Independent Peer Review of Benchmark Stock Assessments for Gulf of Maine Haddock and Sea Scallops

SARC 59 Meeting, July 15 - 18, 2014
Northeast Fisheries Science Center
Woods Hole, Massachusetts

Prepared for:
Center for Independent Experts

Prepared by:
Vivian Haist
1262 Marina Way,
Nanoose Bay, BC
Canada V9P 9C1

August 20, 2014

1. EXECUTIVE SUMMARY

This report presents results of an independent peer review of the Northeast Regional Stock Assessment Review benchmark stock assessments of Gulf of Maine haddock and sea scallops (SARC 59), conducted for the Center for Independent Experts. The primary activity of the review was participation in the July 15 - 18, 2014 SARC 59 Panel in Woods Hole, Massachusetts.

The SARC 59 review process was thorough, effective, and resulted in a comprehensive review of the two stock assessments. The assessments had been developed by Stock Assessment Working Groups (SAWs), who were thorough in developing and evaluating the assessment models and selecting a single model for the assessments, which greatly facilitated the work of the Panel. The SARC 59 Panel reached consensus on all assessment Terms of Reference (ToR).

SARC 59 reviewed a benchmark stock assessment for the Gulf of Maine Haddock resource. The data sources for the assessment were adequate and used correctly. The catch history is relatively well documented, and the Northeast Fisheries Science Center bottom trawl surveys provide a solid basis for stock reconstruction and estimating abundance trends. A statistical catch-age model (ASAP), which is appropriate for the available data, was used for the first time. Numerous sensitivity analyses were conducted to investigate alternative data and model assumptions, and the base model selected by the SAW for determining stock status was appropriate. Biological reference points (BRPs), based on F_{MSY} and B_{MSY} proxies, were updated using results from the new assessment. The SAW conclusion that the stock is not overfished, and overfishing is not occurring is consistent with the analyses presented. The Gulf of Maine haddock stock assessment provides a scientifically credible basis for developing fishery management advice.

The US sea scallop stock assessment was based on separate models for two Georges Bank areas (open and closed) and for a single Mid-Atlantic area. For the purpose of stock status determination, results from the three areas are combined, while stock projections are conducted for 17 subdivisions of the combined area. This approach for stock status determination and stock projections is consistent with the geographical scales used in fisheries management.

The data sources for the sea scallop assessment were adequate and used correctly. The catch is almost entirely from the commercial fishery and is well documented, though there is uncertainty in the magnitude of dredge fishery-induced incidental mortality. A long-term survey (dredge) provides solid information on stock trends and a drop-video survey provides estimates of near-absolute abundance, resulting in relatively high precision in the stock reconstructions. A length-based statistical model (CASA), developed explicitly for the sea scallop assessment, is appropriate for the available data. Sensitivity analyses were conducted to investigate alternative data and model assumptions, and the model selected for determining stock status was appropriate. Biological reference points, based on F_{MSY} and B_{MSY}, were updated using results from the new benchmark assessment. The SAW conclusion that the stock is not overfished, and overfishing is not occurring is consistent with the analyses presented. The sea scallop assessment provides a scientifically credible basis for developing fishery management advice.

2. BACKGROUND

This document reports on an independent peer review of the Northeast Regional Stock Assessment Review Committee (SARC 59) for benchmark stock assessments of Gulf of Maine haddock and sea scallops, conducted for the Center for Independent Experts. The primary activity of the review was participation in the July 15 - 18, 2014 SARC 59 Panel in Woods Hole, Massachusetts.

The CIE *Statement of Work* (Appendix 3) defines the scope of this review. In addition to participation in the SARC 59 Panel, the *Statement of Work* requests a review of draft assessment documents and other pertinent background materials prior to the review meeting and preparation of this report summarizing review findings relative to the terms of reference for the review.

3. DESCRIPTION OF REVIEW ACTIVITIES

The activities undertaken for this review include: 1) review and assimilation of background material and reports provided by the NMFS Project Contact prior to the SARC 59 Panel meeting; 2) conducting some additional model runs using the haddock assessment model; 3) participation in the SARC 59 Panel review meeting and contribution to the Panel Summary report; and 4) preparation of this report.

