working group met the TOR's for this assessment. The WG provided multiple analyses to estimate abundance, total mortality and fishing mortality, including an ADAPT VPA, an analysis of survey data as abundance indices, and a Steele-Henderson production model including predation effects. The analyses indicate that abundance has declined markedly, total mortality is high, non-fishing mortality has recently increased and that the stock is currently in a depleted state. Taken as a whole, the considerable amount of work done by this WG does provide a sound basis for management advice.

Meeting the Terms of Reference

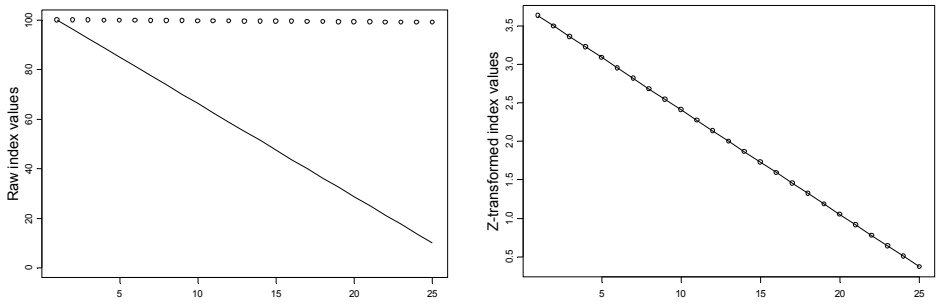
> **TOR 1: Evaluate biases, precision, uncertainty, and sampling methodology of the commercial and recreational catch (including landings and discards) and effort.**

The assessment working group completed this TOR successfully. Data from commercial harvests, commercial discards, recreational harvests, and recreational discards were evaluated. Discard mortality rates were assumed to be 100% of the commercial fisheries and 10% (based on studies) for the recreational fishery. Harvest and discard estimates were stratified by region (north/south), year, and season, and commercial harvests was further stratified by state and gear. Stratum specific biological data was also summarized. Age-length keys were used to convert numbers-at-length to numbers-at-age, which were summed across strata to obtain annual estimates of total numbers of weakfish by age. The WG described several sources of uncertainty and potential biases in the catch-at-age data, such as under-reporting, survey sample sizes, as well as others. Overall, the catch data evaluation was well done.

> **TOR2: Evaluate precision, geographical coverage, representation of stock structure, and relative accuracy of the fisheries independent and dependent indices of abundance. Review preliminary work on standardization of abundance indices.**

The assessment group also completed this TOR well. They evaluated five fishery independent surveys based on geographic coverage, ability to accurately track weakfish abundance, and survey precision, among other factors, as well as ten young-of-the year surveys and two fishery

dependent abundance indices. The RC agreed with the selection of the indices to include in this assessment. The indices appeared to be more or less consistent with each other in the sense that abundance increased and decreased roughly in synchrony, although because the data were Z-transformed prior to the comparison, the magnitudes of the abundance changes were not directly comparable. The effect of this transformation can be shown by comparing two time series before and after transformation: the first declining by 90% (lines in the figure) and the second decreasing by only one percent (points in the figure). As shown in the panel on the left, the two indices convey quite different pictures of abundance trends, whereas after transformation the trends being conveyed are the same:



With one exception, the Z-transformed indices were not used for fitting models. In the one case in which it was used, the WG reran the model using the raw index values and found that the effect of the transformation on the model results was relatively minor. As a result of this evaluation, this technical issue was not considered serious enough to call into question the assessment results. Overall, I agree with the conclusions reached by the WG with respect to this TOR.

TOR3: Evaluate the ADAPT VPA catch at age modeling methods and the estimates of F, Z, spawning stock biomass, and total abundance of weakfish produced, along with the uncertainty and potential bias of those estimates. Review the severity of retrospective pattern.

The assessment working group completed this TOR by presenting several runs using different sets of tuning indices. These runs provided consistent estimates of parameter values for the years 1982 to 2002. More recent estimates were confounded by a prominent retrospective pattern. The WG did not support attempts to correct for the retrospective pattern because estimates of fishing mortality are dependent on natural mortality which was shown to have changed in subsequent analyses (see below). Given the issues associated with the ADAPT model runs, the WG placed greater emphasis on the index-based analyses that do not require assumptions about natural mortality, a position endorsed by the RC.

