Report on the SARC Review of SAW 54 Stock Assessments for Southern New England yellowtail flounder and Atlantic herring, June 2012

N. G. Hall

Unit 2
2 Wexford Street
Subiaco
Western Australia 6008
Australia
Phone: +61 8 9401 6891
Email: nghall@iinet.net.au

June 2012

Contents

1. Executive Summary

The SARC for the 54th Stock Assessment Workshop (SAW 54) met at Woods Hole in June 2012 to examine the stock assessments developed for Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*). Overall, the SARC Review Panel found the assessments to be of a high quality. Decisions relating to stock definition, data to be employed, model structure, base models, and Biological Reference Points (BRPs), and findings on stock status, short term projections, and vulnerabilities by each Working Group (WG) appeared sound and were accepted and endorsed by the Panel. The Panel complimented the yellowtail flounder WG for the bridging analyses that they had conducted to determine the implications of the change from the existing Virtual Population Analysis (VPA) assessment model to a statistical catch-at-age ASAP model. Specific areas of concern in the assessments, which were explored during the review, are discussed below.

As recognised by the WG, the stock definition for the Gulf of Maine/Georges Bank (GoM/GB) Atlantic herring defines a "stock complex", which comprises a number of component sub-stocks. The definition is appropriate for the assessment given the current inability to distinguish fish from the different sub-stocks in commercial and survey catches. There is a risk, however, that stock assessments based on this stock definition may fail to detect whether less-productive stock components are overfished or being subjected to overfishing. Continued research on methods of distinguishing individual fish from component sub-stocks is urged.

For both species, there is a need to explore uncertainty of the assessments to the possibility of interchange of fish and/or larvae among the defined stock of interest and adjoining stocks. The sensitivity of the assessments to such interchange could be explored through the development of alternative catch histories and/or incorporation of alternative stock recruitment relationships, *e.g.*, use of a stock-recruitment relationship that assumes that a portion of the total recruitment arises from the spawning stock biomass (SSB) of an adjoining stock.

The decisions by the WGs to assume that natural mortality declines with weight at age in a pattern similar to that described by Lorenzen (1996), and thus varies with both age and year, were not based on available data, and, indeed, at least in the case of the GoM/GB Atlantic herring, the assumption provides no significant improvement over the simpler assumption of a constant natural mortality over age and year. A similar evaluation of the extent to which the more complex mortality assumption improves the fit for Southern New England/Mid-Atlantic (SNEMA) yellowtail flounder should be undertaken. If there is no significant improvement in the fit of the model, the more parsimonious assumption of a constant natural mortality should be adopted. Note, however, that use of the more complex assumption in the 2012 assessments is unlikely to have affected the stock status determinations that resulted from the assessments.

A key issue in the Atlantic herring assessment was whether the increase in natural mortality since 1995 that had been used in the base model was justified. The wealth of evidence on stomach contents, dietary composition and predator abundance that had been used to estimate the annual consumption of Atlantic herring by fish predators, and the resulting trends in that consumption, provided strong support for the decision that natural mortality had increased. This, combined with the improvement in retrospective patterns that resulted from the increased mortality, justified the assumption in the base model of increased mortality since 1995. An

examination during the review of the change in likelihood, Mohn's Rho, and the average of the estimated consumption between 1996 and 2010 resulting from a range of percentage increases in post-1995 mortality indicated that, although uncertain, the 50% increase that had been used in the base model was appropriate. The Panel concluded that, in the immediate future, continued high levels of natural mortality were likely to occur.

A further issue in the Atlantic herring assessment was that, prior to the review meeting, aspects of uncertainty, such as the robustness of the determination of stock status to uncertainty in the magnitude of the post-1995 increase in natural mortality, had not been fully characterized. This was addressed by the various sensitivity runs of the model that were undertaken during the review meeting. The findings that the stock was not overfished and that overfishing was not occurring were robust to the uncertainties that were explored.

The key issue in the SNEMA yellowtail flounder assessment was whether the marked decline in the estimates of recruitment since 1990, which were obtained when fitting the base model, was simply the result of reduced SSB or was unconnected to current low levels of SSB and represented a decline in productivity associated with some extraneous environmental factor and was thus likely to persist, at least in the immediate future. The WG had undertaken a detailed analysis to explore the hypothesis that recruitment had declined as a result of a shrinking and warmer cold pool. Although recruitment variation was found to be associated with a measure indicative of the state of the cold pool, there was no evidence of a change in the latter measure that would explain the marked decrease in recruitment since 1990. A subjective examination during the review of the changes experienced over recent decades in a broader range of ecological and environmental variables failed to identify any specific factor that might explain the decline in recruitment.

While the weight of evidence presented at the review suggested that it was more likely than not (*i.e.,* a probability that the Panel subjectively assessed as roughly 60%) that the decline in recruitment of SNEMA yellowtail flounder was due to a productivity change associated with some unknown factor, the evidence was not sufficient to rule out the alternative hypothesis that the decline was the result of reduced SSB. The Panel therefore endorsed the WG's decision to present BRPs, stock status determinations, and projections for the two competing hypotheses. While it was concluded that, for both scenarios, overfishing was not occurring, the competing hypotheses produced very different conclusions regarding the overfished status of the stock. If the decline in recruitment since 1990 is the result of reduced SSB, it would be concluded that the stock is still overfished and has not yet been rebuilt, with projections indicating that it would be unlikely to rebuild by 2014 even in the absence of fishing. On the other hand, if, as the Panel concluded was more likely than not, the productivity of the stock has been reduced by some unknown factor, it would be concluded that the stock is rebuilt and not overfished. Note that, in this latter case, "rebuilding" is due to the reduction in the biomass reference point rather than increase in SSB.

The assessments produced for the GoM/GB Atlantic herring and SNEMA yellowtail flounder are of a high quality and provide the best scientific advice regarding the status of these two stocks that is currently available. The WGs are commended for their efforts in developing these assessments.

2. Background

2.1. Overview

A Stock Assessment Review Committee (SARC) meeting to review the 2012 benchmark stock assessments for Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*) was held at the Northeast Fisheries Science Center, Woods Hole, Massachusetts, from 5-9 June, 2012. The SARC Review Panel for the 54th Stock Assessment Workshop (SAW 54) comprised, as chairman, Dr Robert O'Boyle (Beta Scientific Consulting, Canada, member of the New England Fishery Management Council's Scientific and Statistical Committee (SSC)), and, as panel members appointed by the Center for Independent Experts (CIE), Chris Francis (NIWA, NZ), Neil Klaer (CSIRO, AU), and Norman Hall (Murdoch Univ., AU) (Appendix 3). The agenda for the Review Workshop is presented in Annex 3 of Appendix 2.

The Statement of Work provided to Dr Norm Hall by the CIE is attached as Appendix 2. This required that, in addition to satisfying the requirement for SARC Panel members to participate in the review and conduct an independent peer review of each assessment, Review Panel members should assist the Review Chairman in preparing a SARC Summary Report of the review, and each should also prepare an independent CIE report of the assessments and the review process. This CIE report, which is prepared in accordance with the last of these requirements, describes my evaluation of the assessments and the review process.

Prior to the SARC Review Meeting, the stock assessment documents and other background documentation had been made available to Panel members. A list of these documents is presented in Appendix 1.

2.2. Terms of Reference

The terms of reference for the stock assessments of the Southern New England yellowtail flounder and Atlantic herring are presented in the Statement of Work (Appendix 2), together with the terms of reference for the SARC review of these assessments.

2.3. Panel membership

Details of the Panel Membership and of other key participants for the SARC review of the SAW 54 stock assessments for Southern New England yellowtail flounder and Atlantic herring are presented in Appendix 3. In particular, the SARC Review Panel members comprised:

- Robert O'Boyle, Panel Chair, member of NE FMC's SSC
- Chris Francis, CIE
- Neil Klaer, CIE
- Norman Hall, CIE

2.4. Date and place

The SARC met on 5-9 June, 2012, at the Northeast Fisheries Science Center, Woods Hole, Massachusetts, to review the benchmark stock assessments for Southern New England yellowtail flounder and Atlantic herring that had been produced for SAW 54.

2.5. Acknowledgments

Thanks are expressed to the various individuals who participated in the review meeting, and who contributed to the stock assessments, for making the review such an interesting and positive experience The WGs and, in particular, the presenters, J. Deroba and L. Alade, are to be commended for the quality of their stock assessments, and their very competent and professional responses to the Panel's queries and requests. Thanks are also extended to the NEFSC SAW Chairman, J. Weinberg, A. O'Brien, and P. Rago (NEFSC) for their assistance in providing access to review materials and ensuring the smooth running of the review meeting, and to the rapporteurs, T. Chute and J. Blaycock, who greatly assisted the Panel by recording the herring and yellowtail flounder discussions. The valuable insights, comments, and recommendations offered during the review meeting by R. O'Boyle, C. Francis, and N. Klaer are gratefully acknowledged.

3. Description of Reviewer's role in review activities

Prior to the review meeting, I familiarised myself with the background documentation, and the assessment and draft assessment summary reports for the two species that were the subject of the review (Appendix 1). Subsequently I attended and actively participated as a Review Panel member in the SARC meeting that was held at Woods Hole. At this meeting, the lead assessment scientists presented details of the data and the assessments, which I and the other Panel members reviewed and assessed. Together with other Panel members, I requested further details regarding specific aspects of the assessment that were of concern and considered and discussed the results of additional sensitivity runs that were requested during the meeting. I also participated in the Review Panel's discussions regarding the adequacy and soundness of the WG's responses to their various terms of reference, and whether the results of the assessments were of an appropriate scientific standard and thus acceptable as the basis for scientific advice for use in management. With other Panel Members, I contributed my suggestions of points to be considered when preparing the SARC Summary Report. Following the Review Meeting, I drafted those sections of the SARC Summary Report, for which I had been assigned responsibility, and offered comment on the resulting draft report. I then focused on preparing this document, *i.e.*, the CIE report describing my evaluation of the two stock assessments and the SARC review.

4. Summary of findings relevant to SARC review of the Terms of Reference of the stock assessments for SAW 54

In this section of the document, I have attempted to present my own assessment of each of the Terms of Reference for the assessments produced for SAW 54. Note however that, although the Statement of Work calls for an independent assessment by each CIE Panel Member, the review process itself and the process of preparing a SARC Summary Report that encompasses the different perspectives of the members of the Review Panel encourages convergence of views by Panel members and a focus on common issues. Note also that I have not attempted to paraphrase several sections of the SARC 54 Panel Summary Report, which I was responsible for drafting and which I have included in my CIE report.

*4.1 Atlantic herring (*Clupea harengus*).*

AH ToR 1. Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

Was the ToR completed successfully?

