52th Northeast Regional Stock Assessment Review Committee
6 – 10 June 2011
Northeast Fisheries Science Center
Wood's Hole, MA

SARC 52
SUMMARY REPORT

6 July 2011

Review Committee
Patrick J. Sullivan (chair)
Noel Cadigan
John Casey
Cynthia Jones

INTRODUCTION

Background

The 52nd Stock Assessment Review Committee (hereafter referred to as the Review Committee) convened at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA on June 6th through 10th, 2011 to review the benchmark stock assessments of the three winter flounder (*Pseudopleuronectes americanus*) stocks in the Southern New England/Mid-Atlantic (SNE/MA), Georges Bank (GBK), and Gulf of Maine (GOM) regions.

The Review Committee was composed of Dr. Patrick J. Sullivan (Chair and member of the New England Fisheries Management Council's Scientific and Statistical Committee) and three scientists from the Center for Independent Experts: Dr. John Casey (CEFAS), Dr. Noel Cadigan (Dept. of Fisheries and Oceans, Canada), and Dr. Cynthia Jones (Old Dominion University).

The SARC was assisted by the NEFSC SAW Chairman Dr. James Weinberg and his staff, and Dr. Paul Rago (NEFSC). Documentation for all three winter flounder assessments was prepared by the Southern Demersal Working Group (SDWG). The Southern New England/Mid-Atlantic winter flounder assessment was presented at the meeting by Mark Terceiro with Tony Wood acting as rapporteur. The Georges Bank winter flounder assessment was presented by Lisa Hendrickson with Toni Chute acting as rapporteur. The Gulf of Maine winter flounder assessment was presented by Paul Nitschke with Jessica Blaylock acting as rapporteur.

Review of Activities

About two weeks before the meeting, assessment documents and supporting materials were made available to the Review Committee via an ftp server. On the morning of the meeting, the Review Committee met with Drs. Weinberg and Rago to discuss the meeting agenda, reporting requirements, and meeting logistics. During the meeting, all documents were made available electronically and in print.

The meeting opened on Monday afternoon with a welcome and introductions by Drs. Weinberg and Sullivan. The SNE/MA winter flounder assessment was presented that afternoon, after the introductions. On Tuesday morning the GBK winter flounder assessment was presented and on Tuesday afternoon the GOM winter flounder assessment was presented. Wednesday and Thursday allowed for follow-up discussion

on all three assessments and the Review Committee asked further questions and used the remaining time to arrive at a consensus on the content of the review. Friday was spent developing the report.

SARC Process and General Conclusions

The Review Committee reached consensus on all the Terms of Reference for each winter flounder stock. It acknowledges the significant work that the Southern Demersal Working Group had undertaken in preparing and presenting the assessments. It also appreciates the professionalism and cooperation of NEFSC staff at the SARC meeting which significantly assisted the peer review. The reviews of each assessment are briefly summarized below with more details provided in the subsequent sections.

The Terms of Reference for the SNE/MA winter flounder assessment were satisfactorily addressed. The survey age-aggregated indices are declining faster than the preferred model would predict in the last decade for the SNE/MA stock, which may point to a time varying M. There was some concern that the exploitable biomass estimates might be biased high because the weights at age are derived from the catch and thus the size at age may be conflated with selectivity. The statistical catch-age model used for the SNE/MA assessment appeared to be a scientifically credible approach and provided a reasonable basis for fisheries management advice. The Review Committee agreed with the SDWG that in 2010 the SNE/MA winter flounder stock was overfished but overfishing was not occurring.

The Terms of Reference for the GBK winter flounder assessment were satisfactorily addressed. The VPA model used for the GBK assessment appeared to be a scientifically credible approach and provided a reasonable basis for fisheries management advice. In the future, a statistical catch-age model should be considered for the GBK stock as there may be more uncertainty here associated with catch and discards than would be appropriate for the assumption of true known catches made in a VPA type of analysis, however there are pros and cons to using either of these methods. The Review Committee agreed with the SDWG that in 2010 the GBK winter flounder stock was not overfished and overfishing was not occurring.

The Terms of Reference for the GOM winter flounder assessment were partially addressed. The results from the base-case and alternative statistical catch-age models presented and explored at the review meeting appeared inconsistent with trends in the catch and survey data. Uncertainty about data quality, possible changes in productivity,

and possible shifts in distribution of the population led the Review Committee to seek other mechanisms for providing management advice. The SDWG proposed a 30+ cm swept area-biomass estimation method as a basis for providing management advice. There was consensus among the Review Committee to use this method until a model/data configuration can be found that is satisfactory. The Review Committee agreed with the SDWG that, based on the swept-area method, overfishing does not appear to be taking place in 2010 the GOM winter flounder stock, however it is not possible under the current configuration to tell whether or not the stock is overfished.

The Review Committee and NEFSC scientists developed a method at the meeting for combining information on winter flounder across regions to help inform the spawner-recruit relationships used in developing projections and biological reference points. The method is described in Appendix 1 below and uses likelihood-based AIC methods to find a reasonable compromise between a spawner-recruit relationship based on combined data sources and the spawner-recruit estimates associated with the individual stocks of winter flounder. This method maximizes the fit to both the SNE/MA and GBK datasets while minimizing the differences between relationships in the adjoining regions.

Winter flounder

The SARC 52 review was rather unique in that all three assessments reviewed were conducted for the same species (winter flounder), but across stocks inhabiting three different regions. As a consequence there were many similarities as well as differences in the assessments. The Terms of Reference were identical for each stock, and while many of the comments were similar across stocks, many were not. In the review presentations given below, each term of reference will be repeated and then addressed for each region's stock. This will result in some redundancy in the comments, but will facilitate the use of each region's stock assessment review independently of one other. Please also note that the Terms of Reference were those specified for the assessment scientists at the Stock Assessment Workshop. The Review Committee's responsibility here is to ascertain whether or not the Stock Assessment Workshop successfully completed these Terms of Reference.

A: Southern New England / Mid-Atlantic (SNE/MA) Winter Flounder

1. ***Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.***

- The methodology used to derive the catch information appeared appropriate, although the panel did not have time to conduct an extensive examination of these datasets. One might consider the uncertainty in the mortality associated with discards and how that might have affected total landings.
- There is a potential bias in the calculation of spawning stock biomass as the weights at age used are from the landing data and thus are conflated with size selectivity at age especially for the SNE/MA stock.
- Some of the variation in observed weights at age are probably related to variation in sampling rather than actual biological changes in the population. For example, in 1986-1987 for the SNE/MA assessment the weight at age 5 was lower than the cohort weight at age 6. One might consider modeling or smoothing the weights at age in some way to get more consistent estimates of this component of the population.
- One might consider how much and what component of the overall variation is represented in the PSE. For commercial catch, this typically represented error in the allocation process, whereas the PSE determined in the calculation of the discard information from the observer data reflected variation in sampling. If the PSEs were constructed in such a way to better reflect overall uncertainty their use in weighting the data going into the statistical catch at age model might be usefully employed.
- There are several other sources of uncertainty that could be included or accounted for (for example, errors in catch accounting) that are not discussed further here.