In an earlier assessment the WG was not encouraged to pursue alternative age-structured models, primarily because the reviewers did not think it would help with the retrospective pattern. However, a forward projecting, statistical, age-structured model would provide greater flexibility for dealing with changes in natural mortality and I therefore support the WG recommendation that this kind of approach be investigated for future assessments.

TOR4: Evaluate the index-based methods and the estimates of F, ages 1+ stock biomass, surplus production, and time-varying natural mortality of weakfish produced, along with the uncertainty of those estimates. Determine whether these techniques could complement or substitute for age-based modeling for management advice.

The assessment working group successfully completed this TOR. They used an interesting method of estimating total mortality from the index data, ''anchoring'' the index series to abundance by comparison with the converged portion of the VPA, and estimating fishing mortality by combining it with the catch data. I found the analyses to be well thought out and thorough, and thought the analyses provided convincing evidence of a change in natural mortality.

The WG found that the unexpected drop in weakfish surplus production after 1999 coincided with a sharp rise in the coast-wide abundance of two potential predators: striped bass and spiny dogfish. As the WG correctly pointed out, these correlations do not imply that predation is the cause. During the meeting, the WG was encouraged to attempt a simple time series analysis by differencing the weakfish and predator data series (N_{t-1}-N_t) prior to performing a regression. Due to time constraints these kinds of approaches were not fully explored at the SARC. Given that understanding the causes of the changes in natural mortality and surplus production are likely key to affecting their recovery, the WG is encouraged to explore this and other time series approaches (e.g. ARIMA models) to dealing with the trends in these data series.

Although in its current form the index model does not provide management reference points, the model output could be extended to be better linked to management advice. As implemented, if there are errors in the converged portion of the ADAPT VPA, these errors would be carried forward in future assessments if the approach was continued to be used. A statistical model integrating the VPA and index approaches would alleviate both these concerns.

TOR5: Evaluate testing of fishing and additional trophic and environmental covariates and modeling of hypotheses using biomass dynamic models featuring multiple indices blended into a single index with and without a Steele-Henderson (Type III) predator-prey extension. Evaluate biomass dynamic model estimates of F, ages 1+ stock biomass, surplus production, time-varying natural mortality, and biological reference points along with uncertainty of those estimates. Advise on burden of proof necessary for acceptance of alternatives to constant M and whether these biomass dynamic techniques could complement or substitute for age-based modeling for management advice.

TOR's 5 to 7 are very similar in nature and are treated together (see below).

TOR6: Evaluate AIC-based hypothesis testing of fishing and additional predation-competition effects using multi-index biomass dynamic models with and without prey-based, predator-based, or ratio dependent predator-prey extensions. Evaluate biomass dynamic model estimates of F, ages 1+ stock biomass, surplus production, time-varying natural mortality, and biological reference points along with uncertainty of those estimates. Advise on burden of proof necessary for acceptance of alternatives to

constant M and whether these biomass dynamic techniques could complement or substitute for age-based modeling for management advice.

TOR's 5 to 7 are very similar in nature and are treated together (see below).

TOR7: Review evidence for constant or recent systematic changes in natural mortality, productivity, and/or unreported removals.

TOR's 5 to 7 were completed by the assessment working group. The WG provided several analyses to examine the evidence for change in natural mortality including the index-based model discussed above as well as an age-aggregated Steele-Henderson production model used to examine the joint effects of fishing and predation from striped bass, spiny dogfish, bluefish and summer flounder. Environmental conditions were also examined, including decadal shifts in mean sea surface water temperature and deviations in the winter North Atlantic Oscillation Index. The WG concluded that most of the statistical evidence supported the predation hypothesis involving striped bass and spiny dogfish as the primary factor behind the recent and unexpected decline in weakfish productivity. While the WG may well be right, the correlations used as partial support for these conclusions do indicate that weakfish surplus production has gone down while striped bass and spiny dogfish abundance has gone up, but do not provide a casual link. The previously mentioned times series approaches intended to remove the trends from the data would provide stronger evidence in support of this hypothesis. However, if a statistically significant relationship remained after de-trending the data, a counter argument would still exist: a set of factors could exist that are influencing both weakfish and striped bass dynamics concurrently. The WG did acknowledge that although most the evidence does support the predation hypothesis, other factors such as unreported commercial and recreational landings, disease, toxins and parasitism cannot be ruled out at this time. Personally, I found that taken together, the analyses showing that fishing is not the primary cause of the present high total mortality were convincing.