Yes

Reason for acceptance/rejection

Catches from 1964 to 2011 for the Gulf of Maine/Georges Bank Atlantic herring stock complex by US fixed and mobile fishing gears and by the New Brunswick weir fishery were collated and reported. Discards were only included in catch estimates from 1996. As discards represented generally less than 1% of landings, the WG had considered that the lack of earlier historical data was unlikely to affect the results of the assessment. Age compositions (for ages 1 to 14) from 1964 to 2011 were reported for catches by US fixed and mobile fishing gears and (for ages 1 to 11+) from 1965 to 2011 for those from the New Brunswick weir fishery. Weights at age and proportions mature at age were reported. Plots showing the spatial distribution of herring landings per month from 2007 to 2010 were presented. Although the uncertainty of the catch estimates and biological data were not discussed explicitly in the section of the assessment report dealing with this ToR, uncertainty was incorporated in the assessment through iterative reweighting of the catch CV and estimation of effective sample size for the age compositions of the catches.

Strength of analysis

The Panel was advised in the review meeting that, although observer coverage was only about 5 – 15% in the early 2000s, recent observer coverage has been approximately 30%, with 100% coverage in closed area 1 and 80% coverage on Georges Bank. Thus, overall, coverage has been sufficient to conclude that discards have generally been low.

The fact that strong and weak year classes are evident in the age composition data indicates that the age composition data contain information on year class strength that is likely to be of value to the assessment.

Weakness of analysis

No detail is provided in the Assessment Report of the sampling design and methods used to collect biological samples, nor the methods used to analyse and expand the resulting data to produce the tables of age compositions of annual catches, catch weights at age, spawning stock biomass at age, and proportions mature at age. It

was thus not possible to assess how well the reported data were likely to reflect the true age compositions, weights at age, etc.

The uncertainty of the estimates of catch, catch at age, weight at age and maturity data was not discussed in the section of the assessment report dealing with this ToR.

The catch history that has been developed relies on the adequacy of the assumption that the stock complex and its fishery are constrained spatially to lie within the geographic boundaries that were selected by the WG, and that there is no movement of fish across those boundaries or interchange between stocks/fisheries. There would be value in developing alternative catch histories that allow for uncertainty in the northern boundary of the stock, and possible mixing of fish between the Gulf of Main/Georges Bank and Scotian shelf stocks and the effect of such mixing on catches. By exploring the sensitivity of the assessment to these alternative catch histories, insight would be gained on the uncertainty associated with catches, discards, and the spatial distribution of the stock complex and the catches taken from that complex.

Were conclusions and recommendations acceptable?

The decisions of the WG to use the catch data for the fixed and mobile fishing gears from 1965 to 2011 in the assessment and to pool data for older fish into an 8+ age class were acceptable.

Reason for acceptance/rejection

The similarity of the fishing gears and of age composition data for the New Brunswick weir and US fixed gear fisheries supports the decision to combine the data for these fisheries and treat them as a single time series of catches. The length frequencies and age compositions of the catches from the fixed and mobile gears differ markedly, supporting the decision to treat the data for these two gear types as separate time series of catches. The paucity of data for fish older than age 8 supports the decision to pool these data into an 8+ age class.

Does work provide a scientifically credible basis for fishery management advice?

The catch and age composition data provide a sound basis for their use in developing assessment models aimed at providing fishery management advice.

AH ToR 2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

Was the ToR completed successfully?

Yes

Reason for acceptance/rejection

The WG considered and presented abundance and, where available, age composition data for each of the surveys that was considered to be a candidate for use in the assessment, *i.e.*, the NMFS spring and fall bottom trawl surveys, the NMFS winter survey, the NMFS summer shrimp survey, the ichthyoplankton larval surveys, the Massachusetts Department of Marine Fisheries (MA DMF) spring and fall bottom trawl surveys, and the Maine/New Hampshire spring and fall bottom trawl surveys. Predator consumption rates were considered under ToR 6. The potential use of commercial LPUE was considered and discussed by the WG in its assessment report. Uncertainty of abundance data was characterised by the CV of the survey abundance

estimates. Although no details of sample sizes were presented for the age composition data, estimates of effective sample size were determined through the iterative approach described by the WG under ToR 5. The MA DMF and Maine/New Hampshire surveys were eliminated from use in the assessment because of inadequate spatial coverage of the stock complex, while the NMFS winter survey was eliminated by the WG due to inconsistent spatial coverage and lack of fit. Because the relationships of the larval index to both its parent spawning stock biomass and its subsequent age-1 recruitment were unclear, the WG eliminated this index. Factors that affected the consistency of the NMFS spring and fall surveys, such as change of trawl boards or change of survey vessel, were also considered and dealt with by breaking the time series at 1985, in the case of the change in trawl boards, or through adjustment by calibration factors in the case of survey vessel change.

Strength of analysis

The long time series of spring and fall survey data, collected using a sampling protocol that provides consistent spatial coverage, provide valuable information on trends in abundance and age composition of Atlantic herring.

The decision not to apply age-length keys derived from commercial catch data to survey length data collected prior to 1987 is strongly endorsed as such application is inappropriate and would introduce error.

Weakness of analysis

Although unavoidable, the breaks in consistency of the nets, trawl gear and vessels used in the spring and fall surveys have reduced the value of these two long time series.

While the pattern of residuals for the annual spring and fall survey indices supports a decision to break the two time series in 1985, the trend in the residuals of an earlier model that did not break each time series into two sections suggests that the transition occurred over a number of years, and was not entirely consistent in timing with the change in trawl doors. While the decision to break the time series at 1985 is endorsed, consideration should be given to identifying alternative hypotheses that might explain the trend in the residuals from the earlier model.

The trend with length of the length-specific calibration factors used to convert data collected by the FSV Henry B. Bigelow for the fall survey to FRV Albatross IV equivalent indices is unusual and a process to justify such complexity should be identified.

Uncertainty associated with the calibration factors used to convert data collected by the FRV Delaware II and FSV Henry B. Bigelow to FRV Albatross IV equivalent indices is not carried through to the results of the assessment model.

Were conclusions and recommendations acceptable?

Yes

Reason for acceptance/rejection

The decision to use the NMFS spring and fall surveys, and the NMFS summer shrimp survey, in the assessment was justified as each survey was considered to provide good spatial coverage of the stock complex, and thus survey indices and age compositions were likely to be representative of the stock.

Rejection of the state surveys was justified as inadequate coverage of the Atlantic herring stock complex would lead to a non-representative sample of the stock complex with abundance and age composition estimates reflecting only the abundance and age composition of that portion of the stock complex available within the area covered by the state surveys. For the state survey data to serve as a useful indicator, it would be required that the proportion of each age class available within the surveyed

area is constant among years, although possibly varying among age classes. Further exploration should be undertaken to assess the validity of such an assumption before any decision is made to incorporate data from the state surveys as indices of abundance in the assessment.

The decision to reject the NMFS winter survey as an index of abundance is justified as inconsistent spatial coverage will be reflected in the indices of abundance that are calculated. The resulting indices of abundance would not reflect the trend in abundance of the stock complex.

The decision to reject the larval survey index was justified because of the WG's uncertainty that this index reflected the abundance of either the parent spawning biomass or the subsequent age-1 recruitment.

Rejection of commercial LPUE as an index of abundance of Atlantic herring is justified, given the changes in spatial distribution of fishing that have resulted from regulation changes, the potential for hyperstability resulting from the use of sonar to locate and track schools, and the difficulty in identifying fishing effort directed towards Atlantic herring due to target switching during trips.

Does work provide a scientifically credible basis for fishery management advice?

Abundance and age composition indices calculated from the NMFS spring and fall surveys, and the NMFS summer shrimp survey, provide a scientifically credible basis for fishery management advice.

AH ToR 3. Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG presented details of the NEFSC fall acoustic survey and considered the trend in the resulting estimates of abundance of Atlantic herring, which exhibited an abrupt and marked decline from values observed till 2001 to those observed in subsequent years. Plots of survey tracks relative to the spatial extent of the areas of the Georges Bank used in the 1999 and 2012 assessments, which could be compared with plots of the spatial distributions of monthly catches, were presented in the Assessment Report. The trends in the estimates of abundance calculated using the acoustic survey data were compared with those derived from other bottom trawl surveys.

Weakness of analysis

When comparing the estimates for 2006 from the NEFSC fall acoustic survey with those from the long range sonar (OAWRS) study, it is stated in the assessment report that "All approaches were consistent to within 20% or less". This statement appears inconsistent, however, with the values shown in Table A3-4 of the Assessment Report.

Were conclusions and recommendations acceptable?

The decision to exclude the acoustic survey index from the assessment was acceptable.

Reason for acceptance/rejection

The marked decline in the acoustic index from the values observed between 1999 and 2001 to those observed in subsequent years was inconsistent with the trends in other survey indices and fishery monitoring data. Although several hypotheses had been proposed to explain the decline, there was insufficient evidence in support of any specific hypothesis to justify its incorporation in the model. The suggestion that the decrease could be explained by a mis-match between the timing of the survey and the time at which spawning occurred was discounted by the results from the larval study, which back-calculated spawning time.

Other hypotheses considered were differences in fish behaviour and vertical distribution during the different years of the acoustic survey, but the factor(s) involved has/have yet to be identified. It should be noted, however, that the acoustic survey extends over a limited spatial area and indices are thus representative of only a portion of the stock, whereas the other survey indices are representative of the entire stock complex.

Does work provide a scientifically credible basis for fishery management advice?

The decision to exclude the acoustic survey index of abundance was appropriate, as its inclusion would have been likely to have led to biased fishery management advice.

AH ToR 4. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG collated data on the spatial distribution of Atlantic herring derived from fishery-independent survey and ichthyoplankton data, geographic variation in biochemistry, growth, and morphology, and information on movements and migration derived from ichthyoplankton dispersion data and tagging studies (both within and outside the Gulf of Maine/Georges Bank region). It then re-examined those data to determine whether the current stock definitions should be changed.

Strength of analysis

The re-assessment was based on a comprehensive review of available data.

Weakness of analysis

The decision to maintain the current stock definition, *i.e.*, to treat Atlantic herring in the Gulf of Maine/Georges Bank as a stock complex, was based on the fact it is currently not possible to distinguish from which of the different component sub-stocks within this region the fish in commercial and survey catches are drawn. It was also based on the assumption that fish from the Scotian Shelf stock and those from the Gulf of Maine/Georges Bank stock complex remain separate. The WG "acknowledged some degree of mixing of Scotian shelf stocks with U.S. stocks", suggesting that there would be value in considering the sensitivity of the results from the base model for the 2012 stock assessment to an alternative stock structure with some level of mixing of fish from the Scotian shelf in catches and survey data, *e.g.*, through use of alternative catch histories or movement rates from tagging studies. Currently, there is no assessment of the uncertainty associated with the stock structure that has been adopted by the WG.

The WG reported in its Assessment Report that several studies, which had examined morphometric and otolith variation, had successfully identified the stock of origin with considerable (70 to 88%) accuracy. If further research demonstrates that this approach is reliable, it may offer the opportunity to classify sampled fish from commercial and survey catches according to their stock of origin.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The decision to treat the Gulf of Maine and Georges Bank stocks as a single stock complex is justified by the fact that the stock of origin of fish in commercial and survey catches cannot be identified.

Does work provide a scientifically credible basis for fishery management advice?