2. ***Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.***

- The survey data were generally well documented, although some consideration might be given to evaluating whether all surveys provide the same level of information. Some might be considered for elimination, while others might be down weighted relative to their information content and variability.
- Furthermore, consideration should be given as to how or whether surveys should be combined in the model? We appreciated that NEFSC scientists were trying to deal with this problem and that discussions will continue in the U.S. and elsewhere on this issue.
- Information on what area of the stock distribution the particular survey covers should also be provided.
- All indices used in the assessment should be tabulated and provided graphically as time series. For example, there should be standard tables of indices at age from the survey with row totals. These tables should be in a standardized format so that comparisons between tables are straightforward.
- One might consider providing a time series of average survey catch (kg/tow or #/tow) by strata to provide an indication of the changing distribution of the resource. Sometimes this was given by bubble plots on maps, but not always. Appendix C1, Table C3, Page 157 of the GOM WP1 document gives another example of this but only for one year and the strata were aggregated.
- Also a presentation of the standard errors associated with these indices would be helpful for evaluation of the relative uncertainty of each survey. These might eventually be considered for inclusion as weights directly in the assessment.
- Age aggregated indices might be plotted with associated confidence intervals as was done in the GOM document, although it would be good to identify if these are 95% confidence intervals, one standard error, or something else.
- Although there is probably no real consensus on this in the literature, some consideration should be given to the graphical format and type of data presented. If the assessments for the three stocks are considered again together it would be nice to have some consistency in how the data are presented between assessments to see similarities and differences in trends. Some prefer graphical displays and others might prefer tables and, of course, one might need to display the same data in a variety of ways to fully characterize the trends and uncertainty of the data. There were many characteristics of both the data and the model output that were difficult to track simply because the data were not presented in a clear and complete fashion. For reviewers and assessment scientists alike, it is nice to have standardized methods for viewing information, however some creative ways have to be considered to characterize information that is unique to a particular species or stock or stock assessment problem.

- Consideration should be given to presenting relative survey Zs (i.e. relative total mortalities) for each survey to examine trends in exploitation rates independent of the model itself.
- While the length-based net calibrations between vessels were informative and appeared appropriate, this method might be considered for peer review itself, either through publication or through a general peer review process. If this method continues to be used, then the inclusion of the uncertainty in the estimates into the assessment should be encouraged. We note that right now the application of this net re-calibration procedure effectively eliminates either end of the length distribution from Bigelow collections. We recognize that later, if the method is used to expand the Albatross to Bigelow lengths (as opposed to the other way around as is currently done), the expansion would likely create high variability in the tails of the distribution and would be quite unstable for practical use.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

- Estimates for each of the items listed in this term of reference were provided for in the SNE/MA assessment.
- The Review Committee interprets this term of reference to extend to an examination of the quality of the assessment models in general, as model quality is not really addressed in any other of the Terms of Reference provided here. In general, the statistical catch-age model used for SNE/MA assessment is considered to be a scientifically credible approach and provides a reasonable basis for fisheries management advice.
- It would be useful to have a summary table included in the assessment report describing the details of the model configuration for the preferred model being put forward. For these assessments, certain assumptions were made quite clear (e.g. the assumed value for natural mortality, M), while other assumptions were less clear (e.g. selectivity, was it flat topped or domed shape, was it fixed or allowed to vary over time, were there breaks in the series). Ultimately, one would like to have documented all the relevant information needed to replicate the assessment.
- The survey age-aggregated indices are declining faster than the preferred model would predict in the last decade for the SNE/MA stock, which may point to a time varying M. In this assessment, this was explored through various representations of fixed and time varying M. While, it makes sense that M should be higher than the 0.2 value used in previous assessments, it is not clear whether M should be set at a constant 0.3 level, or go from 0.3 to 0.6, or go from 0.3 to 0.6 back to 0.3, or whether the model might be improved by including instead some other time varying

component (catchability, catch reporting, selectivity, etc.). Closer examination of total mortality rates from the surveys should shed some light on this issue.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

- The sensitivity of the assessment results to choices in the allocation of catch appeared to have been thoroughly examined. The results did not seem to be overly sensitive to reasonable alternative allocation decisions.

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

- An analysis of the effects of temperature on departures from mean recruitment levels was provided in a working document by Jon Hare of the NEFSC (SDWG Background WP#13). The analysis was based on spawner-recruit data derived from previous (and not the current) assessment models for each of the three regions considered. The Review Committee notes that sometimes sea surface or coastal air indicators may not be the most indicative of what will eventually affect recruitment (although they are often the only indicators available). The analysis was conducted in a scientifically sound manner and did shed some light on the physical drivers for this species. While there appeared to be higher recruitment rates at lower temperatures, predicting how temperature might change and thus how future recruitments and associated biomasses might vary in projections is not straightforward. Furthermore, a consistent pattern in the temperature indicators was not evident. In the end, the effect of environmental factors was not included in either the assessments or the spawner-recruit models used for developing reference point estimates or making projections. The work conducted here appeared fruitful and efforts in this area should be continued.

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, and FMSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

- Previously adopted stock status definitions for "overfished" and "overfishing" were provided for all stocks. These were updated and redefined for the SNE/MA stock as a

new assessment model was adopted and new methods for determining MSY-based BRPs were employed.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

- The stock status values were examined by the SAW working group and reviewed at the SARC review meeting. The Review Committee is satisfied that the evaluations were performed correctly. As always, one should keep in mind the uncertainty associated with determining the reference points as well as the estimates of stock biomass and exploitation rates. The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.
- The Review Committee agreed with the SDWG that in 2010 the SNE/MA winter flounder stock was overfished but overfishing was not occurring.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.

a. ***Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).***
b. ***Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.***
c. ***Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.***

- It appeared that the SAW working group used standard methods that were reasonable for conducting projections. The SNE/MA assessment used an MCMC approach to generate its initial values for the projections. The written documentation on the SNE/MA projections was unclear about how the uncertainty associated with the recruitment estimates was dealt with.

- Uncertainty in M is not included in the projections.
- Information on stock vulnerability (as may be characterized through indices of productivity and/or susceptibility) was presented in the assessment document. The text in this section explores sensitivity analyses, residual plots and retrospectives in the context of the standard definition of vulnerability. However, it is unclear whether vulnerability issues that are not detectable through standard statistical diagnostics were explored. For example, life history issues such as longetivity of the species, fecundity and overall productivity, resilience to impact, or whether the species or stock is overly susceptible to fishing or environmental conditions (such as ocean warming) were discussed at various points during the review, but no evidence of this information appeared to be provided in the documentation.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

- The research recommendations provided by the SAW working group seemed interesting, constructive and reasonable to try to achieve. We recommend they go forward with these. The working group also reported on previous research recommendations (from GARM III and earlier) and they appear to have made good progress in addressing those that were reasonable to explore. The group should be applauded in this regard.
- In addition, we would like to suggest the following:
- Calibration of the macroscopic identification of maturity relative to the microscopic IDs might be considered through perhaps a ratio estimator to appropriately adjust the maturity at age/length.
- The Review Committee wondered if sexual dimorphism exists for this species? If so, does that create a problem for the assessments? For management?.
- Stock-size indices from the NEFSC Winter, Spring and Autumn surveys were revised to create a consistent set of strata for the whole time series. This means that the survey area was reduced to that which was consistently sampled. This may be ok, but if portions of the population shift into and out of the zones that are no longer sampled (a high possibility for this inshore species) this may adversely affect the indices. So, is this a problem? If so, how might it be best dealt with? We recommend that the number of sets in the excluded strata, the percent biomass in the excluded strata, the percent number of fish per survey strata should be tabulated to see what is missing and see if it is a potential source of bias in the index.
- Text on the changes that have occurred in management regulations seems to have been well documented in the assessment documents. What would be nice to have in addition would be a conceptual model outlining what affects these changes may have on the assessment so that the assessment can be examined with regard to

these hypothesized effects in a straightforward manner (rather than having to pour through all the text in the document to find them).

- Some clear idea of how to respond in a scientific and managerial way to changes in productivity would be useful. For example, if recruitment rates or growth are lower in the last decade, should reference points or projections be based on the last decade, the full time series, or some marriage of the two. Another example relevant to these stocks is the situation where M may have increased. Should reference points be based on the most recent M or projections of what will happen with M? We do not have a specific recommendation for how to do this, but this will clearly continue to be a problem for this and other species, so some plan should be developed for how to deal with this. Whatever approach is used, it should be justified and clearly documented.
- The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.