TOR8: Estimate biological reference points using equilibrium and non-equilibrium assumptions and evaluate stock status relative to these BRPs.

The working group met this TOR and provided several reference points against which status could be assessed. Given the changing natural mortality, the equilibrium reference points requested in the TOR are likely inapplicable to this stock at present. Nonequilibrium reference points calculated with the Steele-Henderson model indicate that F_{MSY} has declined in recent years as predation/competition has eroded the amount of weakfish productivity ''available'' to fishing.

The suggestion to use reference points based on total mortality (as opposed to fishing mortality) is somewhat troublesome as it does not link directly to any specific management action. If the causes of the high total mortality were identified and a plan was developed to reduce it, the use of a total mortality reference point might make sense as the reference value would then be linked to a goal in the plan. At present there appears to be limited scope for reducing total mortality via fishery management.

TOR9: Review stock projections and impacts on the stock under different assumptions of fishing and natural mortality.

This TOR was completed satisfactorily by the WG, in the sense that projections were provided using the three main models used in the assessment. These projections could not be used to accurately predict future abundance given the issues with the models (e.g. the retrospective pattern in the VPA) as well as uncertainty in future trends in natural mortality. The models do show that without a change in natural mortality, rebuilding to past abundance is unlikely. Additionally, given that present fishing mortality appears to be a relatively small component of total mortality, there is little scope for substantially increasing abundance via more restrictive fisheries management (although fisheries closures do have the potential to lead to increased abundance under some of the natural mortality scenarios investigated).

TOR10: Make research recommendations for improving data collection and assessment.

The assessment working group did complete this TOR. They provided 11 high priority, 5 medium priority and three low priority research recommendations. My specific recommendations were discussed under the relevant TORs (statistical catch-at-age modeling and the use of models that account for the time series nature of the data when assessing the causes of the decline).

6.0 Acknowledgments

This assessment review meeting was a very pleasant experience as a result of the hard work of many people. I thank Paul Rago and particularly Jim Weinberg for their hospitality and advice during the meeting. Thanks to Pat Sullivan for an excellent job in chairing the meeting, particularly keeping us on schedule. I also wish to thank the other panel members, Mike Bell and Sven Kupschus, for stimulating discussions both during and around the meeting, Manoj Shivlani for his work coordinating the review and his assistance with travel arrangements, as well as the assessment teams led by Paul Nitschke (tilefish), Larry Jacobson (ocean quahog), and Jeff Brust (weakfish) for providing clear and thorough assessment documents and presentations at the meeting.

7.0 Appendices

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

Appendix 1: Review Panel Membership.

Review Panel Membership

Member	Primary Affiliation
Patrick Sullivan, chair	Cornell University, New York
Mike Bell	Heriot-Watt University, Institute of Petroleum Engineering, Scotland
Jamie Gibson	Department of Fisheries and Oceans, Canada
Sven Kupschus	The Centre for Fisheries and Aquaculture Science, England

Appendix 2: CIE Statement of Work.

Statement of Work

(Subtask T007-05, v 22 December 2008)

External Independent Peer Review by the Center for Independent Experts

SARC 48: Tilefish, Ocean quahog, Weakfish Benchmark Stock Assessments

Meeting Date: June 1-4, 2009

Statement of Work (SOW) for CIE Panelists
(including a description of SARC Chairman's duties)

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract to provide external expertise through the Center for Independent Experts (CIE) to conduct impartial and independent peer reviews of NMFS scientific projects. This Statement of Work (SoW) described herein was established by the NMFS Contracting Officer's Technical Representative (COTR) and CIE based on the peer review requirements submitted by NMFS Project Contact. CIE reviewers are selected by the CIE Coordination Team and Steering Committee to conduct the peer review of NMFS science with project specific Terms of Reference (ToRs). Each CIE reviewer shall produce a CIE independent peer review report with specific format and content requirements (**Annex 1**). This SoW describes the work tasks and deliverables of the CIE reviewers for conducting an independent peer review of the following NMFS project.

Project Description: The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication.