Yes. While the stock definition is considered appropriate for current use in determining management advice, the potential exists that, even though the stock complex may be assessed as not overfished with overfishing not occurring, a less-productive sub-stock within this complex may be overfished or being subjected to overfishing.

AH ToR 5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from AH TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The ASAP model used for the 2009 TRAC assessment was updated by the WG with data through 2011. Because of the retrospective problems encountered with the 2009 assessment, which remained present with the extended time series, all data inputs and model settings were reviewed. While the 1999 ASAP model was fitted to pooled catch data and assumed a selectivity of 1, the revised base model was fitted to the time series of catches for fixed and mobile gears and estimated selectivity. Selectivity was also estimated for the survey data. In contrast to the 1999 assessment, age compositions for survey data prior to 1987, for which no age data had been collected, were not estimated using age-length keys derived from commercial data as had been done in the previous assessment. Age 8 was treated as a plus group. The key modification to the assessment was the replacement of the assumption that natural mortality M was equal to 0.2 by the assumption that, for each year, the average level of natural mortality across all ages was 0.3, and that the relative level of natural mortality at age declined with average weight at age in accordance with relative levels of natural mortality predicted by the equation derived by Lorenzen (1996) for ocean fish. The value of 0.3 was derived from a maximum age for Atlantic herring of 14 years, using Hoenig's (1983) relationship between mortality and maximum age for fish. The resulting estimates for natural mortality for each age were smoothed across years using a General Additive Model (GAM). Age-dependent mortalities from 1996 to 2011 were increased by 50% to reflect the increased consumption of Atlantic herring that had occurred in those years. This increase largely resolved the retrospective pattern in SSB that had been present.

Estimates of annual fishing mortality, age-1 recruitment, spawning stock biomass and total biomass calculated using the base model were reported by the WG in their Assessment Report. The posterior probability distributions of SSBs and Fs were determined using MCMC and time series plots of these variables, with 80% confidence limits, were plotted. Similar results for age-1 recruitment and total biomass were not presented in the Assessment Report, however. Sensitivities of SSB and F to the inclusion/exclusion of the winter, larval and acoustic surveys and to use of F=0.2 for all ages and years were explored. The WG also presented a plot showing the sensitivity of the estimates of SSB, F, and age-1 recruitment, but not total biomass, to an ASAP run using fish consumption as a fleet and a run without the 50% increase in natural mortality from 1996 to 2011. The sensitivity of the values of F and SSB relative to respective reference points for varying values of steepness of the Beverton and Holt stock-recruitment relationship was also examined.

Retrospective patterns for SSB, F, and age-1 recruitment were explored for the base run, and the time series of estimated SSBs and Fs from the 1995, 2005, and 2009 assessment models were compared with those from the base run. No comparison with previous projections was made.

Strength of analysis

Use of the ASAP model, which has been well tested, was appropriate for this assessment. It was also pleasing to note that other model structures, such as SS3, are being explored.

The uncertainty in the assessment was explored using MCMC and alternative runs to assess sensitivity to model assumptions.

Examination of the likelihood profile demonstrated that the steepness parameter of the stock-recruitment relationship was very imprecise, with a 95% confidence interval ranging from less than 0.35 to 0.85 and with associated biological reference points and estimates of SSB and fishing mortality that varied greatly with the changes in steepness. Despite this, the conclusions regarding stock status remained unchanged from the findings obtained from the base model over the range of steepness values except in the case of an atypically low steepness of 0.35, where the results indicated that overfishing was occurring.

Weakness of analysis

The use of the Lorenzen (1996) model to introduce an age-dependent mortality is not justified as it increases model complexity and, in the case of the Atlantic herring model, is based on an assumption rather than support from observed data. Furthermore, advice was provided at the review meeting that the model using Lorenzen-based age-dependent mortalities failed to improve the fit relative to a model that assumed constant M at age.

The Panel recommended that effective sample sizes for age compositions should be calculated using iterative reweighting based on mean age, and possibly reflecting relative magnitude of initial sample sizes.

It is recommended that a more appropriate CV be imposed on recruitment deviations. Mertz and Myers (1996) report that, for most fish, the value of $\sigma_{\log_e R}$ falls within the range from 0.3 to 1. The base model imposed a minimal constraint on the deviations, *i.e.*, CV=1, which may have influenced the 2009 age-1 recruitment estimate, which was estimated to be approximately twice as great as the maximum of recruitment estimates for earlier years and which, as a consequence, influenced estimates of recent biomass.

The substantial uncertainties in the calibration factors employed to adjust the survey data for the effect of changes in survey vessel could not be carried through to the estimates produced by the model.

The effect of increased or decreased weighting for different survey or age composition data on stock assessment results was not examined, and thus it was not possible to assess whether the different survey indices exhibited tension in the support they exhibited for different values of key model parameters.

The extent to which survey catchability was estimated to increase when the trawl doors were changed in 1985 was very different in the fall and spring surveys, *i.e.*, 2.64 for the spring survey *cf*. 13.6 for the fall survey. This, combined with the fact that the residuals for the spring and fall indices for an earlier model that assumed constant catchability for each time series exhibited a trend suggesting transition over a number of years and timing that was not entirely consistent with the trawl door change, suggests that another factor may have been involved. This deserves further investigation.

Were conclusions and recommendations acceptable?

The Panel accepted the base case, which was based on M=0.3, with a 50% increase for years after 1995, as the most plausible model for provision of management advice.

Reason for acceptance/rejection

The temporal trend in consumption data for Atlantic herring justifies the assumption that M would have increased in the later years of the time series. While inclusion of the consumption data in an alternative model was explored, the WG concluded that "the inter-annual variation of the fish predator consumption estimates was not well understood and was beyond what would be expected from a relatively constant predator fleet", and such a model was therefore eliminated from further consideration. Nevertheless, the average of the consumption estimates was clearly greater from about 1996, and thus, although unusual for an assessment model, an assumption that M increased in this later period is justified.

The Panel examined the results obtained during the review meeting by fitting the base model but using alternative values of percentage increase in post-1995 natural mortality. Estimates of likelihood, degree of retrospective pattern (Mohn's Rho), and extent to which the average level of implied consumption in the post-1995 period matched the estimates of consumption were compared. From this, the Panel concluded that the 50% increase in post-1995 mortality assumed in the base model was appropriate and plausible, producing post-1995 consumption estimates that were realistic and retrospective patterns that were acceptable.

The Panel was concerned that the estimate of recruitment for the large 2008 year class might have been affected by the relatively large CV that had been imposed on recruitment deviations in the base case model. However, although an exploratory run using a reduced CV lessened the estimated strength of the 2009 age-1 recruits, the resulting recruitment estimate for the 2008 year class remained the highest in the time series.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The base case model produced by the WG provides a credible basis for use in developing fishery management advice. Continued high levels of consumption of Atlantic herring by fish predators is expected in the immediate future, but monitoring should be undertaken to assess whether this persists in the longer term.

AH ToR 6. Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

Was the ToR completed successfully?

The implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship were considered and the results were used to inform the stock assessment. The uncertainty of the estimates of consumption by fish predation was not characterized.

Reason for acceptance/rejection

The WG undertook a comprehensive examination of the data on predator abundance and associated consumption data. The predators, which were considered, included fish, two highly migratory species of fish, marine mammals and sea birds. Estimates of consumption were calculated by combining estimates of predator abundance with estimates of total consumption per predator and dietary composition. Although estimates of consumption of Atlantic herring were produced, the uncertainty of these consumption estimates was not addressed or carried into the assessment. Estimates of consumption were included in an exploratory run of the assessment model but did not improve model fit. Estimates of consumption were used, however, to inform the decision to increase post-1995 natural mortality in the base assessment model.

Strength of analysis

The comprehensive data that are available from stomach analyses for fish from the Gulf of Maine/Georges Bank region provides a strong basis for developing estimates of the consumption of Atlantic herring by the various fish predators. This, coupled with estimates of predator abundance from long time series of fishery-independent survey data, has allowed the construction of a time series of annual consumption of Atlantic herring, which, given the uncertainties involved, appears reasonably sound.

It was noted by the Panel that most of the annual signal in the consumption of herring was related to stomach contents rather than predator abundance.

Weakness of analysis

Abundance estimates of some predator species were calculated from swept area calculations, rather than assessment models.

The later peak in the time series of annual consumption estimates was driven by the very high abundances of two individual predator species, but similar high levels of abundance for those species were not present in adjacent years. Greater inter-annual consistency might have been expected.

The uncertainty of the consumption estimates needs to be assessed.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

It is unusual to consider annual consumption in stock assessment as, in most cases, a time series of estimates of this variable is not available. In the case of the Gulf of Maine/Georges Bank Atlantic herring, however, the wealth of available annual information from stomach contents from a wide range of herring predators (fish, mammals, other), and annual abundance estimates of those predators, has afforded the opportunity to calculate a time series of the annual consumption for this

stock complex. Despite high uncertainty in the data, the consumption estimates provide strong evidence that consumption of herring has increased since 1995. It was noted, however, that the consumption estimates, which have been calculated, are likely to be an underestimate of total consumption and that inclusion of estimates of consumption in an exploratory run of the assessment model did not improve model fit.

Advice received by the Panel during the review meeting indicated that, given the current trends in abundance of predator populations, consumption of Atlantic herring is likely to remain high.

Does work provide a scientifically credible basis for fishery management advice?

Yes, but the uncertainty of consumption estimates needs to be recognised.

AH ToR 7. 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, FMSY and 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.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

Existing stock status definitions for "overfished" and "overfishing", which were based on the fit of a Fox surplus production model, were reported in the Assessment Report. Updated MSY reference points, which were based on the Beverton and Holt (BH) stock-recruitment relationship fitted internally in the base model, were derived. Values of these reference points, together with 80% confidence limits determined from an MCMC analysis, were reported. The WG commented on the scientific adequacy of the existing and new reference points and concluded that the new reference points were an improvement on the existing reference points.

Strength of analysis

The analysis was sound. Uncertainty in the reference points associated with age-specific M, the percentage increase in M since 1995, and steepness of the Beverton and Holt stock-recruitment relationship, was explored.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel noted that the existing biological reference points were derived using an age-aggregated rather than age-structured model, and based on an analysis on catch data aggregated over gear type rather than an analysis that employed catch by mobile or fixed gear. The Panel endorsed the redefinition of the biological reference points to those derived using the base model from this assessment.

It was recognised that, as these biological reference points were based on conditions prevailing in 2011, the WG assumed a continued level of high natural mortality due to consumption by predators. This assumption was considered appropriate for the immediate future, *i.e.*, for the next three to five years, but there was uncertainty as to whether such high mortality would continue to prevail in the longer term. Continued monitoring of predation was recommended.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The updated estimates of the BRPs, *i.e.*, Fmsy, SSBmsy, MSY, and the value of ½MSY used as the reference level when assessing whether the stock is overfished, which were determined using the Beverton and Holt stock-recruitment relationship fitted internally in the new base model, provide a scientifically credible basis for fishery management advice.

AH ToR 8. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from AH TOR-7).