B: Georges Bank (GBK) Winter Flounder

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.

- The methodology used to derive the catch information appeared appropriate, although the panel did not have time to conduct an extensive examination of these datasets. One might consider the uncertainty in the mortality associated with discards and how that might have affected total landings.
- There is a potential bias in the calculation of spawning stock biomass as the weights at age used are from the landing data and thus are conflated with size selectivity at age.
- Some of the variation in observed weights at age are probably related to variation in sampling rather than actual biological changes in the population. One might consider modeling or smoothing the weights at age in some way to get more consistent estimates of this component of the population.
- One might consider how much and what component of the overall variation is represented in the PSE. For commercial catch, this typically represented error in the allocation process, whereas the PSE determined in the calculation of the discard information from the observer data reflected variation in sampling. If the PSEs were constructed in such a way to better reflect overall uncertainty their use in weighting the data going into the statistical catch-age model might be usefully employed.
- It was noted that prior to 2002 all fish below the minimum landing size were assumed to be discarded. It is not clear whether this is appropriate or not, but one might reexamine that assumption to see if it is valid and consistent with current practices for handling the data.

- The issue of identification of maturity at age at sea came up during the meeting and although the at sea methods seemed challenging but appropriate, post-landing association to appropriate maturity categories might be conducted using some kind of ratio estimator based on laboratory analyses.

2. ***Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.***

- The survey data were generally well documented, although some consideration might be given to evaluating whether all surveys provide the same level of information. The Canadian survey might be down weighted relative to their information content and variability.
- Information on what area of the stock distribution the particular survey covers should also be provided.
- All indices used in the assessment should be tabulated and provided graphically as time series. For example, there should be standard tables of indices at age from the survey with row totals. These tables should be in a standardized format so that comparisons between tables are straightforward.
- One might consider providing a time series of average survey catch (kg/tow or #/tow) by strata to provide an indication of the changing distribution of the resource. Sometimes this was given by bubble plots on maps, but not always. Appendix C1, Table C3, Page 157 of the GOM WP1 document gives another example of this but only for one year and the strata were aggregated.
- Also an assessment of the standard errors associated with these indices would be helpful for evaluation of the relative uncertainty of each survey. These might also be considered for inclusion as weights directly in the assessment.
- Age aggregated indices might be plotted with associated confidence intervals as was done in the GOM document, although it would be good to identify if these are 95% confidence intervals, one standard error, or something else.
- Although there is probably no real consensus on this in the literature, some consideration should be given to the graphical format and type of data presented. If the assessments for the three stocks are considered again together it would be nice to have some consistency in how the data are presented between assessments to see similarities and differences in trends. Some prefer graphical displays and others might prefer tables and, of course, one might need to display the same data in a variety of ways to fully characterize the trends and uncertainty of the data. There were many characteristics of both the data and the model output that were difficult to track simply because the data were not presented in a clear and complete fashion. For reviewers and assessment scientists alike, it is nice to have standardized methods for viewing information, however some creative ways have to be

considered to present information that is unique to a particular species or stock or stock assessment problem.

- Consideration should be given to presenting relative survey Zs (i.e. relative total mortalities) for each survey to examine trends in exploitation rates independent of the model itself.
- While the length-based calibrations between vessels were informative and appeared appropriate, this method might be considered for peer review itself, either through publication or through a general peer review process. If this method continues to be used, then the inclusion of the uncertainty in the estimates into the assessment should be encouraged. We note that right now the application of this net re-calibration procedure effectively eliminates either end of the length distribution from Bigelow collections. We recognize that later, if the method is used to expand the Albatross to Bigelow lengths (as opposed to the other way around as is currently done), the expansion would likely create high variability in the tails of the distribution and would be quite unstable for practical use.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

- Estimates for each of the items listed in this term of reference were provided for in the GBK assessment.
- The reviewers take this term of reference to extend to an examination of the quality of the assessment model in general, as model quality is not really addressed in any other of the other Terms of Reference provided here. In general, the VPA model used for GBK assessment was a scientifically credible approach and provides a reasonable basis for fisheries management advice. A statistical catch-age model should be considered for the GBK stock as there may be more uncertainty here associated with catch and discards than would be appropriate for the assumption of true known catches as is made in a VPA type of analysis. The challenge in using a statistical catch-age model over a VPA for this stock is in tracking the changes in selectivity that likely have occurred as a result of changes in management and in the fishery. While a statistical catch-age model should be explored, this review found no issues that would necessarily indicate that one would get different results than were found in the VPA analysis.
- It would be useful to have a summary table included in the assessment report describing the details of the model configuration for the preferred model being put forward. Ultimately, one would like to have documented all the relevant information needed to replicate the assessment.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

- The sensitivity of the assessment results to choices in the allocation of catch appeared to be thoroughly examined. The results did not seem to be overly sensitive to reasonable alternative allocation decisions.

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

- An analysis of the effects of temperature on departures from mean recruitment levels was provided in a working document by Jon Hare of the NEFSC (SDWG Background WP#13). The analysis was based on spawner-recruit data derived from previous (and not the current) assessment models for each of the three regions considered. The Review Committee notes that sometimes sea surface or coastal air indicators may not be the most indicative of what will eventually affect recruitment (although they are often the only indicators available). The analysis was conducted in a scientifically sound manner and did shed some light on the physical drivers for this species. While there appeared to be higher recruitment rates at lower temperatures, predicting how temperature might change and thus how future recruitments and associated biomasses might vary in projections is not straightforward. Furthermore, a consistent pattern in the temperature indicators was not evident. In the end, the effect of environmental factors was not included in either the assessments or the spawner-recruit models used for developing reference point estimates or making projections. The work conducted here appeared fruitful and efforts in this area should be continued.

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, and FMSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

- Previously adopted stock status definitions for "overfished" and "overfishing" were provided for all stocks. These were updated and redefined for the GBK stock as an updated assessment model was adopted and new methods for determining MSY-based BRPs were employed.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

- The stock status values were examined by the SAW working group and reviewed at the SARC review meeting. The Review Committee is satisfied that the evaluations were performed correctly. As always, one should keep in mind the uncertainty associated with determining the reference points as well as the estimates of stock biomass and exploitation rates. The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.
- The Review Committee agreed with the SDWG that in 2010 the GBK winter flounder stock was not overfished and overfishing was not occurring.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.

a. ***Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).***
b. ***Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.***
c. ***Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.***

- It appeared that the SAW working group used standard methods that were reasonable for conducting projections. The GBK assessment used a bootstrap approach to incorporate the assessment uncertainty into the initial values the projections.
- Uncertainty in M is not included in the projections.
- Information on stock vulnerability (as may be characterized through indices of productivity and/or susceptibility) was presented in the assessment document. The text in this section explores sensitivity analyses, residual plots and retrospectives in the context of the standard definition of vulnerability. However, it is unclear whether vulnerability issues that are not detectable through standard statistical diagnostics were explored. For example, life history

issues such as longetivity of the species, fecundity and overall productivity, resilience to impact, or whether the species or stock is overly susceptible to fishing or environmental conditions (such as ocean warming) were discussed at various points during the review, but no evidence of this information appeared to be provided in the documentation.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

- The research recommendations provided by the SAW working group seemed interesting, constructive and reasonable to try to achieve. We recommend they go forward with these. The working group also reported on previous research recommendations (from GARM III and earlier) and they appear to have made good progress in addressing those that were reasonable to achieve. The group should be applauded in this regard.
- In addition, we would like to suggest the following:
- Calibration of the macroscopic identification of maturity relative to the microscopic IDs might be considered through perhaps a ratio estimator to appropriately adjust the maturity at age/length.
- The Review Committee wondered if sexual dimorphism exists for this species? If so, does that create a problem for the assessments? For management?
- Georges Bank is a unique area, and a retention index showing the influence on larval drift might be a valuable metric to have.
- Text on the changes that have occurred in management regulations seems to have been well documented in the assessment documents. What would be nice to have in addition would be a conceptual model outlining what affects these changes may have on the assessment so that the assessment can be examined with regard to these hypothesized effects in a straightforward manner (rather than having to pour through all the text in the document to find them).
- Some clear idea of how to respond in a scientific and managerial way to changes in productivity would be useful. For example, if recruitment rates or growth are lower in the last decade, should reference points or projections be based on the last decade, the full time series, or some marriage of the two. Another example relevant to these stocks is the situation where M may have increased. Should reference points be based on the most recent M or projections of what will happen with M? We do not have a specific recommendation for how to do this, but this will clearly continue to be a problem for this and other species, so some plan should be developed for how to deal with this. Whatever approach is used, it should be justified and clearly documented.
- The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.