The SARC48 review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the Science and Statistics Committee (SSC) of the New England or Mid-Atlantic Fishery Management Council. The panel will convene at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts during June 1-4, 2009 to review three assessments (tilefish (*Lopholatilus chamaeleonticeps*), ocean quahog (*Arctica islandica*), and weakfish (*Cynoscion regalis*)). In the days following the review of the assessment, the panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**. The summary report format is attached as **Annex 4.**

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs 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. CIE reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models and Biological Reference Points. Expertise should include statistical catch-at-age and catch-at-length models, traditional VPA approaches, delay-difference models, and the implications of spatial harvesting patterns. Experience with comparative studies of these approaches is especially valuable. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Experience with the biology and population dynamics of species on the agenda would be useful.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during June 1-4, 2009.

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

The CIE's deliverables shall be provided according to the schedule of milestones listed below**.** The CIE reviewers, along with input and leadership from the SARC Chairman, will write the SARC Summary Report. In addition, each CIE reviewer will write an individual independent review report. These reports will provide peer-review information for a presentation to be made by NOAA Fisheries at meetings of the New England and Mid-Atlantic Fishery Management Councils. The SARC Summary Report shall be an accurate representation of the SARC panel viewpoint on how well each SAW Term of Reference was completed (please refer to Annex 2 for the SAW Terms of Reference).

The three CIE reviewers shall conduct an impartial and independent peer review in accordance with the Terms of Reference (ToR) herein. The three SARC CIE reviewers' duties shall occupy a maximum of 14 days per person (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; and several days following the open meeting to contribute to the SARC Summary Report and to produce the Independent CIE Reports).

Not covered by the CIE, the SARC chair's duties should occupy a maximum of 14 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation).

Charge to SARC panel

The panel is to determine and write down whether each Term of Reference of the SAW (see Annex 1) was or was not completed successfully during the SARC meeting. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. Where possible, the chair shall identify or facilitate agreement among the reviewers for each Term of Reference of the SAW.

If the panel rejects any of the current Biological Reference Point (BRP) proxies for B_{MSY} and F_{MSY}, the panel should explain why those particular proxies are not suitable and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs are the best available at this time.

Roles and responsibilities

(1) Prior to the meeting

(SARC chair and CIE reviewers)

Review the reports produced by the Working Groups and read background reports.

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, and contact details) 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, foreign national security clearance, and information concerning other pertinent meeting arrangements. The NMFS Project Contact is also 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 COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations (available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will attempt to provide the CIE reviewers all necessary background information and reports for the peer review. This will be done by electronic mail or an FTP site. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. The CIE reviewers shall read all documents in preparation for the peer review.

(2) During the Open meeting

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs. **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.** Each CIE 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 ToRs as specified in the contract SoW. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

(SARC chair)
Act as chairperson, where duties include control of the meeting, coordination of presentations and discussion, making sure all Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For the assessment, review both the Assessment Report and the Assessment Summary Report.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to discuss the stock assessment and to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(SARC CIE reviewers)
For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a reviewer's point of view, determine whether each Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a reviewer considers any existing Biological Reference Point proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(3) After the Open meeting

(SARC CIE reviewers)
Each CIE reviewer shall prepare an Independent CIE Report (see Annex 1). This report should explain whether each Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the "Charge to SARC panel" statement.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.

During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent CIE Report produced by each reviewer.

The Independent CIE Report can also be used to provide greater detail than the SARC Summary Report on specific Terms of Reference or on additional questions raised during the meeting.

(SARC chair)
The SARC chair shall prepare a document summarizing the background of the work to be conducted as part of the SARC process and summarizing whether the process was adequate to complete the Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report.

(SARC chair and CIE reviewers)
The SARC Chair and CIE reviewers will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar or a consensual view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion.

The SARC Summary Report (please see Annex 3 for information on contents) should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

The contents of the draft SARC Summary Report will be approved by the CIE reviewers by the end of the SARC Summary Report development process. The SARC chair will complete all final editorial and formatting changes prior to approval of the contents of the draft SARC Summary Report by the CIE reviewers. The SARC chair will then submit the approved SARC Summary Report to the NEFSC contact (i.e., SAW Chairman).

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.