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The existing model was updated by the WG with data through 2011 and stock status was evaluated with respect to the existing reference points. Stock status was also evaluated using the estimates and new MSY reference points determined using the base model for the 2012 assessment.

Strength of analysis

The status of the stock was determined for a range of both age-specific and post-1995 values of M, and for a range of values of the steepness parameter of the stock-recruitment relationship. Apart from the unlikely case in which steepness was assumed to be 0.35, the conclusion was that the stock was not overfished and overfishing was not occurring.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel agreed that, while there is uncertainty in the values of the reference points and estimates of SSB and F, the conclusions that the stock is not overfished and that overfishing is not occurring appeared robust to that uncertainty. The conclusion using the previous model was consistent with that from the new base model and new reference points.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The assessment indicates that the stock is not overfished and that overfishing is not occurring.

AH ToR 9. Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

Was the ToR completed successfully?

No, although some initial work is underway.

Reason for acceptance/rejection

Until now, it appears that the focus of the WG has been on developing the ASAP model and exploring its uncertainties. Although research projects have been initiated to explore some aspects of model fitting and assessment of stock status, *e.g.*, the consequences of ignoring retrospective patterns, no studies to undertake a formal management strategy evaluation have yet been initiated.

The Panel recommended that development of a management strategy evaluation framework for the Atlantic herring should be pursued, but suggested that, as a first step, this would require the engagement of fishery managers and stakeholders to consider the specification of formal harvest control rules and identify the alternative harvest policies that should be explored. Management Strategy Evaluation would assist in exploration of strategies to achieve management objectives under the largest uncertainties that exist, *e.g.*, whether high values of natural mortality will persist.

AH ToR 10. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The base model was employed by the WG to project the likely outcomes in 2013, 2014, and 2015 of employing alternative fishing mortalities or catches. Probabilities of exceeding threshold BRPs for F or falling below threshold BRPs for SSB were computed, and the implications of uncertainty in steepness and natural mortality were explored. Factors such as the unknown contribution of fish from the Scotian shelf and the possibility that the strength of the 2008 year class is overestimated, which would affect the stock's vulnerability to becoming overfished, were discussed.

Strength of analysis

Assessment uncertainties were well described.

Drawing numbers-at-age in 2012 from the results of the MCMC simulations of the base ASAP run ensures that uncertainty of the initial age composition at the start of the projection period is well characterised.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The WG has explored the effect on projections of a range of assumptions and the most important sources of uncertainty, *i.e.*, age-specific M, percentage increase in post-1995 M, steepness of the stock-recruitment relationship and size of the 2008 year class.

Vulnerability of the stock to various factors, such as (1) whether the high natural mortality experienced in recent years is likely to continue over the next three to five years, (2) whether the size of the 2008 year class is as large as has been estimated, and (3) the fact that the assessment is based on data for a stock complex rather than a single stock and that fish from the Scotian shelf make an unknown contribution to catches, were considered.

Does work provide a scientifically credible basis for fishery management advice?

Yes. Projections made using the base model for a range of alternative harvest scenarios provide a credible basis for fishery management advice.

AH ToR 11. For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG determined that no research recommendations had been proposed at previous assessments. Research recommendations arising from the current assessment were reported.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The list of recommendations proposed by the WG in the Assessment Report needs to be sorted into priority order. Consideration should be given to modifying recommendation f to "Determine the depth distribution of Atlantic herring and the factors that influence that distribution". Recommendation g appears too vague and requires clarification. The recommendations to collect additional data and to improve the discrimination of stocks and sub-stocks, and to monitor the abundance of the various sub-stocks are endorsed, as modelling the Gulf of Maine/Georges Bank Atlantic herring as a stock complex masks the possibility that a less-productive sub-stock might be overfished or experiencing overfishing. The recommendations associated with improving the estimation of the consumption of Atlantic herring are also endorsed as they offer the potential that future models might explicitly incorporate consumption of herring as an explanatory variable affecting natural mortality, thereby allowing the variability and uncertainty associated with this component of natural mortality to be carried through to estimates of BRPs and determination of stock status. Other recommendations related to improvements to data and methods for stock assessment.

The Review Panel identified a number of research areas that should also be considered. These are listed below.

Data

- Develop alternative catch histories to allow for interchange of fish from the Gulf of Maine/Georges Bank stock complex with those from the Scotian shelf stock, and use these to explore the uncertainty associated with the boundary of the Gulf of Maine/Georges Bank stock area.
- Use the alternative catch histories developed in the previous research item to explore whether catch under-estimation (*e.g.,* inclusion of Scotian shelf catch) can contribute to the reduction in the retrospective pattern and contribute to or explain the need for increased M.

Model

- Explore the effect of imposing a penalty to introduce greater similarity in the relative increase in catchability experienced by the spring and fall survey data in response to the change in trawl doors.

Management Strategy Evaluation (MSE)

- Jointly with managers and stakeholders, identify formal harvest control rules and harvest strategies to be explored, and major uncertainties to be considered, and based on this information, undertake a management strategy evaluation.

*4.2 Southern New England yellowtail flounder (*Pleuronectes ferrugineus*)*

YF ToR 1. Estimate landings and discards by gear type and where possible by fleet, from all sources. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG collated and reported landings and estimated discards of SNEMA yellowtail flounder between 1935 and 2011. Of these data, only the time series of total catches (landings + dead discards) from 1973 to 2011 was employed in the assessment. For 1994 to 2011, the Assessment Report provided estimates of landings (and CVs) by commercial fishers employing trawl (88 to 99% of total landings), scallop dredge, gillnet, and other/unknown fishing gear and fishing in the area over which the stock is distributed. Discards (and CVs) were also collated and reported (by gear type) for the same period. Catch at age of commercial landings between 1994 and 2011 was calculated using a revised length-weight relationship. Plots of the time series of landings reported from catches from different statistical areas were included in the Assessment Report.

Strength of analysis

The data reported for the 2012 assessment reflect a marked revision of data and assumptions from the previous assessment in 2008.

The precision of estimates of discards has been improved by use of spatial stratification.

The revised length-weight relationships have improved estimates of weight at age, and, through their effect on conversion of catch weights to numbers at age, have led to improved estimates of age compositions.

An improved estimate of discard mortality was determined using results from a Reflex Action Mortality Predictor (RAMP) study. The decision to use 90% as the

estimate of discard mortality in the model, and thereby to allow for additional post-release mortality of discarded fish, was endorsed.

Weakness of analysis

There is considerable uncertainty associated with the quantities of fish that are discarded. Between 1994 and 2011, the CVs of estimated discards have ranged from 14 to 75%. Since 2005, partly in response to restrictive trip limits, discards have become a significant proportion (~60%) of the catch of SNEMA yellowtail flounder, and thus uncertainty associated with discards affects the uncertainty associated with total removals from the stock.

Observer coverage was limited in the late 1990s and early 2000s, but has subsequently improved, particularly with the contribution of the ASMs since 2010, in which year approximately 30% of total groundfish trips were covered. This increased observer coverage should improve the precision of estimates of discards.

It is recommended that the sensitivity of the assessment to uncertainty in the estimates of total catch is explored through the construction and use in the ASAP model of alternative catch histories that vary the estimated proportion of discards.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

Although estimates of total catch were affected by the uncertainty associated with discards, the values appeared sound and appropriate for use in the assessment.

Does work provide a scientifically credible basis for fishery management advice?

Yes.

YF ToR 2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of commercial or recreational LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG considered the four survey indices that were available, *i.e.*, the NMFS spring and fall bottom trawl survey, the winter survey, and the larval survey, and whether these were likely to provide indices that were representative of the stock of yellowtail flounder. Calibration factors, which were used to adjust for the effects of net change, change in trawl gear, and change in survey vessel, were described in the Assessment Report. The WG also considered whether it was appropriate to employ commercial CPUE as an index of abundance of yellowtail flounder.

Strength of analysis

The decision by the WG that commercial CPUE was unlikely to be a useful indicator of abundance because of changes in management regulations, changes in reporting methodology, and the change from a directed to a bycatch fishery was endorsed by the Panel.

Weakness of analysis

The value of the spring and fall bottom trawl survey indices is reduced by the inconsistencies that were introduced through the changes that occurred in net type (Yankee 36/41), trawl boards, and survey vessel. Although calibration factors were

determined to adjust for these inconsistencies and convert the data to "Albatross IV" equivalent data, the uncertainties associated with these calibration coefficients are currently unable to be carried through to the assessment and reflected in estimates of biological reference points and determination of stock status.

The conclusion that estimates of abundance for bottom trawl survey (BTS) tows undertaken during the day were similar to those taken during the night and that an aggregated index could therefore be used appears to be incorrect. The fact that the abundance estimates for evening tows consistently exceeded those for daylight tows indicates that improved indices of abundance could be obtained by taking the time of day, *i.e.*, day or night, into account.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The decision by the WG that, because of poor sampling, the southern strata should be excluded when calculating abundance indices for the winter survey was endorsed.

The Panel endorsed the use of data from the NMFS spring and fall bottom trawl survey, the winter survey (with exclusion of the southern strata), and the larval survey. It was noted that the net type used for the larval surveys changed following the MARMAP survey (1977-1987), which justifies the subsequent decision to split the series when fitting the ASAP model.

Does work provide a scientifically credible basis for fishery management advice?

Yes.

YF ToR 3. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG examined data relating to the geographic distribution of fishing, abundance data from fishery-independent surveys, spawning and ichthyoplankton distributions, distributions of juveniles and adults determined from bottom trawl surveys, genetics, life history patterns, morphology, ichthyoplankton dispersion, tagging observations, and patterns of parasite infection. In assessing whether the existing stock definition was appropriate, knowledge of migration among stock areas was taken into account.

Strength of analysis

A comprehensive review was undertaken of the available data.

Weakness of analysis

Tagging data demonstrate that some mixing occurs among stock components. The possibility also exists that some recruitment to the SNEMA stock arises from a contribution of larvae from outside the boundaries of the SNEMA stock area. The sensitivity of the assessment to such mixing was not assessed.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel endorsed the WG's conclusion that, for the purposes of management and assessment, it is appropriate to manage the yellowtail flounder that

occupy the waters off the east coast of the U.S. as three separate stocks, of which the SNEMA stock is the southernmost.

Does work provide a scientifically credible basis for fishery management advice?

Yes. There is a strong case for yellowtail flounder that occupy the southern New England (mid-Atlantic Bight) region to be considered a single stock for management purposes.

YF ToR 4. 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 a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The forward-projecting statistical catch at age ASAP model, which was accepted as the base case model, was used to produce estimates of annual fishing mortality, age-1 recruitment and both spawning stock and total biomass for the time series. MCMC was used to explore the uncertainty in estimates of SSB and F in 2011. Estimates of uncertainty were not reported for other values in this time series. A large number of alternative model runs were undertaken, however, to explore uncertainty/sensitivity associated with alternative model assumptions. A retrospective analysis was undertaken for SSB, F, and age-1 recruitment. The time series of SSB, F, and age-1 recruitment produced by the base model in the 2012 assessment was compared with the time series estimated using the base case models from previous assessments. No comparison was made with previous projections.