The materials provided to prepare for the SARC Panel meeting included; draft stock assessment documents and assessment summaries for Gulf of Maine haddock and sea scallops, previous haddock and sea scallop assessment documents, and other background documents pertaining to the data and methods used for the SARC 59 benchmark assessments (Appendix 2). Additionally, an executable of the stock assessment model (ASAP) and data file used for the Gulf of Maine haddock data assessment were provided.

The primary focus for the SARC 59 Panel members (Appendix 1) during the July 15 - 28, 2014, meeting included:

- Determining whether data were adequate and used properly, the analyses and models were carried out correctly, and conclusions are reasonable and consistent with the analyses presented.
- Determining whether each stock assessment Term of Reference was completed successfully.
- Determining whether the scientific assessments are adequate to serve as a basis for developing fishery management advice.
- Reviewing and agreeing text for the Gulf of Maine and sea scallop stock assessment summary documents.

A SARC 59 Panel report, summarizing the Panel's views and conclusions relative to the meeting ToR, was prepared by Panel members during and after the meeting. This report, prepared for the CIE, reflects my own views, which are consistent with the Panel's conclusions on all substantive issues.

4. SUMMARY OF FINDINGS

4.1 OVERVIEW

The SARC 59 review process for Gulf of Maine haddock and sea scallops was thorough, effective, and resulted in a comprehensive review of the two stock assessments. The assessments had been developed by Stock Assessment Working Groups (SAWs): through a series of six meetings for the haddock assessment, which used a statistical catch-age model (ASAP) for the first time and through a single meeting for the sea scallop assessment which continued to use a model (CASA) that had been explicitly developed for the sea scallop resource. The SAWs had done a thorough job in developing and evaluating the assessment models, selecting a "base case" model for the assessment, and reporting the results. This greatly facilitated the work of the SARC 59 Panel.

The SARC 59 Panel made only minor requests of the assessment scientists for additional background and supporting documentation. The Panel did not request any additional model runs or suggest changes to the SAW agreed stock assessments. The Panel reached consensus on all assessment Terms of Reference (ToR), and concluded that both the Gulf of Maine haddock and sea scallop assessments provide a scientifically credible basis for developing fishery management advice.

4.2 GULF OF MAINE HADDOCK

SARC 59 reviewed a benchmark stock assessment for the Gulf of Maine (GoM) haddock resource. The data sources for the assessment were adequate and used correctly. The catch history is relatively well documented, and the Northeast Fisheries Science Centre (NEFSC) bottom trawl surveys provide a solid basis for stock reconstruction and estimating abundance trends. A statistical catch-age model (ASAP), which is appropriate for the available data, was used for the first time. Numerous sensitivity analyses were conducted to investigate alternative data and model assumptions, and the base model selected by the SAW for determining stock status was appropriate.

Biological reference points (BRPs), based on F_{MSY} and B_{MSY} proxies, were updated using results from the new assessment. The SAW conclusion that the stock is not overfished, and overfishing is not occurring is consistent with the analyses presented.

Some suggestions about alternative approaches and assumptions of the stock assessment modelling that may improve future assessments are suggested below, but these are unlikely to change the assessment of stock status.

Findings relative to each of the Terms of Reference (ToR) follow:

1. *Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance.*

This ToR was met.

The assessment team did a very thorough job completing this ToR. The assessment document contains a comprehensive discussion of all sources of landings and discards and problems associated with them. Analyses of trends in the spatial and temporal patterns in catch are provided.

Historically, commercial landings have dominated the GoM haddock catch but over the past decade recreational fishing has become more important and these landings are now of similar magnitude to the commercial fishery. Discard rates have generally been low in the commercial fishery (<<10%) but are sometimes much higher in the recreational fishery (>50%), in particular when strong year-classes recruit.

Uncertainty in commercial landing arises from a number of sources, and allocating haddock catch to the correct stock may introduce bias in catch estimates. One study that compared vessel trip report data (VTR) with vessel monitoring system data (VMS) estimated that misreporting on GoM haddock could result in as much as 20% underestimation in the landings, however that value is considered an upper bound on the potential error in landings estimates.