B: Gulf of Maine (GOM) Winter Flounder

1. ***Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.***

- The methodology used to derive the catch information appeared appropriate, although the panel did not have time to conduct an extensive examination of these datasets. However, a more thorough examination of recreational catch data in the early part of the time series is needed. The estimated catch and CPUE for 1982 appeared to be high and variable.
- As with the other stocks, one might consider the uncertainty in the mortality associated with discards and how that might have affected total landings.
- When a statistical catch-age model is eventually developed, some consideration should be given to appropriately estimating weight at age in the spawning population. Using the weight at age from the landings may not be appropriate. Also, some of the variation in observed weights at age are probably related to variation in sampling rather than actual biological changes in the population. One might consider modeling or smoothing the weights at age to get more consistent estimates of this component of the population.
- When using a statistical model, one might consider how much and what component of the overall variation is represented in the PSE. For commercial catch, this typically represented error in the allocation process, whereas the PSE determined in the calculation of the discard information from the observer data reflected variation in sampling. If the PSEs were constructed in such a way to better reflect overall uncertainty their use in weighting the data going into the statistical catch at age model might be usefully employed.
- There are several other sources of uncertainty that could be included or accounted for (for example, errors in catch accounting) that are not discussed further here.

2. ***Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.***

- The survey data were generally well documented, although some consideration might be given to evaluating whether all surveys provide the same level of information. Some might be considered for elimination, while others might be down weighted relative to their information content and variability.
- Furthermore, consideration should continue to be given as to how surveys should be combined in the assessment? We appreciated that NEFSC scientists started to address this problem and that discussions will continue in the U.S. and elsewhere on this issue.

- Information on what area of the stock distribution the particular survey covers should also be provided.
- All indices used in the assessment should be tabulated and provided graphically as time series. For example, there should be standard tables of indices at age from the survey with row totals. These tables should be in a standardized format so that comparisons between tables are straightforward.
- One might consider providing a time series of average survey catch (kg/tow or #/tow) by strata to provide an indication of the changing distribution of the resource. Sometimes this was given by bubble plots on maps, but not always. Appendix C1, Table C3, Page 157 of the GOM WP1 document gives another example of this but only for one year and the strata were aggregated.
- Also an assessment of the standard errors associated with these indices would be helpful for evaluation of the relative uncertainty of each survey. These might also be considered for inclusion as weights directly in the assessment.
- Age aggregated indices have been plotted with associated confidence intervals . in the GOM document; it would be good to identify if these are 95% confidence intervals, one standard error, or something else.
- Although there is probably no real consensus on this in the literature, some consideration should be given to the graphical format and type of data presented. If the assessments for the three stocks are considered again together it would be nice to have some consistency in how the data are presented between assessments to see similarities and differences in trends. Some prefer graphical displays and others might prefer tables and, of course, one might need to display the same data in a variety of ways to fully characterize the trends and uncertainty of the data. There were many characteristics of both the data and the model output that were difficult to track simply because the data were not presented in a clear and complete fashion. For reviewers, it is nice to have standardized methods for presenting information, however some creative ways have to be considered to present information that is unique to a particular species or stock or stock assessment problem.
- Consideration should be given to presenting relative survey Zs (i.e. relative total mortalities) for each survey to examine trends in exploitation rates independent of the model itself.
- While the length-based calibrations between vessels were informative and appeared appropriate, this method might be considered for peer review itself, either through publication or through a general peer review process. If this method continues to be used, then the inclusion of the uncertainty in the estimates into the assessment should be encouraged. We note that right now the application of thisnet re-calibration procedure effectively eliminates either end of the length distribution from the Bigelow collections. We recognize that later, if the method is used to expand the Albatross to Bigelow lengths (as opposed to the other way around as is currently done), the expansion would likely create high variability in the tails of the distribution and would be quite unstable for practical use.

- Tables of indices at age (and totals) were unavailable for the GOM assessment.

3. ***Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.***

- Because the GOM model was not accepted for use, only estimates of stock biomass and exploitation rates in 2009 and 2010 were available.
- The reviewers take this term of reference to extend to an examination of the quality of the assessment model in general, as model quality is not really addressed in any of the other Terms of Reference provided here. The GOM statistical catch-age model proved to be inadequate to account for conflicting trends in the catch and survey information, however the fall back analysis of the area swept method provides a reasonable gauge of overfishing status and provides time trends in biomass to assess the direction the stock is headed.
- For future assessments, where a model is used, it would be useful to have a summary table included in the assessment report describing the details of the model configuration for the preferred model being put forward. For the other winter flounder assessments, certain assumptions were made clear (e.g. assumed value for M), while other assumptions were less clear (e.g. selectivity, was it flat topped or domed shape, was it fixed or allowed to vary, were their breaks in the series). Ultimately, one would like to have documented all the relevant information needed to replicate the assessment.
- Several statistical catch-age models were explored for this stock. Because of the conflict in the information provided by the survey relative to the age information in the catch and the scale of the total catch none of the models adequately accounted for trends in both indices. The statistical catch-age model looked promising as a way of synthesizing all the information available and if the application were found to be appropriate would provide useful biological reference points, but work is still needed in this area.
- The GOM assessment is now based on a swept-area method. Because this was a fallback assessment, the Review Committee did not get to explore this approach as fully. However, it is a simple and straightforward method that can be used. We simply note that the survey catchability used in this analysis is more of an assumption than a rigorously estimated parameter. Nevertheless, the F reference point, and the finding that overfishing is not taking place, is robust to reasonable choices of the survey catchability q.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

- The sensitivity of the assessment results to choices in the allocation of catch appeared to be thoroughly examined. The results did not seem to be overly sensitive to reasonable alternative allocation decisions.

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

- An analysis of the effects of temperature on departures from mean recruitment levels was provided in a working document by Jon Hare of the NEFSC (SDWG Background WP#13). The analysis was based on spawner-recruit data derived from previous (and not the current) assessment models for each of the three regions considered. The Review Committee notes that sometimes sea surface or coastal air indicators may not be the most indicative of what will eventually affect recruitment (although they are often the only indicators available). The analysis was conducted in a scientifically sound manner and did shed some light on the physical drivers for this species. While there appeared to be higher recruitment rates at lower temperatures, predicting how temperature might change and thus how future recruitments and associated biomasses might vary in projections is not straightforward. Furthermore, a consistent pattern in the temperature indicators was not evident. In the end, the effect of environmental factors was not included in either the assessments or the spawner-recruit models used for developing reference point estimates or making projections. The work conducted here appeared fruitful and efforts in this area should be continued.