Strength of analysis

The runs that were undertaken to serve as a bridge between the existing VPA model and the new ASAP model provided a valuable demonstration of the consequences of the various modifications to the data that were made and the effect of the change to the new model structure.

Recognising the changes to fisheries regulations that have taken place, but noting that an iterative process comparing model fit and residuals from a number of alternative runs was employed, the final decision to use six fishery selectivity blocks that broadly matched the periods over which the various fishery regulations applied appears reasonable

Weakness of analysis

The decision to use a Lorenzen (1996) curve to introduce a declining trend in mortality with age should be tested against the alternative hypothesis that natural mortality is age independent to ensure that unnecessary complexity has not been introduced to the model's structure.

The sensitivity of model outputs to the choice of the value of M to which the Lorenzen (1996) curve was scaled needs to be assessed.

The Panel recommended that effective sample sizes for age compositions should be calculated using iterative reweighting based on mean age, and possibly reflecting relative magnitude of initial sample sizes.

The base model did not employ an internal stock-recruitment relationship as high values of recruitment for the 1980 and 1987 year classes at very low to moderate stock sizes made it difficult to include such a relationship.

The substantial uncertainties in the survey calibration factors used to "standardize" the spring and fall bottom trawl survey indices to "Albatross IV" values were not carried forward into the estimates produced by the new base model.

Uncertainty in catch history, particularly in discards, was not explored.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel endorsed the decision to use ASAP rather than VPA for the assessment and the use of the proposed base model (run 26) as the model to be employed to determine reference points and stock status.

The trends in SSB and fishing mortality estimated by the new base model were comparable to those produced in previous assessments.

The fact that, with the exception of the estimate for one year, model estimates of age-1 recruitment since 1990 have been consistently very poor, supports the view that, because of some unknown factor, survival from spawning to recruitment has declined from that experienced prior to 1990. Without a mechanism to explain the decline in recruitment, however, and additional evidence, it would be inappropriate to conclude that there has been a regime shift. The possibility that, in recent years, the stock has been held at low spawning stock biomass by some other factor, such as additional unrecognised catches, cannot be entirely dismissed. The Panel therefore considered it to be more likely than not that the fishery has moved into a low recruitment state, and therefore endorsed the decision to carry forward two scenarios when considering biological reference points and stock status, *i.e*., a two-stanza stock-recruitment scenario in which recruitment for SSB values less than 4,300 mt reflect recent recruitment levels and recruitment for greater SSB values reflect the recruitment levels experienced prior to 1990, and a recent-recruitment scenario in which recruitment levels reflect those for the period since 1990. The Panel also agreed that projections should be based on the recent recruitment scenario.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The proposed base model was accepted as providing a credible basis for fishery management advice, although there was uncertainty as to whether recent low recruitment reflected a response either to low SSB or to reduced productivity resulting from some other, unknown factor.

YF ToR 5. Investigate causes of annual recruitment variability, particularly the effect of temperature. If possible, integrate the results into the stock assessment (YF TOR-4).

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG investigated the hypothesis that recruitment of yellowtail flounder was related to the dynamics of the cold pool, and that the effect on the cold pool of increasing temperatures had led to a decline in recruitment for this stock. Such a hypothesis is consistent with the fact that a species at the southern edge of its distributional range, such as SNEMA yellowtail flounder, is likely to be sensitive to oceanographic conditions The first principal component (PCA1) of a Principal Component Analysis (PCA) of 15 variables that were considered to be potential indicators of the dynamics of the cold pool was found to act as a useful measure, with a positive value reflecting a small, warm cold pool and a negative value reflecting a

large, cold pool. The WG then explored the relationship between recruitment and PCA1 by fitting an alternative ASAP model to the data for the yellowtail flounder with a Beverton and Holt stock-recruitment relationship and examining the improvement in fit obtained by including PCA1 as a covariate in the stock-recruitment relationship.

Strength of analysis

The analysis undertaken by the WG demonstrated that the dynamics of the cold pool affected the variability of recruitment of SNEMA yellowtail flounder, but that this did not provide an explanation for the sudden decline in recruitment in 1990 and the subsequent long run of low annual recruitment.

Weakness of analysis

Over a time period longer than the time series considered in the stock assessment, PCA1 displayed an increasing trend with time. Such a result is consistent with the hypothesis that the cold pool has become smaller and warmer with increasing water temperature. However, in the shorter period represented by the time series of data included in the stock assessment, no trend in PCA1 is evident. Moreover, PCA1 exhibited no marked change that would explain the "step-like" decline in recruitment that occurred in 1990.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel noted that, while inclusion of PCA1 as a covariate in the stock-recruitment relationship improved the fit of an ASAP model that assumed a Beverton and Holt stock-recruitment relationship, it did not explain the marked decline in recruitment that occurred in 1990, It was recognised, however, that the base model (run 26) did not include a stock-recruitment relationship, as high values of recruitment for 1980 and 1987 year classes at very low to moderate stock sizes had made it difficult to include such a relationship. The Panel agreed with the decision of the WG that PCA1 should not be incorporated into the base model for the assessment.

During the review, the Panel was presented with broader evidence of oceanographic change within a number of time series of environmental and biological data. While these data, and particularly multiple temperature data sets, revealed a general warming trend, there was no evidence of a step-like change in 1990 that would offer an explanation for the decline in recruitment in 1990 and subsequent low levels of annual recruitment. That is, a link between reduced recruitment and oceanographic changes to explain the sudden and sustained drop in recruitment could not be found.

The Panel agreed with the WG's conclusion that the model represented by "run 26" should continue to serve as the base model.

Does work provide a scientifically credible basis for fishery management advice?

Yes. Provided that it is recognised that the SNEMA yellowtail flounder stock is currently experiencing a long run of low annual recruitment, use of the base (run 26) model provides a scientifically credible basis for fishery management advice.

YF ToR 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, FMSY and 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.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG reported the existing stock status definitions for "overfished" and "overfishing", and then considered whether these reference points should be updated or redefined. Because of the changes in length-weight relationships, the assumption of age-dependent natural mortality scaled to a new average level, and changes in discard data, the WG determined that the reference points needed to be updated. As analytic model-based estimates were not available, the WG recommended and calculated the values (and 90% confidence intervals) of alternative proxies for the BRPs. The adequacy of the existing and new BRPs was discussed.

Strength of analysis

Two recruitment scenarios were considered when calculating the values of the BRPs for the SNEMA yellowtail flounder. The first scenario, which was considered to be more likely than not, reflected the recent low recruitments that had been experienced over the last two decades, *i.e.*, reduced productivity since 1990. The second scenario reflected both earlier and more recent recruitment, with recruitment being dependent on SSB and a break point at 4,300 mt, *i.e.*, recruitment reflecting recent low levels of recruitment when SSB < 4,300 mt, otherwise reflecting the recruitment levels experienced prior to 1990. The Panel adopted the term "recent" to describe the former scenario and "two-stanza" to describe the latter. The Panel considered that the evidence was insufficient to rule out either scenario and endorsed the WG's decision to present BRPs for both recruitment scenarios.

Weakness of analysis

Despite the WG's attempts, it was not possible to explain the recent long run of low recruitment values. There was thus considerable uncertainty as to whether the recent series of low recruitment was due to low SSB or whether some unknown factor had produced a reduction in productivity.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel accepted continued use of $F_{40\%}$ and $SSB_{40\%}$ proxies for F_{msy} and SSB_{msy}.

The analyses undertaken by the WG demonstrated that the F reference points were more robust to the possibility that there had been a productivity shift than the SSB reference points, as the value of Bmsy depends on whether or not there has been such a change. Based on the levels of recruitment that have been experienced over the last two decades, however, the reduced recruitment scenario is more likely than not to be the appropriate scenario on which management advice should be based.

The Panel noted that, when a Beverton and Holt stock-recruitment relationship was included in the ASAP model, and fitted internally within the model, very different values of MSY and Bmsy reference points were estimated.

The dilemma that the WG faced when determining the BRPs for the SNEMA yellowtail flounder stock was whether or not the stock had experienced a productivity shift. In making such a decision, the Panel suggested that it might be appropriate to employ a weight-of-evidence approach. The Panel suggested that the decision that the

stock had experienced a productivity shift might be justified if, for example, there had been:

1. A long run of years with observed above or below average annual recruitment.
2. A long period of positive or negative recruitment residuals from the predicted values of the stock-recruitment relationship that cannot be explained by re-specification of that stock-recruitment relationship.
3. Error in estimated model inputs such as total catch, abundance indices, or catch age/size composition can be ruled out as a cause of the above or below average/expected recruitment.
4. A plausible mechanism based on environmental or ecological conditions has been identified.

Based on the above list and consideration of the modelling results, the weight of evidence led to the Panel to conclude that, subjectively, there was about a 60% chance that a productivity shift best explains the recent series of low recruitment estimates. The further development and refinement of a list such as the above by a larger group of experienced fishery scientists would provide a valuable tool for future situations in which consideration is being given to significant departure from conventional hypotheses.

The fact that, with the exception of the estimate for one year, model estimates of age-1 recruitment since 1990 have been consistently very poor, supports the view that, because of some unknown factor, survival from spawning to recruitment has declined from that experienced prior to 1990. Without a mechanism to explain the decline in recruitment, however, and additional evidence, it would be inappropriate to conclude that there has been a productivity shift. The possibility that, in recent years, the stock has been held at low spawning stock biomass by some other factor, such as additional unrecognised catches, cannot be entirely dismissed. The Panel therefore considered it more likely than not that the fishery has moved into a low recruitment state, and therefore endorsed the decision to carry forward two scenarios when considering biological reference points and stock status, *i.e.*, a two-stanza stock-recruitment scenario in which recruitment for SSB values less than 4,300 mt reflect recent recruitment levels and recruitment for greater SSB values reflecting the recruitment levels experienced prior to 1990 and a recent-recruitment scenario in which recruitment levels reflect those for the period since 1990. The Panel also agreed that projections should be based on the recent recruitment scenario.

Does work provide a scientifically credible basis for fishery management advice?

Insufficient evidence exists to rule out either recruitment scenario. While the estimated value of $F_{40\%}$ was the same for both recruitment scenarios, the estimated value of $SSB_{40\%}$ and associated yield for the recent recruitment scenario were far smaller than the corresponding estimates for the two-stanza model, *i.e.*, only 13% of the magnitude of those values,

YF ToR 7. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from YF TOR-6).

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG employed the updated version of the existing VPA model, which included revised length-weight relationships and updated numbers at age, and extended the data to 2011, to assess stock status relative to existing BRPs. It then compared the estimates of SSB and F from the new base model with the BRPs that were determined using both the recent recruitment and two-stanza recruitment scenarios to evaluate the status of the stock.

Strength of analysis

Two scenarios were considered when assessing stock status using the new model, *i.e.*, a recent low-recruitment scenario and a two-stanza recruitment scenario, which reflects the historical recruitment that was experienced in the fishery. Although the former scenario is considered more likely than not, the evidence is insufficient to rule out either scenario.