The assessment document contains substantial discussion on whether it is better to use landings or survey based length-weight relationships to convert landings in weight to landings in numbers, for use in the assessment model. I assume this is for the VPA analysis, not for the ASAP modelling. For the ASAP modelling, an alternate and preferable approach would be to fit the biomass of the catch components directly in the model. That is, the model's predicted catch in numbers-at-age is converted to catch in weight-at-age based on annual input weight-at-age data. This would avoid the need for length-weight relationships and directly adjusts for changes in mean weights at age. The document states that recreational catch in numbers is converted to catch in weight for the ASAP analyses, so it is likely that this suggested approach had also been used for the commercial landings. Note, however, that if the recreational and commercial fisheries are treated separately (i.e. as two "fleets") it would be possible to fit to recreational catch in numbers and commercial catch in biomass, thereby maintaining their natural units.

Age information is not available for the recreational fishery or for the commercial fishery discards, so their age structures are estimated using age-length keys (ALK) from the commercial fishery and from the surveys. This has the potential to bias age-composition estimates for those fishery components, in particular if they have different selectivity than that of the "borrowed" ALK. Given non-commercial landings are becoming a large component of the total catch, this potential for bias should be investigated. Sampling for ageing structures in the recreational fishery is likely warranted.

The assessment scientists went to considerable effort to supplement the observer-collected discard length frequency data (LF) with survey LFs for years where the observer data was considered inadequate. Because the assessment model (ASAP) combines all catch (commercial and recreational catch and discards) into a single "fleet", it is necessary to have age-composition data for all catch components. However, if each of the four catch components were treated separately in the assessment model, it would not be necessary to impute age compositions for years where sampling data are inadequate or not available.

Discard mortality rates are assumed to be 100% for the commercial fishery and 50% for the recreational fishery. There is no direct information to support these estimates, so they are highly uncertain.

Landings-per-unit-effort (LPUE) analyses were conducted for both the commercial and recreational fisheries using GLM models to estimate standardized annual indices. The assessment team determined that these indices were unlikely to reflect changes in abundance, in part because changes in management and other factors have affected selectivity, so they are not useful for use in the stock assessment. This conclusion is reasonable given the plethora of management changes that have occurred over the history of the fisheries.

2. *Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). If available, consider whether tagging information could be used in estimation of stock size or exploitation rate. Characterize the uncertainty and any bias in these sources of data.*

This ToR was met. The use of tagging data for estimating abundance was not explicitly dealt with, however the discussion about using the available tagging data to estimate movement rates (ToR 3) made it clear that the design of the haddock tagging program was inappropriate for estimating abundance.

The main survey data used in the GoM haddock assessment are the NEFSC spring and fall bottom trawl surveys, which have been conducted annually since 1968 and 1963, respectively. These surveys have used 3 different vessels and 3 different door configurations. Consistent survey time series are generated by converting the data to "Albatross IV/Polyvalent door" equivalents based on calibration coefficients estimated from calibration experiments. The largest change in the survey time series occurred when a new research vessel (FSV Henry B. Bigelow) and new survey protocols were introduced in 2009. Calibration coefficients for the Bigelow surveys are length-based, unlike the other calibration coefficients which are independent of length.

The conversion of all survey data to the "Albatross IV/Polyvalent door" equivalent hides the additional uncertainty associated with the vessel- and door-specific conversion factors. This uncertainty (and potential bias) would be best accounted for by modelling the relative gear and door effects directly in the stock assessment model and fitting to the actual observations (see ToR 4).

The assessment team investigated day/night differences in catch rates for the NEFSC surveys and found higher rates for tows conducted during the day. The differences can be substantial and appear largest when survey abundance peaks (strong year-classes are recruiting). Because abundance trends were similar between the day and night tows and there would be a large reduction in the amount of data if only day or night tows were used, the SAW decided to use all the data in developing abundance indices. While this is a reasonable conclusion, I think the day/night differences warrant further investigation. One potential approach is to use GLMs on the survey data to develop abundance indices that account for the day/night differences.

Two additional surveys have been conducted in state waters; the Massachusetts Department of Marine Fisheries bottom trawl survey (MADMF) and the Maine-New Hampshire (MENH) inshore groundfish survey. These surveys encompass only a small component of the GoM haddock stock, and capture primarily juvenile fish. They were, appropriately, not included in the haddock assessment as they are unlikely to capture trends in abundance of the entire stock.

3. *Evaluate the hypothesis that haddock migration from Georges Bank influences dynamics of GOM stock. Consider role of potential causal factors such as density dependence and environmental conditions.*

This ToR was met to the extent possible given the available data.