Other Tasks – Contribution to Summary Report: Each CIE reviewer will assist the Chair of the panel review meeting with contributions to the Summary Report. CIE reviewers are not required to reach a consensus, and should provide a brief summary of their views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

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 Woods Hole, MA, from June 1-4, 2009 as called for in the SoW, and conduct an independent peer review in accordance with the ToRs (Annex 2);
3) No later than June 19, 2009, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to David Sampson 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;

Appendix 2: CIE Statement of Work.

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

27 April 2009	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
18 May 2009	NMFS Project Contact will attempt to provide CIE Reviewers the pre-review documents by this date
1-4 June 2009	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
4 June 2009	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
19 June 2009	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
22 June 2009	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
29 June 2009	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
2 July 2009	CIE submits CIE independent peer review reports to the COTR
9 July 2009	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

* The SARC Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the SARC chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion.

NEFSC staff and the SAW Chairman will make the final SARC Summary Report available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers, which will serve as a SAW Assessment Report.

Modifications to the Statement of Work: Requests to modify this SoW must be made through the Contracting Officer's Technical Representative (COTR) who submits the modification for approval to the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the CIE within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and Terms of Reference (ToR) of the SoW as long as the role and ability of the CIE reviewers to complete the SoW deliverable in accordance with the ToRs and deliverable schedule are not adversely impacted. The SoW and ToRs cannot be changed once the peer review has begun.

Appendix 2: CIE Statement of Work.

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. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (the CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards: (1) each CIE report shall have the format and content in accordance with Annex 1, (2) each 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 notification of acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the approved CIE reports to the NMFS Project Contact and regional Center Director.

Key Personnel:

William Michaels, Contracting Officer's Technical Representative (COTR)
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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

NMFS Project Contact:

Dr. James Weinberg
National Marine Fisheries Service, NOAA
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
james.weinberg@noaa.gov Phone: 508-495-2352

Dr. Nancy Thompson, NEFSC Science Director
National Marine Fisheries Service, NOAA
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
nancy.thompson@noaa.gov Phone: 508-495-2233

Appendix 2: CIE Statement of Work.

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 whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and Conclusions and Recommendations in accordance with the ToRs. For each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, the Independent Review Report should state why that Term of Reference was or was not completed successfully. To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice.

a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a concise summary of whether they accept or reject the work that they reviewed, and to explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

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

c. Reviewers should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification.

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

e. The CIE independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include as separate appendices as follows:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of the CIE Statement of Work
Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Appendix 2: CIE Statement of Work.

ANNEX 2:

Assessment Terms of Reference for SAW/SARC-48, June 1-4, 2009

(file: 4/27/09)

A. Tilefish

1. Characterize the commercial catch including landings, effort and discards. Characterize recreational landings. Evaluate utility of study fleet results as improved measures of CPUE.

2. Estimate fishing mortality and total stock biomass for the current year, and for previous years if possible, and characterize the uncertainty of those estimates. Incorporate results of new age and growth studies.

3. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}). Comment on the scientific adequacy of existing and redefined BRPs.

4. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

5. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).
    a. Provide numerical short-term projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (alternate states of nature).
    b. If possible, comment on the relative probability of the alternate states of nature and on which projections seem most realistic.
    c. For a range of candidate ABCs, compute the probabilities of rebuilding the stock by November 1, 2011.
    d. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.

6. Review, evaluate and report on the status of the research recommendations offered in recent SARC reviewed assessments. Identify new research recommendations, including recruitment estimation.

B. Ocean quahog

1. Characterize commercial catch including landings, effort, and discards.

2. Estimate fishing mortality, spawning stock biomass, and stock biomass for the current and previous years. Characterize uncertainty of the estimates.

3. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}). Comment on the scientific adequacy of existing and redefined BRPs.

4. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

5. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).
    a. Provide numerical short-term projections (3-4 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (alternate states of nature).
    b. If possible, comment on the relative probability of the alternate states of nature and on which projections seem most realistic.
    c. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.

6. Review, evaluate and report on the status of SARC/Working Group research recommendations listed in recent SARC reviewed assessments. Identify new research recommendations.

Appendix 2: CIE Statement of Work.

C. Weakfish (Final weakfish TORs approved by Weakfish Management Board 4-24-09)

1. Evaluate biases, precision, uncertainty, and sampling methodology of the commercial and recreational catch (including landings and discards) and effort.