Weakness of analysis

The overfished status of the stock, and determination of whether or not the stock is rebuilt, depends on which of the recent recruitment or two-stanza recruitment scenarios is correct. Although the Panel concluded that the former recruitment scenario was more likely than not, in which case the stock would be considered to be rebuilt despite its current low biomass, there is insufficient evidence to rule out either recruitment scenario.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

Based on the existing BRPs and using an updated version of the existing VPA model, it was concluded that the stock is overfished but overfishing is not occurring. In the previous GARM III assessment, it had been concluded that overfishing was occurring.

Based on the new model outputs and BRPs estimated using the two-stanza recruitment scenario, the stock is assessed as being overfished but recruitment overfishing is not occurring. In contrast, if the BRPs associated with the recent recruitment scenario are used to determine stock status, it is concluded that the stock is not overfished (and is "rebuilt") and overfishing is not occurring. The "rebuilt" status arises not from a marked increase in SSB, however, but from the reduction in the SSB reference point that results from the reduced productivity of the stock. There is thus considerable uncertainty as to whether or not the stock is rebuilt.

Despite the great difference in BRPs from the 2-stanza and Beverton-Holt models, the stock status was the same.

Does work provide a scientifically credible basis for fishery management advice?

While the results of the assessment indicate that overfishing is not occurring, it is not possible, given the evidence, to determine whether or not the stock is overfished. If low recruitment is the result of low SSB, then the stock is not yet rebuilt and remains overfished. If, however, low recruitment is the result of reduced productivity due to some unknown factor, the stock is rebuilt and is not overfished.

YF ToR 8. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and recruitment as a function of stock size).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

Was the ToR completed successfully?

Yes, in all but one item. Although the majority of the items listed in the TOR were successfully accomplished, the WG did not estimate the annual probability that the estimate of SSB fell below the threshold BRP.

Reason for acceptance/rejection

Short term projections were made under both the recent and two-stanza recruitment scenarios assuming that F was either 0, $F_{75\%msy}$, or F_{msy}. Estimates of SSB and yield (and their lower and upper 90% confidence limits) were reported. As F was assumed to be fixed at values less than or equal to the BRP, the probability of F exceeding the BRP was zero. The probability that the annual estimate of SSB fell below the BRP for biomass was not calculated. The major uncertainty in the assessment is whether the current productivity of the stock is better represented by the recent recruitment or two-stanza recruitment scenario. The WG commented (briefly) on the uncertainties of the assessment and sensitivity of the projections to various assumptions, and offered advice regarding which of the two recruitment scenarios was currently more likely. Factors that could influence the vulnerability of the stock were identified and discussed.

Strength of analysis

The projection method was suitable,

Weakness of analysis

Although the main source of uncertainty, *i.e.*, whether the current productivity of the stock was better represented by the recent recruitment or two-stanza recruitment scenario, was addressed, it would have been useful if more sources of uncertainty had been considered in the projections.

There would be value in considering the consequences of assuming that the recent recruitment scenario is true when current stock productivity is better represented by the two-stanza recruitment scenario, or vice versa, in terms of lost yield or impact on SSB.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

Projections were made using both the recent-recruitment scenario and the two-stanza scenario. Under the two stanza scenario, rebuilding is not expected by 2014 even with F = 0. Under the recent recruitment scenario, the stock is already rebuilt. In

this latter case, if it is now fished with a fishing mortality equal to $F_{40\%}$, SSB will continue to decline. It would be interesting to assess how many years would be required before, with a given level of fishing mortality, it is possible to determine which of the two recruitment scenarios is more likely to be correct.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The projections provide information regarding the potential yields and expected values of future SSB under the two recruitment scenarios that had been identified. It would have been useful to assess the consequences of incorrectly assuming that one recruitment scenario was true when in fact the other was actually true.

YF ToR 9. Review, evaluate and report on the status of research recommendations listed in most recent peer reviewed assessment and review panel reports. Identify new research recommendations.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG reviewed, evaluated and reported on the status of research recommendations that had been proposed earlier, and then identified new research recommendations (most of which were completed prior to the Review Meeting).

Strength of analysis

The WG provided a succinct but thorough comment on research recommendations from previous assessments Most of the previous research recommendations had been met either completely or to some degree, or were no longer considered applicable as a result of changes to model structure.

Many of the new research recommendations had been fully or partially addressed prior to the Review Meeting.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The research recommendations that had been proposed in earlier assessments were identified and reviewed. Comment was made on progress. The research recommendations that remained to be addressed were identified as:

- Explore whether sea sampling of otter trawl and scallop vessels, and sampling frequency of bottom trawl surveys, is adequate for such temporal resolution. Collect adequate numbers of commercial samples for determination of length and age composition.
- Explore the use of "windows" of biomass rather than the breakpoint to create the stanzas in the stock-recruitment relationship, as this might better address inconsistencies in building plans that might arise as the biomass grows from the lower to the higher stanza.

New research recommendations were listed. Although many of these had been fully or partly addressed prior to the meeting, further work was required on the following research recommendations:

- Update the length-weight parameters used to convert commercial landings (in weight) into numbers of fish. This could be accomplished by expanding existing data collection programs (*e.g.,* Cooperative Research, Industry Based Surveys, NEFSC port sampling) to collect individual fish weights while

collecting length and age data. This research recommendation is applicable to numerous species/stocks in the northeast, not just SNE/MA yellowtail flounder. (Partially completed in this assessment based on data available).

- The work on the influence of the cold pool and associated environmental parameters on yellowtail population dynamics has not been fully developed, and merits further research. (The WG explored the application of the cold pool index in this assessment by explicitly incorporating the cold pool index in the ASAP model. Further work will continue to explore the application of environmental data in the assessment).
- If the volume of commercial landings increases in the future, ensure that adequate samples of the landings are obtained for all market categories on at least a quarterly basis. (Quarterly resolution was not explored in this assessment for deriving fishery catch data).

It would have been useful if the outstanding research recommendations from earlier studies had been merged with the outstanding new recommendations and the items arranged in order of perceived priority.

The Review Panel identified a number of research areas that should also be considered. These are listed below.

Data

- Exploration of alternative catch history (a possible explanation of apparent productivity shift).
- Adjust NMFS survey for day/night effect using GAM or GLM.
- Explore inclusion of pre-1973 data in ASAP.
- Compile independent evidence for or against productivity shift (including trends in other commercial species in the same stock region).

Model

- Further explore stock-recruitment models that incorporate oceanographic processes.
- Explore why 2-stanza BRPs are so different from those produced using a model with a Beverton-Holt stock-recruitment relationship (a more conventional approach in many fisheries assessments in other locations), and determine which approach is more appropriate for this stock.
- Explore the uncertainty associated with the calibration factor for a change such as the change in trawl doors, by breaking the time series of survey indices into two sections, *i.e.,* preceding and following the change, then estimating catchabilities separately for each component, and calculating the ratio of those catchabilities. A term would be added to the objective function representing the likelihood of this ratio given the empirically determined calibration factor and its standard error, and assuming a normal probability distribution.
- Employ iterative reweighting of effective sample size based on fit to mean age rather than the McAllister and Ianelli (1997) approach.
- Undertake further research to investigate whether a model incorporating a stock-recruitment relationship, with the possibility of a change in parameter values in 1990, might be a viable alternative to the base model. Such a model might allow uncertainty associated with a possible productivity change to be carried through to the estimates of BRPs and the determination of stock status and avoid the need to present alternative recruitment scenarios.

5. Conclusions and recommendations

The SARC 54 Review Panel examined the 2012 stock assessments developed for the Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*). After considering the information relating to stock structure, the data that were available, and the details of the assessment for each species, and discussing areas of concern, the Panel accepted the base model that had been proposed by the WG for each assessment. The accepted base models, which had been developed using the statistical catch-at-age ASAP model, differed from those which had been applied in the previous assessments, *i.e.*, the 2009 TRAC assessment for Atlantic herring and the 2008 GARM assessment for yellowtail flounder.

The WG's decision to maintain the current stock definition for the Gulf of Maine/Georges Bank Atlantic herring was endorsed by the Panel, which recognised that this "stock" was a complex that comprised at least three "sub-stocks". In the case of this stock, the assumption of a constant natural mortality of $M=0.2$ year^{-1} in the ASAP model, which had been used in the previous 2009 assessment, had been replaced by an assumption that natural mortality declined with weight at age in a pattern similar to that described by Lorenzen (1996) and thus varied with both age and year, that the average value of natural mortality over age in the years up to 1995 was 0.3 year^{-1}, and that, since 1995, the average value of natural mortality over age had increased by 50%. Rather than fitting the model to total catch, as was done in the 2009 assessment, catches of fixed (US fixed gear and New Brunswick weir fishery) and mobile gears were used in the 2012 assessment. Maturity at age, which was constant in the 2009 assessment, varied with age in the new assessment.

The Panel was advised that, in the case of Atlantic herring, introduction of age and time dependent natural mortality, based on Lorenzen (1996), failed to improve the overall likelihood of the model relative to a model that employed a constant mortality over age and time. On this basis, the added complexity of the age and time dependent mortality is not justified. However, its inclusion is not likely to affect the results. A more important issue is whether or not natural mortality increased by 50% since 1995. This unusual increase in natural mortality was justified by increased annual consumption by predators of Atlantic herring, and by the fact that inclusion of the increased natural mortality reduced retrospective patterns. The very considerable body of data relating to dietary composition and predator abundance suggests that the trends in consumption are real, and that it is appropriate to increase natural mortality in the latter portion of the time series. Accordingly, after considering other results, diagnostic outputs, and output from various sensitivity runs, the Panel accepted the base ASAP assessment model for Atlantic herring as an appropriate model for developing management advice. While there is some uncertainty regarding the extent to which natural mortality increased, examination of the results obtained by applying a range of alternative percentage increases suggested that the value of 50% was plausible. It should be recognised, however, that the uncertainty associated with this percentage is not carried through into the uncertainty associated with estimates of spawning stock biomass (SSB), fishing mortality (F), or biological reference points (BRPs), or into projections.

As it is very likely that high levels of consumption of Atlantic herring will persist in the immediate future, the decision by the WG to employ the increased level of natural mortality when calculating BRPs was endorsed by the Panel. Likewise, the Panel concurred with the finding that the GoM/GB Atlantic herring stock complex

was not overfished and that overfishing was not occurring. Exploration of the sensitivity of this finding to a range of values of M and post-1995 increase in M, and to a range of values of steepness of the BH stock-recruitment relationship employed in the ASAP model for the Atlantic herring demonstrated that this finding was robust to such uncertainty.

The major uncertainties of the 2012 assessment for GoM/GB Atlantic herring are the extent to which natural mortality increased since 1995, whether the high level of consumption by predators is likely to persist in the longer term, the extent to which mixing of fish from different stocks across the boundaries of the Gulf of Maine/Georges Bank region has influenced the assessment, whether any of the various sub-stocks in the Gulf of Maine/Georges Bank stock complex are overfished or experiencing overfishing, and whether the abundance of the 2008 year class, which is currently estimated to be twice as great as the largest year class previously encountered (1994) and which is projected to make a considerable contribution to future catches, has been over-estimated.