A number of approaches were followed to investigate the hypothesis that Georges Bank (GB) haddock migrate to the GoM. These included: a literature review focussed on historic haddock tagging studies; examination of the degree of recruitment synchrony between GB and GoM haddock; year-class tracking in survey data and the GoM haddock assessment diagnostics; re-analysis of tagging data from the Northeast Consortium Cooperative Haddock Tagging Program (NCCHT); and explicit modelling of movement between GB and GoM in a stock reconstruction model (SCAA).

An investigation of the potential impact of "spillover" of GB haddock to the GoB stock was explored using the SCAA model. The model was fitted to same data as used in the GoM stock assessment (see ToR 4) and estimates of GB haddock numbers-at-age came from the 2012 stock assessment. The SCAA model was used to explore alternative hypotheses of movement between GB and GoM, including both permanent and temporary movement assumptions. The movement runs assumed that movement from GB to GoM occurs at a constant rate when GB biomass is above a fixed critical level, with all fish aged 2 and older (sensitivities for ages 1++ and 3++) moving at that rate. Results supported movement from GB to GoM, but at low rates (<0.8%). This is an interesting and useful approach for investigating the potential effect of "spillover" of haddock from GB to GoM, though results from this modelling exercise are not definitive in terms of movement rate estimates. That is, only a single movement parameter was estimated (appropriately, given the limited data to support movement estimation), and if there is movement from GB to GoM the process is likely much more complex than that modelled.

Overall, results from the various approaches were not definitive. Haddock movement between GB and GoM clearly occurs, however the magnitude of movement and whether it is constant or in response to density is unclear. Analyses of the HCCHT tagging program data are not useful to inform this ToR because the program was not designed to address movement at the stock level. Although there is some recruitment synchrony between GB and GoM, that does not imply movement. Given the GB stock is almost an order of magnitude larger than the GoM stock, any significant amount of movement from GB to the GoM should be apparent in the assessment model diagnostics (but they are not). So, while it can be concluded that the rate of movement of GB haddock to the GoM is relatively small, the absolute magnitude of that movement is unclear.

4. *Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.*

This ToR was met.

A statistical catch-age model (ASAP) was used to reconstruct the GoM haddock stock and estimate recruitment, stock biomass and fishing mortality rates. The previous benchmark assessment had used a VPA catch-age model which was updated to provide a bridge to the new approach. For the ASAP assessment, a single fleet representing commercial and recreational catch and discards was modeled. The model was fitted to the NEFSC spring and fall bottom trawl survey data. An alternative statistical catch-age model (SCAA) was used to explore assumptions about movement between the GoM and GB haddock stocks.

This was the first GoM haddock assessment to use the ASAP model, and many runs were conducted to investigate alternative approaches and assumptions to find a suitable base case. The primary model components explored to develop the base case were data weighting, including weighting of at-age data and total catch, and constraints on recruitment deviations to restrict the magnitude of apparently strong recent year-classes. The basis for selecting a final model was appropriate, though it is surprising that the final weighting for the fishery at-age data is higher than that of the survey at-age data. The single "fleet" fishery age composition data are comprised of both commercial and recreational fishery discards and landings, and there have been many changes in fisheries regulations that would affect the age-selectivity of the fisheries, whereas the trawl surveys should have relatively stable age-selectivity (other than the recent change to the RV Bigelow). Hence, I would have expected greater consistency in the survey age-composition data than in the fishery age-composition data. However, the weighting approach used (iterative re-weighting for effective sample size) should ensure the data weights are consistent with the

ability of the model to fit the various data sources so there is no reason to believe the final weightings are inappropriate.

The ASAP assessment assumed a constant natural mortality rate of 0.2, consistent with previous GoM assessments. This value was justified based on the maximum observed age in catch and survey data, and equations that use the maximum age (i.e. Hoenig 1983; Hewitt and Hoenig, 2005). Those equations, however, provide estimates of total mortality (*Z*), not natural mortality (*M*).