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, and FMSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

- Previously adopted stock status definitions for "overfished" and "overfishing" were provided for all stocks. The GOM assessment, using a swept-area method, was able to provide a proxy estimate of the "overfishing" level, but could not provide an estimate of "overfished" status.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

- The stock status values were examined by the SAW working group and reviewed at the SARC review meeting. The Review Committee is satisfied that the evaluations were performed correctly. As always, one should keep in mind the uncertainty associated with determining the reference points as well as the estimates of stock biomass and exploitation rates. The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.
- The Review Committee agreed with the SDWG that, based on the swept-area method, overfishing does not appear to be taking place in 2010 the GOM winter flounder stock, however it is not possible under the current configuration to tell whether or not the stock is overfished.
- Depending on what assessment methods are available for future assessments, a length-based yield-per-recruit analysis should be considered to help determine biological reference points. If this is taken on, then updated information on growth should be used. The Review Committee notes that this approach was, in fact, developed and implemented at end of the meeting by the SDWG and the results are now included in the revised summary for GOM. And, as requested by the Review Committee, the updated growth parameters were used.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.

a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).

b. Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.

c. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

- The GOM assessment provided no projections.

- Consider the comments for conducting projections provided for the other two stocks when the time comes to create projections from an acceptable assessment model.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

- The research recommendations provided by the SAW working group seemed interesting, constructive and reasonable to try to achieve. We recommend they go forward with these. The working group also reported on previous research recommendations (from the GARM III and earlier) and they appear to have made good progress in addressing those that were reasonable to achieve. The group should be applauded in this regard.
- In addition, we would like to suggest the following:
- Calibration of the macroscopic identification of maturity relative to the microscopic IDs might be considered through perhaps a ratio estimator to appropriately adjust the maturity at age/length.
- The Review Committee wondered if sexual dimorphism exists for this species? If so, does that create a problem for the assessments? For management?
- Stock-size indices from the NEFSC Winter, Spring and Autumn surveys were revised to create a consistent set of strata for the whole time series. This means that the survey area was reduced to that which was consistently sampled. This may be ok, but if portions of the population shift into and out of the zones that are no longer sampled (a high possibility for this inshore species) this may adversely affect the indices. So, is this a problem? If so, how might it be best dealt with? We recommend that the number of sets in the excluded strata, the percent biomass in the excluded strata, the percent number of fish per survey strata should be tabulated to see what is missing and see if it is a potential source of bias in the index.
- Text on the changes that have occurred in management regulations seems to have been well documented in the assessment documents. What would be nice to have in addition would be a conceptual model outlining what affects these changes may have on the assessment so that the assessment can be examined with regard to these hypothesized effects in a straightforward manner (rather than having to pour through all the text in the document to find them).
- Some clear idea of how to respond in a scientific and managerial way to changes in productivity would be useful. For example, if recruitment rates or growth are lower in the last decade, should reference points or projections be based on the last decade, the full time series, or some marriage of the two. Another example relevant to these stocks is the situation where M may have increased. Should reference points be based on the most recent M or projections of what will happen with M? We do not have a specific recommendation for how to do this, but this will clearly

continue to be a problem for this and other species, so some plan should be developed for how to deal with this. Whatever approach is used, it should be justified and clearly documented.

- The Review Committee recommends that future Terms of Reference might include an evaluation of the probability of being overfished or overfishing taking place rather than simply using a point estimate based on the model.

Appendix 1.

The Review Committee and NEFSC scientists developed a method at the meeting for combining information on winter flounder across regions to help inform the spawner-recruit relationships used in developing projections and Biological Reference Points. The method is described below and uses likelihood-based AIC methods to find a reasonable compromise between a spawner-recruit relationship based on combined data sources and the individual spawner-recruit estimates associated with the individual stocks of winter flounder. This method maximizes the fit to both the SNE/MA and GBK datasets while minimizing the differences between relationships in the adjoining regions.

FMSY, SSBMSY, and MSY were estimated using a spawner-recruit model applied over a range of values for steepness (defined as the slope of the stock recruitment curve near the origin). It was assumed, based on the biology of the species, that steepness should be similar between the different stocks. These stocks are neighbouring populations of the same species that share common reproductive strategies. Fecundities at size are similar, although larval survivorship and recruitment to the fishery may vary between areas. Because the data available for any one stock may not be sufficient to fully parameterize a spawner-recruit relationship, some method of bringing additional information to bear on the estimates would be useful. Initially estimates of steepness from the work of Myers et al. (1999) were used as a prior for estimating the spawner-recruit relationship, but because the Myers et al. data include only more distantly related Pleuronectids than those present in these assessments it was felt that some way of using information available in the adjacent stocks would be more appropriate.

The objective was to find values of steepness chosen to be as similar as possible between stocks within the constraints of the information content available within each stock. A strategy was outlined that allowed the steepness parameters to be chosen among a range of values that provided reasonable fits to the spawner-recruit data for each individual stock, but were also reasonably close in the parameter space to each other. A profile of ΔAICs (difference between AIC and the minimum value found) over a range of steepness values was developed for each of the two available stocks. The profiles are provided in Figure 1 below. It was considered that values of steepness associated with the AIC values that are within 2 units of the minimum AIC for each stock (i.e. ΔAIC <=2) would be within a range of realistic values (Burnham and Anderson, 2002).

Once the profiles were generated, the steepness value for each stock was selected so that the difference in steepness between the SNE/MA and GBK stocks was as small as possible within the constraint that the choices were not outside the ΔAIC = 2 bound for the given (original) stock's minimum fit. For the SNE stock this means steepness was set at the largest value possible within ΔAIC = 2 of its minimum fit (steepness = 0.61). For the GBK stock this means steepness was set at the smallest value such that ΔAIC = 2 of its minimum fit (steepness=0.78). Thus, the model estimates were shrunk towards each other, making steepness as similar as possible without losing the stock specific characteristics of the recruitment process.

The BRP estimates derived for the winter flounder stocks based on the spawner-recruit relationship specified in this way are direct MSY-based estimates and we believe are the most appropriate for use in informing management decisions at this time.

Burnham, K. P., and Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag.

Myers, R. A., Bowen, K. G., Barrowman, N. J. 1999. Maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56: 2404-2419.