2. Evaluate precision, geographical coverage, representation of stock structure, and relative accuracy of the fisheries independent and dependent indices of abundance. Review preliminary work on standardization of abundance indices.

3. Evaluate the ADAPT VPA catch at age modeling methods and the estimates of F, Z, spawning stock biomass, and total abundance of weakfish produced, along with the uncertainty and potential bias of those estimates. Review the severity of retrospective pattern.

4. Evaluate the index-based methods and the estimates of F, ages 1+ stock biomass, surplus production, and time-varying natural mortality of weakfish produced, along with the uncertainty of those estimates. Determine whether these techniques could complement or substitute for age-based modeling for management advice.

5. Evaluate testing of fishing and additional trophic and environmental covariates and modeling of hypotheses using biomass dynamic models featuring multiple indices blended into a single index with and without a Steele-Henderson (Type III) predator-prey extension. Evaluate biomass dynamic model estimates of F, ages 1+ stock biomass, surplus production, time-varying natural mortality, and biological reference points along with uncertainty of those estimates. Advise on burden of proof necessary for acceptance of alternatives to constant M and whether these biomass dynamic techniques could complement or substitute for age-based modeling for management advice.

6. Evaluate AIC-based hypothesis testing of fishing and additional predation-competition effects using multi-index biomass dynamic models with and without prey-based, predator-based, or ratio dependent predator-prey extensions. Evaluate biomass dynamic model estimates of F, ages 1+ stock biomass, surplus production, time-varying natural mortality, and biological reference points along with uncertainty of those estimates. Advise on burden of proof necessary for acceptance of alternatives to constant M and whether these biomass dynamic techniques could complement or substitute for age-based modeling for management advice.

7. Review evidence for constant or recent systematic changes in natural mortality, productivity, and/or unreported removals.

8. Estimate biological reference points using equilibrium and non-equilibrium assumptions and evaluate stock status relative to these BRPs.

9. Review stock projections and impacts on the stock under different assumptions of fishing and natural mortality.

10. Make research recommendations for improving data collection and assessment.

Appendix 2: CIE Statement of Work.

Appendix to the TORs:

Clarification of Terms
used in the SAW/8SARC Terms of Reference

(The text below is from DOC National Standard Guidelines, Federal Register, vol. 74, no. 11, January 16, 2009)

On "Acceptable Biological Catch":

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability":

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce MSY and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Appendix 2: CIE Statement of Work.

Annex 3: Tentative Agenda

48th Northeast Regional Stock Assessment Workshop (SAW 48)
Stock Assessment Review Committee (SARC) Meeting

June 1-4, 2009

Stephen H. Clark Conference Room – Northeast Fisheries Science Center
Woods Hole, Massachusetts

AGENDA* (version: 5-27-09)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Monday, 1 June

10:00 – 10:30 AM
Opening
Welcome — **James Weinberg**, SAW Chairman
Introduction — **Patrick Sullivan**, SARC Chairman
Agenda
Conduct of Meeting

10:30 - Noon Tilefish Assessment Presentation (A)
Paul Nitschke/ Mike Palmer/ Tiffany Vidal — **Jamie Gibson** — **Palmer/Vidal**

Noon – 1:00 PM Lunch

1:00 – 2:30 PM SARC Discussion of Tilefish (A)
Patrick Sullivan, SARC Chairman

2:30 – 3:00 PM Break

3:00 - 5:00 PM Ocean quahog Assessment Presentation (B)
Larry Jacobson/ Toni Chute — **Mike Bell** — **Ralph Mayo**

5:00 – 6:00 PM SARC Discussion of Ocean quahog (B)
Patrick Sullivan, SARC Chairman

Appendix 2: CIE Statement of Work.