The Panel endorsed the WG's conclusion that yellowtail flounder in the southern New England/Mid-Atlantic (SNEMA) region should be considered as a separate stock for management purposes. For this stock, the assumption of a constant natural mortality of 0.2 $year^{-1}$ in the VPA model, which had been used in the previous 2008 assessment, was replaced by an assumption that natural mortality declined with weight at age in a pattern similar to that described by Lorenzen (1996) and thus varied with both age and year, and that the average value of natural mortality over age was 0.3 $year^{-1}$. The data series used had also been revised considerably since the previous assessment. Discards for the 2012 assessment were estimated using spatial and temporal stratification of observer data, estimates of discard mortality were revised, and, from 1994, revised weight-length relationships were used to convert catches to numbers at age. Estimates of weights and maturity at age were also updated.

The WG is commended for the bridging analyses that they undertook to demonstrate the effect of each step of the transition from the previous VPA model for yellowtail flounder to the new ASAP model with its revised data.

The Panel did not support the WG's conclusion that estimates of abundance from daylight survey trawls were sufficiently similar to those from trawls at night that an aggregate index could be used and advised that improved precision would be gained by taking the time of day, *i.e.*, day or night, into account. The assumption in the new model that natural mortality is age and time dependent, and follows a pattern such as that described by Lorenzen (1996) introduces additional complexity to the model for yellowtail flounder. A simpler model that assumes constant natural mortality is to be preferred unless it can be demonstrated that the more complex model improves the likelihood significantly. Data weightings used in fitting the model, *e.g.*, doubling the larval CVs and adding 0.1 to survey CVs, appeared arbitrary.

After considering the results of the base assessment model and its diagnostic outputs, and the results and diagnostic output from various sensitivity runs, the Panel accepted the base ASAP assessment model for SNEMA yellowtail flounder as an appropriate model for developing management advice. A key result from this base model, however, was the marked decline in recruitment that had occurred since 1990. The WG proposed two scenarios to account for this. Firstly, the WG proposed a "two-stanza recruitment scenario" that assumed that recruitment is a function of SSB, with high but very variable recruitment at larger values of SSB and very low recruitment when SSB is less than about 4,300 mt. The WG's second proposal was a "recent

recruitment scenario", which assumed that, since 1990, annual recruitment is low because of environmental factors that are unrelated to SSB. The Panel agreed that, given the pattern of recruitment predicted by the base model, these two recruitment scenarios were appropriate alternatives. The Panel also recommended that, in the future, consideration might also be given to exploring whether a BH model fitted internally within the ASAP model, might provide a further alternative. By allowing for a change in the values of the parameters of the stock-recruitment relationship in 1990, such a model could be used to explore the extent to which steepness or asymptotic recruitment had changed.

The Panel accepted the WG's finding that variables indicative of the size and temperature of the cold pool influenced the strength of recruitment of yellowtail flounder, but concluded that these variables provided no explanation of the step-like decline in recruitment in 1990. Likewise, data relating to broader ecosystem and environmental change in the region provided no explanation for the sudden decline in recruitment.

Based on the weight of the evidence that had been presented, the Panel concluded that, while it was not possible to rule out either of the two recruitment scenarios, it was more likely than not that productivity of the yellowtail flounder stock had declined since 1990, *i.e.*, there was a qualitatively-assessed probability of about 60% that current recruitment will follow the recent recruitment scenario rather than the two-stanza scenario. Note that the value of 60% is based on the Panel's subjective opinion, and has no quantitative basis. Recognising that the evidence was insufficient to rule out either scenario, however, the Panel agreed with the WG that both recruitment scenarios should be considered when evaluating reference points, determining stock status, and considering the results of short term projections.

Results from the base assessment model for SNEMA yellowtail flounder indicate that, for either recruitment scenario, overfishing is not occurring. If recruitment follows the two-stanza scenario, however, the stock is not rebuilt, is overfished, and is unlikely to rebuild by 2014 even in the absence of fishing. On the other hand, if the productivity of the stock has been reduced by some unknown factor and, as the Panel considers is more likely than not, recruitment follows the recent recruitment scenario, the stock is rebuilt and not overfished. Note that the fact that the stock is considered to be rebuilt in this latter case is due to the reduction in the biomass reference point rather than increase in SSB.

Key uncertainties in the case of the SNEMA yellowtail flounder assessment include the identity of the factor or factors that have led to and would explain the decline in recruitment since 1990, the question of whether low recruitment is likely to persist, the possibility of error in catches and discards (which could be explored through development of alternative catch histories), and the possibility that the SNEMA yellowtail flounder stock is not entirely closed to interchange of fish or larvae across the boundary of the SNEMA stock area.

The unusual features of the stock assessments undertaken for SWA 54, *i.e.*, that, since 1995, natural mortality of Atlantic herring had increased markedly, and that, since 1990, there had been a decline in the recruitment of yellowtail flounder, highlight the fact that the assumptions of a constant level of natural mortality and a constant stock-recruitment that are typically used in stock assessment models are probably over-simplistic. With global climate change, it would not be unexpected that natural mortality, growth, weight at age, maturity at age, and stock-recruitment of other stocks are also experiencing change. In most cases, however, the change would be expected to be of a more gradual nature than the changes that appear to have

occurred in the cases of natural mortality of GoM/GB Atlantic herring or recruitment of SNEMA yellowtail flounder.

Rather than constructing separate models to describe alternative scenarios, as was done in the case of yellowtail flounder, there would be value in developing a model that allows exploration of hypotheses regarding the types of change in parameters such as natural mortality and stock-recruitment that might be expected. For example, in the case of the SNEMA yellowtail flounder, a BH model with a change in steepness or asymptotic level of recruitment following 1990 could be incorporated in the ASAP model, with the parameter describing the change in steepness or asymptotic recruitment being estimated when the model is fitted to the data. Such a model would provide greater ability to characterize the uncertainty associated with the possible change in the stock-recruitment relationship and allow determination of BRPs and stock status that take that uncertainty into account. By employing such a model, the need to consider two alternative recruitment scenarios, *i.e.,* low recruitment associated with either low SSB or reduced productivity, would be avoided.

A major concern of the current assessment for GoM/GB Atlantic herring is that, other than through the use of sensitivity runs, the uncertainty associated with the estimate of the post-1995 percentage increase in natural mortality is not carried through into the assessments of BRPs and determination of stock status. Further refinement of the time series of annual estimates of consumption of Atlantic herring, and particularly characterization of the uncertainty of the annual estimates, and continued development of an assessment model that incorporates this source of natural mortality are recommended, as this would facilitate assessment of the uncertainty associated with the consumption of Atlantic herring by predators.

Inadequate detail was provided in the assessment reports of the statistical designs of the bottom trawl surveys and the protocols used to obtain biological samples for these surveys and for the commercial catch data. Descriptions of the methods used to expand the collected data to estimates of catch at age or to combine the survey data from different samples to form a single index of abundance or overall age composition for each survey were either missing or inadequate.

The SARC meeting facilities and logistical support for the meeting were excellent. Information Technology support during the meeting ensured that access was available to the file server. The FTP site provided convenient access to electronic copies of background papers and assessment documents. The assistance provided by the rapporteurs, who recorded details of the discussions during the review, was greatly appreciated. It was pleasing to find that access to the meeting was available through WebEx, allowing a broader audience to participate.

6. References

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. *Fishery Bulletin*, **82**: 898-903.

McAllister, M. K., and Ianelli, J. N. 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Science*, **54**: 284–300.

Mertz, G., and Myers, R. A. 1996. Influence of fecundity on recruitment variability of marine fish. *Can. J. Fish. Aquat. Sci.*, **53**: 1618 - 1625.

Appendix 1: Bibliography of all material provided

Yellowtail Flounder

Background Papers

1 Wood, A. D., Cadrin, S. X., Alade, L. A., Martins, D., Moser, J., & Westwood, A. D. 2012. Mortality and movement of yellowtail flounder, *Limanda ferruginea*, tagged off New England.

2 McBride, R., Press, Y., & Wuenschel, M. 2012. Classifying female yellowtail flounder maturity: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method.

3 McElroy, W. D., Towle, E. K., Press, Y. K., McBride, R. S., & Wuenschel, M. J.. 2012. Comparison of fecundity among stocks of female yellowtail flounder, *Limanda ferruginea*.

4 McElroy, W. D., Press, Y. K., & Wuenschel, M. J. 2012. Reproductive effort as a predictor of the natural mortality rate for southern New England yellowtail flounder: the Gunderson method.

5 Alade, L., & Cadrin, S, 2012. A review of yellowtail flounder stock structure off New England.

6 Alade, L., Hart, D., & Legault, C. 2012. Influence of spatial stratification on discard estimation for the Southern New England Mid-Atlantic yellowtail flounder.

7 Alade. L. 2012. Southern New England-Mid Atlantic yellowtail length-weight relationship.

8 Barkley, A., & Cadrin, S. 2012. Results of the application of reflex action mortality predictors onboard commercial fishing vessels.

9 Gervelis, B., Burchard, K., & Hoey, J. 2011. A comparison of discard estimates of yellowtail flounder using study fleet self-reported data and NEFOP observer data.

10 Alade, L. A. 2012. Southern New England Mid-Atlantic yellowtail flounder commercial landings.

11 Hare, J., Richardson, D., & Mountain, D. 2012. Indices of the Mid-Atlantic cold pool and relationship to Southern New England/ Mid-Atlantic yellowtail flounder recruitment.

12 Richardson, D. E., Walsh, H., & Hare, J. 2012. Southern New England yellowtail flounder larval data.

13 Anonymous. 2012. Revised Spatial Stratification for estimating discards for SNEMA yellowtail flounder.

Previous Assessments

14 GARM. 2002. Southern New England yellowtail flounder.

15 GARM. 2005. Southern New England-Mid Atlantic yellowtail flounder.

16 GARM. 2008. Southern New England/Mid Atlantic yellowtail flounder.

17 Anonymous. 2003. Yellowtail Flounder Stock Structure. 36th SAW Advisory.

18 Northeast Fisheries Science Center. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-06; 453 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

19 Anonymous. 2003. Southern New England - Mid Atlantic Yellowtail

Flounder. 36th SARC Consensus Summary.

Working Papers

1 Southern Demersal Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Mid-Atlantic (*Limanda ferruginea*) stock assessment updated through 2011. A report of the Southern Demersal Working Group for SARC 54, June 5th – 9th, 2012. Northeast Fisheries Science Center. Woods Hole, MA. SAW/SARC 54.

2 Anonymous. 2012. Southern New England Yellowtail Flounder Assessment Summary for 2012. SAW/SARC 54.

Atlantic Herring

Background Papers

1 Shepherd, G., Cieri, M., Power, M., & Overholtz, W. 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic Herring Stock Assessment Update. TRAC Reference Document 2009/04.

2 Transboundary Resources Assessment Committee. 2009. Gulf Of Maine-Georges Bank Herring Stock Complex. Status Report 2009/04.