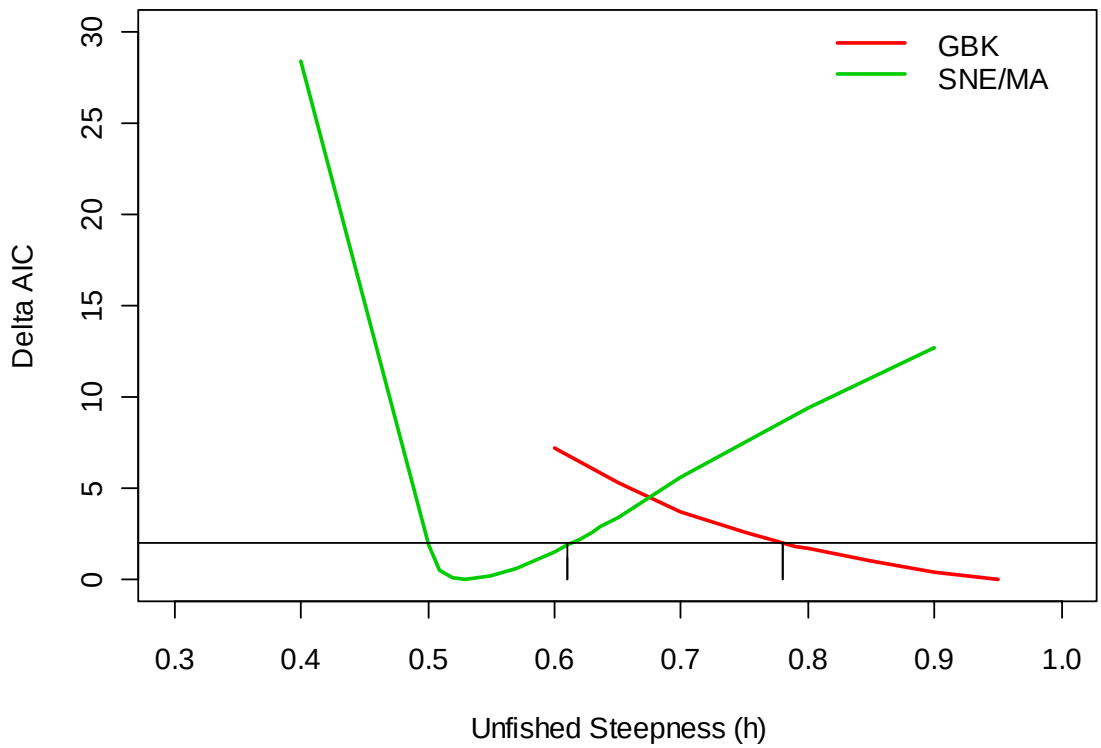

Figure 1. Delta AIC (AIC – min(AIC)) for each region's fitted spawner-recruit relationship over a range of fixed steepness parameter values. The curves correspond to the ΔAIC values from the fits for the two regions (Georges Bank and Southern New England / Mid-Atlantic). The black horizontal line corresponds to the ΔAIC threshold of 2. Steepness values corresponding to a ΔAIC below 2 are not considered statistically different from one another with a region. The vertical black lines show the locations of the most similar steepness parameters that are still within the range of best estimates for each model. The steepness values corresponding to this criterion are 0.61 for SNE/MA and 0.78 for GBK.

SARC 52 - Bibliography – List of References

Working Papers for Review

SDWG. 2011. Southern New England/Mid-Atlantic (SNE/MA) Winter Flounder: Stock Assessment for 2011

SDWG. 2011. Southern New England / Mid-Atlantic (SNE/MA) Winter Flounder Assessment Summary

SDWG. 2011. Georges Bank Winter Flounder Stock Assessment for 2011

SDWG. 2011. Georges Bank Winter Flounder Assessment Summary For 2011

SDWG. 2011. Gulf of Maine (GOM) Winter Flounder Stock Assessment for 2011

SDWG. 2011. Gulf of Maine (GOM) Winter Flounder Assessment Summary for 2011

Background Papers

Southern Demersal Working Group Papers

1. DeCelles, G, Cadrin, SX. ICES CM 2007/L:18, An Interdisciplinary Assessment of Winter Flounder (Pseudopleuronectes americanus) Stock Structure

2. Terceiro, M. Impacts of reduced inshore strata sampling on NEFSC trawl survey indices for SNE/MA winter flounder

3. Southern Demersal Working Group. 2011. Maturity

4. Southern Demersal Working Group. 2011. Response to 2008 GARM3 Research Recommendations for winter flounder

5. Southern Demersal Working Group. 2011. SNE/MA Winter Flounder TOR 4

6. Southern Demersal Working Group. 2011. Management Regulations

7.

   7a. Miller, T. 2011. Winter Flounder Length-based Survey Calibration

7b. Southern Demersal Working Group. 2011. Winter Flounder Calibration: WP 7

8. McBride, R, Wuenschel, M, McElroy, D, Rowinski, Y, Thornton, G., Nitschke, P. 2011. Classifying female winter flounder maturity during NEFSC resource surveys: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method. A Working Paper for SARC 52

9.

   9a. SDWG52 WP 9: Validating the stock apportionment of commercial fisheries landings using positional data from Vessel Monitoring Systems (VMS): Impacts on the winter flounder stock allocations. (Update of Palmer, MC Wigley, SE 2007. Validating the stock apportionment of commercial fisheries landings using positional data from Positional Monitoring Systems (VMS). US Dept Commer, Northeast Fisheries Sci Cent Ref Doc. 07-22.

   9b. Palmer MC, Wigley SE Using positional data from vessel monitoring systems (VMS) to validate the logbook-reported area fished and the stock allocation of commercial fisheries landings, 2004-2008. (Update of Northeast Fisheries Science Center Reference Document 07-22)

10. DeCelles, G, Roman S, Martins, D, Wood, A, Cadrin S. Results of an Industry-Based Survey for Winter Flounder in the Great South Channel. (Presentation in December 2010 Distribution and Abundance of Winter Flounder in the Great South Channel included in NEFSC CRD 1021.)

11. Wigley SE, Palmer, M., Legault, C. A 2011. Comparison of Discard Rates Derived from At-Sea Monitoring and Observer Trips. A Working Paper in support of SARC 52 Winter Flounder TOR 1:"Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data."

12. McElroy, WD, Rowinski, YK, Towle, McBride RS, Wuenschel MJ. Reproductive potential of female winter flounder, *Pseudopleuronectes americanus*: Comparison of fecundity and skipped spawning among three stocks

13. Hare, J. ToR 5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function). Development of environmentally-explicit stock-recruitment models for three stocks of winter flounder (*Pseudopleuronectes americanus*) along the northeast coast of the United States

14. Wigley, SE, Blaylock, J, Palmer, M. 2011. Measures of Uncertainty in the Trip-based Allocated Landings. A Working Paper in support of SARC 52 Winter Flounder TOR 4 "Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR 5)."

15. Southern Demersal Working Group. Anthony Woods. 2011. Winter flounder natural mortality derived from data in Howe and Coates (1975) using instantaneous rates tagging models. Working Paper: Re-analysis of Howe and Coates (1975)

16. Southern Demersal Working Group. 2011. Consensus Statement on Biological Reference Points (Term of Reference 6) and Vulnerability (Term of Reference 8b) for Winter Flounder Stocks (also called "16_D")

Garm III Background Papers

1. GARM III. Summary NEFSC 2008

2. Terceiro, M. GARM III Report. NEFSC 2008. J. Southern New England/Mid-Atlantic winter flounder

3. Hendrickson, L. GARM III Report. NEFSC 2008. K. George's Bank winter flounder

4. Nitschke, P. GARM III Report. NEFSC 2008. I. Gulf of Maine winter flounder

5. O'Boyle R, Bell, Crecco V, Van-Eeckhaute L, Kahn D, Needle C, Rothschild, B, Smith S, Helge Volstad, J . GARM III Report. NEFSC 2008. Report of the Groundfish Review Assessment Meeting (GARM III). Part I. Data Methods.

6. O'Boyle R., De Oliveira J, Gavaris S, Ianelli J, Jiao Y, Jones C, Medley P. GARM III Report. NEFSC 2008. Panel Summary Report of the Groundfish Review Assessment Meeting (GARM III). Part II Assessment Methodology (Models)

7. O'Boyle R, Bell M, Gavaris S, Haist V, Reeves S, Thompson G.GARM III Report. NEFSC 2008. Panel Summary Report of the Groundfish Review Assessment Meeting (GARM III). Part 3. Biological Reference Points.

Misc. Background Papers

1. Nies TA. 2011. Use of Reference Points in the Northeast Multispecies Fishery Management Plan (Applicable to New England Groundfish, Including Winter Flounder)

2. Nies TA. 2011. Management of the Northeast Multispecies Fishery 1977-2010

3. Rademeyer RA, Butterworth DS. April 2010. Initial Applications of Statistical Catch-at-Age Assessment Methodology to the Gulf of Maine Winter Flounder Resource
4. Rademeyer RA, Butterworth DS. April 2010. Initial Applications of Statistical Catch-at-Age Assessment Methodology to the Southern New England/Mid-Atlantic Winter Flounder Resource
5. Rademeyer R, Butterworth D. SNE Winter Flounder: Application of SCAA. Marine Resource Assessment and Management Group, University of Cape Town

6. Rademeyer RA, Butterworth DS. April 2010. Update of the Southern New England/Mid-Atlantic Winter Flounder Resource New Base Case SCAA using updated data

7. Rothschild Brian J. and Jiao Yue, May 2011 'Characterizing Uncertainty in Fish Stock Assessments: the Case of the Southern New England-Mid-Atlantic Winter Flounder', Transactions of the American Fisheries Society, 140: 3, 557 — 569, First published on: 18 May 2011 (iFirst)

8. Brooks, E. N., Powers, J. E., and Corte´s, E. 2010. Analytical reference points for age-structured models: application to data-poor fisheries. – ICES Journal of Marine Science, 67: 165–175.