Tuesday, 2 June

9:00 – 10:15 AM	Revisit Tilefish Assessment with Presenters (A)
10:15 – 10:30 AM	Break
10:30 - Noon	Revisit Ocean Quahog Assessment with Presenters (B)
Noon – 1:00 PM	Lunch
1:00 – 3:45 PM	Weakfish Assessment Presentation (C) **Jeff Brust/ Vic Crecco/ Jim Uphoff** **Sven Kupschus** **Russ Allen**
3:45 – 4:00 PM	Break
4:00 – 5:30 PM	SARC Discussion of Weakfish (C) **Patrick Sullivan**, SARC Chairman

Wednesday, 3 June

9:00 – 10:15 AM	Revisit Weakfish Assessment with Presenters (C)
10:15 – 10:30 AM	Break
10:30 - Noon	Tilefish follow up + review Assessment Summary Report (A)
Noon – 1:00 PM	Lunch
1:00 – 3:00 PM	Ocean qua. follow up + review Assessment Summary Report (B)
3:00 – 3:15 PM	Break
3:15 – 5:15 PM	Weakfish follow up + review Assessment Summary Report (C)

Thursday, 4 June

9:00 – 10:15 AM	Final Revisits with presenters, if needed (A, B, C)
10:15 – 10:30 AM	Break
10:30 AM – 5 PM	SARC Report writing. (closed meeting)

*Times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public, except where noted.

ANNEX 4: Contents of SARC Summary Report

1.

The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If the CIE reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

2.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, include recommendations and justification for alternative proxies. If such alternatives cannot be identified, then indicate that the existing BRPs are the best available at this time.

3.

The report shall also include the bibliography of all materials provided during the SAW, and any papers cited in the SARC Summary Report, along with a copy of the CIE Statement of Work.

The report shall also include as a separate appendix the Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 3: Bibliography of Materials Provided for Review.

Stock	Working Paper	Title	Author(s)	Number of Pages
Tilefish:	A-1	Assessment of Golden tilefish	Southern Demersal Working Group	121
	A-1, Appendix 1	An overview of the tilefish data collected through the NEFSC Study Fleet Project	Palmer, Ball, Anderson, Conboy, Moser	37
	A-1, Appendix 2	Evaluating shifts in size and age at maturity of Golden tilefish from the Mid-Atlantic Bight	Vidal	23
	A-1, Appendix 3	Model Output	Nitschke	11
	A-2	Golden tilefish Assessment Summary Report	Nitschke	14
Tilefish Background Papers:	A-3	Assessment of Golden tilefish (2005)	Southern Demersal Working Group	101
	A-4	Golden tilefish Assessment Summary Report for 2005		8
	A-5	SARC 41 Chair's Report to the CIE (2005)	Jones	29
Ocean quahog:	B-1	Stock Assessment for Ocean quahogs	Invertebrate Subcommittee	175
	B-1a	Ocean quahog Appendix Report	Invertebrate Subcommittee	100
	Appendix 1	Invertebrate Working Group		
	Appendix 2	Ocean quahog resources in Maine waters		
	Appendix 3	Clam dredge performanc		
	Appendix 4	2008 Cooperative Industry Surfclam/Ocean quahog survey		
	Appendix 5	Maps of clam survey catches 1980-2008		
	Appendix 6	KLAMZ assessment model details		
	Appendix 7	West Coast Harvest Policy		
	Appendix 8	Updated shell length/meat weight		
	B-2	Assessment Summary Report for Ocean quahogs		

Appendix 3: Bibliography of Materials Provided for Review.

Stock	Working Paper	Title	Author(s)	Number of Pages
Ocean quahog Background Papers				
	B-3	SARC 44 Assessment Report (2005)	Invertebrate Subcommittee	271
	B-4	2006 Ocean quahog Assessment Summary Report		13
	B-5	SARC 44 Summary Report for CIE (2006)	Jones	64
	B-6	F35% Revisited 10 Years Later	Clark	7
Weakfish	C-1	Weakfish Stock Assessment Report	ASMFC Weakfish Technical Committee	281
	C1a (App C1-C5) C-1	Weakfish Tech. Committees response to Data Poor Meeting comments	ASMFC Weakfish Technical Committee	143
	Appendix C-2	Proportional Stock Density Indices for Weakfish		
	Appendix C-3	SAS-based application of the Harvest Control Model to conduct Weakfish stock projections		
	Appendix C-4	Index Standardization		
	Appendix C-5	Preferred Run ADAPT Output		
	C-2	Weakfish Assessment Summary Report		10
Weakfish Background Papers	C-3	2004 Assessment		419
		2006 Assessment		
		Estimating Discards		
		Population Structure		
	C-4	Report by the Peer Review Panel for the Northeast Data Poor Stocks Working Group	Miller	9