3 Transboundary Resource Assessment Committee. 2006. Gulf of Maine-Georges Bank Herring Stock Complex. Status Report 2006/01.

4 Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting held 10-14 February 2003. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2003/014.

5 O'Boyle, R., & Overholtz, W. 2006. Proceedings of the Transboundary Resources Assessment Committee (TRAC). Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring 2 – 5 May 2006 Woods Hole, Massachusetts.

6 O'Brien, L., & T. Worcester, T. 2008. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Gulf of Maine/Georges Bank Herring, Eastern Georges Bank Cod and Haddock, Georges Bank Yellowtail Flounder. Report of Meeting held 8 - 11 June 2009 St. Andrews Biological Station, St. Andrews, New Brunswick, Canada.

7 Jech, J. M., & Stroman, F. 2012. Aggregative patterns of pre-spawning Atlantic herring on Georges Bank from 1999-2010. *Aquat. Living Resour.*, **25**: 1–14.

8 Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., & Nero, R. W. Fish population and behavior revealed by instantaneous continental shelf–scale imaging. *Science*, **311**: 660-663.

9 Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., Nero, R. W., & Jech, J. M. 2009. Supporting online material for "Critical population density triggers rapid formation of vast oceanic fish shoals". *Science*, **323**: 1734.

10 Link. J. S., & Almeida, F. P. 2000. An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Technical Memorandum NMFS-NE-159. 60 pp.

11 Kanwit, J. K., & Libby, D. A. 2009. Seasonal movements of Atlantic herring (*Clupea harengus*): results from a four year tagging study

conducted in the Gulf of Maine and Southern New England. *J. Northw. Atl. Fish. Sci.*, **40**: 29–39.

12 Guan, W., Cao, J., Chen, Y., & Matthew Cieri, M. 2012. A simulation study to evaluate impacts of spatial structure of Atlantic herring fishery on retrospective errors in stock assessment.

13 Reid, R. N., Cargnelli, L. M., Griesbach, S. J., Packer, D. B., Johnson, D. L., Zetlin, C. A., Morse, W. W. & Berrien, P. L. 1999. Essential Fish Habitat Source Document: Atlantic Herring, *Clupea harengus*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-126. 48 pp.

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

15 Miller, T. J., Richardson, D. E., & Hare, J. A. In prep. Maximum likelihood estimation of larval production indices from length frequency and growth information collected on ichthyoplankton surveys.

16 Cao, J., Guan, W., Chen, Y., & Cieri, M. In prep. Evaluating factors influencing retrospective errors in estimating stock biomass for Atlantic herring *Clupea harengus*.

Working Papers

1 Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

1 Appendices 3-6

Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. Appendices 3-6. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

2 Anonymous. 2012. Atlantic Herring Assessment Summary for 2012.

Appendix 2: Copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Norm Hall

External Independent Peer Review by the Center for Independent Experts

54th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Southern New England yellowtail flounder and Atlantic herring.

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 Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*). Yellowtail flounder is a demersal flatfish distributed from Labrador to Chesapeake Bay generally at depths between 40 and 70 m (20 to 40 fathoms). Off the U.S. coast, three stocks are considered for management purposes: Cape Cod/Gulf of Maine, Georges Bank, and Southern New England/ Mid-Atlantic. The principal fishing gear used to catch yellowtail flounder is the otter trawl. The last peer reviewed assessment of Southern New England yellowtail flounder was in 2008 as part of the GARM III. Atlantic herring is a pelagic fish that is widely distributed in continental shelf waters of the Northeast Atlantic, from Labrador to Cape Hatteras. Important commercial fisheries for juvenile herring (ages 1 to 3) exist along the coasts of Maine and New Brunswick. Development of large-scale fisheries for adult herring is comparatively recent, primarily occurring in the western Gulf of Maine, on Georges Bank, and on the Scotian Shelf. The last peer reviewed assessment of Atlantic herring was in 2009 as part of the TRAC. Yellowtail flounder and Atlantic herring are managed by the New England Fishery Management Council. Results of the 2012 peer 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 stock assessment Terms of Reference (ToRs), which are carried out by the SAW Working Groups, are attached

in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is described in **Annex 4.**

The SARC 54 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 Mid-Atlantic 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 of the stock assessments that are provided, and this review should be in accordance with this SoW and stock assessment ToRs herein. CIE reviewers shall have working knowledge and recent experience in fish stock assessments. For yellowtail, familiarity with forward projecting models and estimation is desirable. For herring, familiarity with pelagic fish and acoustic surveys is desirable. For both stocks, experience with time- and sex-specific natural mortality rate is desirable.

In general, CIE reviewers for SARCs shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise shall include statistical catch-at-age, state-space and index methods. Reviewers shall also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers shall 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 BRPs.

Each CIE reviewer's duties shall not exceed a maximum of 15 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 15 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 and Date of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during June 5-9, 2012 (Tuesday through Saturday).

Charge to SARC panel: During the SARC meeting, the panel is to determine and write down whether each stock assessment Term of Reference of the SAW (see **Annex 2**) was or was not completed successfully. 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 SARC chair shall identify or facilitate agreement among the reviewers for each stock assessment Term of Reference of the SAW.

If the panel rejects any of the current Biological Reference Points (BRP) or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies 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 stock assessment 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/.
http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.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 (i.e., working papers) 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 with 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 stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the 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 stock assessment 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 stock assessment Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For each 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 stock assessment 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 or BRP proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist. 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 request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

3. After the Open meeting
(SARC CIE reviewers)

Each CIE reviewer shall prepare an Independent CIE Report (see **Annex 1**). This report should explain whether each stock assessment 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 stock assessment 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 stock assessment 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, with the assistance from the CIE reviewers, will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each stock assessment 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 stock assessment 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 Points (BRP) or 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 stock assessment ToR listed 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 5-9, 2012 (Tuesday through Saturday).
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than June 25, 2012, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and 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 assessment 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.

30 April 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
22 May 2012	NMFS Project Contact will attempt to provide CIE Reviewers the pre-review documents by this date
5-9 June 2012	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
9 June 2012	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
25 June 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
26 June 2012	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
29 June 2012	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)

9 July 2012	CIE submits CIE independent peer review reports to the COTR
16 July 2012	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

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

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

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

Modifications to the Statement of Work: Requests to modify this SoW must be 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 stock assessment ToR listed 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)

Dr. William Karp, Acting NEFSC Science Director
National Marine Fisheries Service, NOAA
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
Bill.Karp@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: Stock Assessment Terms of Reference for SAW/SARC54 (June 5-9, 2012) (to be carried out by SAW Working Groups) (file vers.: 10/21/11)

A. Atlantic herring

1. Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

3. Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

4. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

6. Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

7. 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}, $B_{THRESHOLD}$, F_{MSY} and 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.

8. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).
    a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).

9. Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

10. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).
    a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

11. For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

B. SNE/Mid-Atlantic Yellowtail Flounder

1. Estimate landings and discards by gear type and where possible by fleet, from all sources. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of commercial or recreational LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

3. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

4. 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 a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

5. Investigate causes of annual recruitment variability, particularly the effect of temperature. If possible, integrate the results into the stock assessment (TOR-4).

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}, $B_{THRESHOLD}$, F_{MSY} and 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 with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).
    a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
    b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-6).

8. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).
    a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a

sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and recruitment as a function of stock size).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

9. Review, evaluate and report on the status of research recommendations listed in most recent peer reviewed assessment and review panel reports. Identify new research recommendations.

Appendix to the SAW Assessment TORs:

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

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

> *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" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

> *"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)

Rules of Engagement among members of a SAW Assessment Working Group:

> Anyone participating in SAW assessment working group meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

(~:\sarc\sarc54…\TORs\DRAFT SAW-SARC-54_[date].doc)

(END OF ANNEX 2)

Annex 3: Draft Agenda

54th Northeast Regional Stock Assessment Workshop (SAW 54)
Stock Assessment Review Committee (SARC) Meeting

June 5-9, 2012

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

Draft AGENDA* (version: 14 March 2012)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR
Tuesday, June 5			
1 – 1:30 PM			
Welcome	**James Weinberg**, SAW Chair		
Introduction	**Robert O'Boyle**, SARC Chair		
Agenda			
Conduct of Meeting			
1:30 – 3:30	Assessment Presentation (A. Herring)		
	Jon Deroba, others	**TBD**	**TBD**
3:30 – 3:45	Break		
3:45 – 6	Assessment Presentation (A. Herring)		
	Jon Deroba, others	**TBD**	**TBD**
Wednesday, June 6			
9 – 11:45	SARC Discussion w/ presenters (A. Herring)		
	Robert O'Boyle, SARC Chair		**TBD**
11:45 – 1	Lunch		
1:00 – 3:15	Assessment Presentation (B. SNE YT)		
	Larry Alade	**TBD**	**TBD**
3:15 – 3:30	Break		
3:30 – 5:30	SARC Discussion w/ presenters (B. SNE YT)		
	Robert O'Boyle, SARC Chair		**TBD**

Thursday, June 7

9 - 11	Revisit w/ presenters (A. herring) **Robert O'Boyle**, SARC Chair	**TBD**
11 – 11:15	Break	
11:15 – 12:30	Revisit w/ presenters (B. SNE YT) **Robert O'Boyle**, SARC Chair	**TBD**
12:30 – 1:45	Lunch	
1:45 – 2:15	(cont.) Revisit w/ presenters (B. SNE YT) **Robert O'Boyle**, SARC Chair	**TBD**
2:15 -2:30	Break	
2:30 – 5:30	Review/edit Assessment Summary Report (A. herring) **Robert O'Boyle**, SARC Chair	**TBD**

Friday, June 8

9 - 12	Review/edit Assessment Summary Report (B. SNE YT) **Robert O'Boyle**, SARC Chair	**TBD**
12 – 1:15	**Lunch**	
1:15 – 5	SARC Report writing. (closed meeting)	

Saturday, June 9

9:00 - 3 PM (cont.) 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 Working Group was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

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

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

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

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

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

Appendix 3: Panel membership and Key Participants

SARC Meeting for SAW 54
Southern New England yellowtail flounder and Atlantic Herring

Appointee	Function	Affiliation
Independent Review Panel		
Robert O'Boyle	Chair and Reviewer	NE FMC SSC
Chris Francis	Independent Reviewer	CIE
Neil Klaer	Independent Reviewer	CIE
Norman Hall	Independent Reviewer	CIE
Presenters		
John Deroba	Lead Analyst, Atlantic herring	
Larry Alade	Lead Analyst, yellowtail flounder	
Coordination		
Jim Weinberg	SAW 54 Chairman	NEFSC
Anne O'Brien	Administrative Support	NEFSC
Paul Rago	Chief, Population Dynamics Branch	NEFSC
Rapporteurs		
T. Chute	Rapporteur	
I. Blaycock	Rapporteur	

Among those present at the meeting were Larry Jacobsen, Jason Link, Mark Terciero, Chris Legault, Liz Brooks, and Matthew Cieri. Numerous other individuals, who were present or viewing the meeting via WebEx, participated and offered valuable comment during the review meeting.