9. Brooks, E. N., and Powers, J. E. 2007. Generalized compensation in stock-recruit functions: properties and implications for management. – ICES Journal of Marine Science, 64: 413–424.

10. Gedamke T*1 Hoenig JM, Dupaul WD, and Musick JA. Stock–Recruitment Dynamics and the Maximum Population Growth Rate of the Barndoor Skate on Georges Bank. North American Journal of Fisheries Management. 29:512–526, 2009

11. Iles, TC. A Review of Stock-Recruitment Relationships with Reference to Flatfish Populations. Netherlands Journal of Sea Research 32 (3/4): 399-420 (1994)

12. NEFSC. NEFSC Stat Area

13. NEFSC. NEFSC Strata

14. Gislason H, Daan N, Rice JC, Pope JG. Size, Growth, Temperature and the natural mortality of marine fish. Fish and Fisheries. 2010 11, Issue 2, PP 149-158

15. Pereira JJ, Goldberg R, Zizkowski JJ, Berrien PL, Morse WW, Johnson DL. Winter Flounder, Pseudopleuronectes americanus, Life History and Habitat Characteristics. September 1999. NOAA Technical Memorandum NMFS-NE-138. 48 pp.

16. Lux FE, Peterson AE Jr, Hutton RF. 1970. Geographical Variation n Fin Ray Number in Winter Flounder, Pseudopleuronectes americanus (Walbaum), Off Massachusetts. Trans Amer Fish Soc 1970, No. 3

17. Howe AB, Coates PG. 1975. Winter Flounder Movements, Growth, and Mortality off Massachusetts. Trans. Amer. Fish. Soc., 1975, No 1

18. Lux FE. 1973. Age and Growth of the Winter Flounder, Pseudopleuro Americanus, on Georges Bank. Fishery Bulletin: Vol. 71, No. 2

19. Berry RJ, Saila SB, Horton DB. 1965. Growth Studies of Winter Flounder, Pseudopleuronectes americanus (Walbaum), in Rhode Island. Transactions of the American Fisheries Society. Volume 94, Issue 3, PP 259-264.

20. Saila, SB. 1961. A Study of Winter Flounder Movements. Limnol. Oceanogr. 6:292-298

21. Witherell DB, Burnett J. 1993. Growth and Maturation of winter flounder, Pleuronectes americanus, in Massachusetts. Fishery Bulletin 91 (4), 1993

22. Pierce DE, Howe AB. 1977. A Further Study on Winter Flounder Group Identification Off Massachusetts. Trans. Am. Fish. Soc., Vol. 106, No. 2, 1977.

23. Perlmutter A. 1947. The Blackback Flounder and Its Fishery in New England and New York. Bull. Bingham Oceanogr. Collect., Yale Univ. 11(2): 1-92.

24. NEFSC. 2003. 36th Northeast Regional Stock Assessment Workshop (36th SAW) Stock Assessment Review Committee (SARC) Consensus Summary of Assessments.

25. NEFSC. Mayo RK and Terceiro M, Eds. 2005. A report of the Stock Assessment Workshop (SAW) Northern and Southern Demersal Working Groups.

Assessment of 19 Northeast Groundfish Stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM). NEFSC Reference Document 05-13.

26. NEFSC. 1996. A Report of the 21st Northeast Regional Stock Assessment Workshop. 21st Northeast Regional Stock Assessment Workshop (21st SAW) Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Reference Document 96-05d.

27. Wigley SE, Palmer MC, Blaylock J, Rago PJ, 2008. A Brief Description of the Discard Estimation for the National Bycatch Report. NEFSC Reference Document 08-02.

28. Ampela, K. 2009. The Diet and Foraging Ecology of Gray Seals (Halichoerus Grypus) in United States Waters. A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

29. Hoenig JM. 1983. Empirical Use of Longevity Data to Estimate Mortality Rates. Fishery Bulletin: Vol. 82, No. 1, 1983

30. Hewitt DA, Hoenig JM. 2005. Comparison of two approaches for estimating natural mortality based on longevity. Fishery Bulletin 103(2)

31. DFO. 2011. Impacts of Grey Seals on Fish Populations in Eastern Canada. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2010/071

32. Ryer CH. 2008. A review of flatfish behavior relative to trawls. Fisheries Research 90(2008) 138-146.

33. Ryer, CH, Rose CS, Iseri PJ. 2009. Flatfish herding behavior in response to trawl sweeps: a comparison of diel responses to conventional sweeps and elevated sweeps. Fishery Bulletin 108(2).

34. Trzcinski MK, Mohn R, Bowen WD. 2006. Continued Decline of an Atlantic Cod Population: How Important is Gray Seal Predation? Ecological Applications, 16 (6), (Dec., 2006), pp. 2276-2292.

35. Brodziak J, Cadrin SX, Legault CM, Murawski SA. 2009. Goals and Strategies for rebuilding New England groundfish stocks. Fisheries Research 94 (2008) 355-366.

36. ASMFC Winter Flounder Technical Committee. SAW 28 Report. 1998. Assessment of the Southern New England/Mid-Atlantic and Gulf of Maine Winter Flounder stocks.

37. DeCelles G, Cadrin SX. Stock Structure of winter flounder in U.S. waters. SARC 52 Southern Demersal Working Group Paper #1.

38. McElroy WD, Rowinski YK, Towle EK, Wuenschel MJ, McBride R. 2011. Reproductive potential of female winter flounder, Pseudopleuronectes americanus: comparison of fecundity and skipped spawning among three stocks. Abstract. 30th Annual meeting of the Florida Chapter of the American Fisheries Society.

39. NEFSC. Overview of winter flounder ageing.

40. Wigley SE, Blaylock J, Palmer M, NEFSC. 2011. Measures of Uncertainty in Trip-based Allocated Landing. Working Paper 14 SARC 52 Winter Flounder.

41. Zhang Y, Cadrin SX. Analysis of Winter Flounder Survey Indices for stock assessment –SNE. Department of Fisheries Oceanography School for Marine Science & Technology University of Massachusetts, Dartmouth.

42. NOAA- NMFS. 2009. US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2008. NOAA Technical Memorandum NMFS-NE-210.

43. Dickie LM, McCracken FD. 1955. Isopleth Diagrams to Predict Equilibrium Yields of a Small Flounder Fishery. J. Fish. Res. BD. Canada, 12(22), 1955.

44. Saila SB, Horton DB, Berry RJ. 1965. Estimates of the Theoretical Biomass of Juvenile Winter Flounder, Pseudopleuronectes americanus (Walbaum) Required for a Fishery in Rhode Island. J. Fish. Res. BD. Canada., 22(4).

45. Magnuson-Stevens Act Text. Available from: **http://www.nmfs.noaa.gov/msa2005/docs/MSA_amended_msa%20_20070112_FINAL.pdf**

46. NEFSC. 2002. Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. NEFSC Reference Document 02-04

47. Jeffries HP, Terceiro M. 1985. Cycles of changing abundances in the fishes of Narragansett Bay area. Mar. Ecol. Prog. Ser 25: 239-244, 1985

48. M_Equations_WFL

49. WFL_Predation

50. Nye JA, Link JS, Hare JA, Overholtz WJ. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Mar Ecol Prog Ser 393: 111-129, 2009.

51. Sissenwine MP, Shepherd JG. 1987. An Alternative Perspective on Recruitment Overfishing and Biological Reference Points. Can. J, Fish. Aquat. Sci., Vol 44: 913-918.

52. Clark W, 1991. Groundfish Exploitation Rates Based on Life History Parameters. Can. J. Fish. Aquat. Sci. 48: 734-750.

53. Shepherd JG. 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. J. Cons. Int. Explor. Mer, 40(1): 67-75.

54. Clark WG. 1993. The Effect of Recruitment Variability on the Choice of a Target Level of Spawning Biomass Per Recruit. Management of Exploited Fish. Alaska Sea Grant 1993.

Statement of Work

External Independent Peer Review by the Center for Independent Experts

52st Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Winter flounder (Southern New England Stock), Winter flounder (Georges Bank Stock), Winter flounder (Gulf of Maine Stock).

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 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: The purpose of this meeting will be to provide an external peer review of stock assessments for three stocks of winter flounder (*Pseudopleuronectes americanus*): Southern New England, Georges Bank, and Gulf of Maine. Winter flounder, also known as blackback or lemon sole, is a demersal flatfish distributed in the Northwest Atlantic from Labrador to Georgia. U.S. commercial and recreational fisheries exist from the Gulf of Maine to the Mid-Atlantic Bight. Winter flounder stocks are managed in federal waters under the New England Fishery Management Council's Northeast Multispecies Fishery Management Plan (FMP), and in state waters under Atlantic States Marine Fisheries Commission's Fishery Management Plan for Inshore Stocks of Winter Flounder. The last assessment of these three winter flounder stocks was carried out at the Groundfish Assessment Review Meeting (GARMIII) in 2008. Results of the 2011 review will form the scientific basis for fishery management in the northeast region. Duties of reviewers are explained below in the "**Requirements for CIE Reviewers**", in the "**Charge to the SARC Panel**" and in the "**Statement of Tasks**". The Terms of Reference (ToRs) for the assessment scientists are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is attached as **Annex 4.**

The SARC 52 review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the SSC of the New England or MidAtlantic Fishery Management Council. The SARC panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Requirements for CIE Reviewers: Three 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 fish stock assessments. Reviewers should be familiar with winter flounder (or comparable species) life history and population dynamics.

In general, CIE reviewers for SARCs shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include statistical catch-at-age, state-space and index methods. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of biological reference points that includes an appreciation for the varying quality and quantity of data available to support estimation of biological reference points.

Each CIE reviewer's duties shall not exceed a maximum of 16 days to complete all work tasks of the peer review described herein.

Not covered by the CIE, the SARC chair's duties should not exceed a maximum of 16 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).

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during 6-10 June, 2011.

Charge to SARC panel: The panel is to determine and write down whether each Term of Reference of the SAW (see **Annex 2**) 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 <u>and</u> 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.

Statement of Tasks:
1. Prior to the meeting (SARC chair and CIE reviewers)
Review the reports produced by the Working Groups and read background reports.

Each CIE reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein: 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, and FAX number) 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 other information concerning 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 by FAX the requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the 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.

2. During the Open meeting

Panel Review Meeting: 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 shall 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 herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. 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 draft 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.

4. **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 Points (BRP) or their 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 (see **Annex 4**).

(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 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 SARC 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 4** 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**.

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 at the Woods Hole, Massachusetts during June 6-10, 2011. 3) Conduct an independent peer review in accordance with the ToRs(**Annex 2**). 4) No later than June 24, 2011, 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 to 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.

* 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 approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR 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 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).

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 be completed with 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 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 CIE reports to the NMFS Project Contact and Center Director.

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: 301-713-2363 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 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:

Dr. James Weinberg, NEFSC SAW Chairman Northeast Fisheries Science Center 166 Water Street, Woods Hole, MA 02543 James.Weinberg@noaa.gov (Phone: 508-495-2352) (FAX: 508-495-2230)

Mr. Frank Almeida, Acting NEFSC Science Director National Marine Fisheries Service, NOAA Northeast Fisheries Science Center 166 Water St., Woods Hole, MA 02543 frank.almeida@noaa.gov phone: 508-495-2233

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 a concise summary of whether they accept or reject the work that they reviewed, and 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 others read the SARC 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 the following appendices:

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.

Annex 2: Assessment Terms of Reference for SAW/SARC52

A. Winter flounder (Southern New England Stock)

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.

2. Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.
a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).
b. Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.
c. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

B. Winter flounder (Georges Bank Stock)

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.

2. Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.
a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).
b. Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.
c. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

C. Winter flounder (Gulf of Maine Stock)

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.

2. Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

8. 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) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.
a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). 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 (e.g., terminal year abundance, variability in recruitment).
b. Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.
c. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

Appendix to the SAW TORs:

Clarification of Terms used in the SAW/SARC 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)

Annex 3: Draft Agenda

52nd Northeast Regional Stock Assessment Workshop (SAW 52) Stock Assessment Review Committee (SARC) Meeting June 6-10, 2011

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

DRAFT AGENDA (version: 20 April 2011)

TOPIC PRESENTER(S) SARC LEADER RAPPORTEUR

Monday, June 6

1 – 1:15 PM Welcome **James Weinberg**, SAW Chair Introduction **Patrick Sullivan**, SARC Chair Agenda Conduct of Meeting

1:15 – 3:15 Assessment Presentation (A. SNE Winter flounder) **Mark Terceiro TBD TBD**

3:15 – 3:30 Break

3:30 – 5:30 SARC Discussion w/ presenters (A. SNE Winter flounder) **Pat Sullivan**, SARC Chair **TBD**

Tuesday, June 7

8:30-10:30 AM Assessment Presentation (B. GBK Winter flounder) **Lisa Hendrikson TBD TBD**

10:30-10-45 Break

10:45 – 12:30 SARC Discussion w/ presenters (B. GBK Winter flounder) **Pat Sullivan**, SARC Chair **TBD**

12:30 -1:45 Lunch

1:45 – 3:45 Assessment Presentation (C. GOM Winter flounder) **Paul Nitschke TBD TBD**

3:45 – 4:00 Break

4:00 – 5:45 SARC Discussion w/ presenters (C. GOM Winter flounder) **Pat Sullivan**, SARC Chair **TBD**

(Evening Social/Dinner at **TBD**, 7pm)

Wednesday, June 8

8:45 – 11 11 - 11:15	Revisit w/ presenters (A.) **Pat Sullivan**, SARC Chair Break	**TBD**
11:15 – 12:30	Revisit w/ presenters (B.) **Pat Sullivan**, SARC Chair	**TBD**
12:30 – 1:45	Lunch	
1:45 – 2:45	cont. Revisit w/ presenters (B.) **Pat Sullivan**, SARC Chair	**TBD**
2:45 - 3	Break	
3 – 5:15	Revisit w/ presenters (C.) **Pat Sullivan**, SARC Chair	**TBD**

Thursday, June 9

8:45 – 11 Review/edit Assessment Summary Report (A.)
Pat Sullivan, SARC Chair **TBD 11 - 11:15** Break

11:15 – 12:30 Review/edit Assessment Summary Report (B.) **Pat Sullivan**, SARC Chair **TBD**

12:30 – 1:45 Lunch

1:45 – 2:45 cont. Review/edit Assessment Summary Report (B.) **Pat Sullivan**, SARC Chair **TBD**

2:45 - 3 Break

3 – 5:15 Review/edit Assessment Summary Report (C.) **Pat Sullivan**, SARC Chair **TBD**

Friday, June 10
9:00 -5:30 PM SARC Report writing. (closed meeting)

*All 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.

1. 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.
2. 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.