A profile likelihood on *M*, conducted with the original ASAP Base run (ASAP_BASE), indicated an improvement in fit of 6 negative log-likelihood units when *M* was fixed at 0.1 rather than 0.2. Runs with lower *M* were not explored further in the assessment, although a decrease of 2 negative log-likelihood units or greater is considered significant with either a likelihood ratio test or the AIC criterion. Using the final model run proposed as the basis for the assessment (ASAP_final_temp10) and profiling on *M*, I found a decrease of 18 negative log-likelihood units when *M* was changed from 0.2 to 0.1, and a preference for even lower values of *M*. When a particularly large at-age residual was removed from the data (plus group for the 1994 fall trawl survey) the model fit improved by 30 negative log-likelihood units with an *M* of 0.1 rather than 0.2.

The model clearly shows a preference for lower *M*, and it appears that the trawl survey data provide the basis for this preference (rather than aliasing for some other model mis-specification). Survey-based estimates of year- and age-specific *Z* for ages assumed to be fully selected indicate fairly low rates (<0.25) in recent years. Though the SARC 59 Panel felt there was reasonable support for a lower M, they did not think it would be possible to fully explore this assumption in the time available. *M* is likely age-dependent and potentially variable over time, and estimation of *M* will be confounded with survey selectivity assumptions. Future GoM haddock assessments should evaluate alternative assumptions about *M*. In particular, the George Bank multi-species VPA may be informative about age and time-varying *M*.

The final ASAP model selected for the assessment models a single fleet representing commercial and recreational catch and discards. One of the sensitivity runs had separated the commercial and recreational catch and this run indicated different age-selectivity for the two catch components. Separation of the commercial and recreational catch, and possibly of landings and discards as well, would likely be a better approach for future GoM haddock assessments. Each of these catch components has a distinct selectivity and separating them would allow this to be modelled explicitly. Also, if each catch component was modelled separately missing data would not need to be imputed.

The assessment model assumes relatively large errors in the catch estimates on the basis that there is uncertainty in the magnitude of the catches. However, errors in the catch data are unlikely to be independent and identically distributed (IID, as assumed in the model), rather they are likely correlated and potentially biased. For example, the document suggests that gutted-whole weight conversion factors may be incorrect and there may be potential under-reporting of commercial landings. These types of factors are likely to lead to bias rather than IID errors, so it is unlikely that the model can correctly estimate the true catch. An alternate approach would be to develop alternative catch streams, based on potential biases in the catch estimates, and investigate the sensitivity of the assessment to those uncertainties.

The effects of changes in the vessels and doors used in the NEFSC bottom trawl survey on age- or size-selectivity and catchability is dealt with by adjusting the data to common "Albatross IV/Polyvalent" equivalents. The adjustments are based on estimated vessel and door effects from paired trawl experiments. This approach ignores the uncertainty and potential bias that results from uncertainty in the conversion factors. An approach that is more consistent with the actual data would be to develop priors for each of the components (i.e. door and vessel effects), and then model those effects while fitting to the

actual (unadjusted) survey indices. The full uncertainty associated with the vessel and door conversions would then be carried into the assessment, in particular in marginal posterior distributions.

Uncertainty in the stock assessment was explored through sensitivity runs and by estimating the marginal posterior distributions of key parameters using Markov-chain Monte-Carlo simulation (McMC). The range of sensitivity runs explored was extensive and included: alternative methods for calculating survey indices; inclusion/exclusion of survey indices including the state surveys; treating the RV Bigelow trawl time series as separate indices; inclusion of LPUE indices; alternative recreational discard mortality assumptions; alternative starting years; explicit treatment of catch fleets; and different terminal recruitment assumptions. Overall, model results were relatively insensitive to the alternative model configurations indicating conclusions about stock status are robust to these uncertainties. The SAW did a thorough job exploring uncertainty in the assessment results, and the final selected base model provides an appropriate basis for informing management decisions. The SAW proposed an alternative model formulation, which penalizes the magnitude of recent year-classes, also be used for projections to demonstrate the effect if these year-classes are overestimated in the base model. Given the magnitude of recent recruitment, this is the largest source of uncertainty in the assessment and the alternative model is a reasonable approach to bracket this uncertainty.

The potential impact of movement of GB haddock to the GoB stock was explored using the SCAA model. SCAA was fitted to same data as the ASAP model, and gave similar results to ASAP when run with similar assumptions. The stock reconstructions for runs with the movement hypotheses (permanent or temporary movement, see ToR 3) were similar to those for the isolated stock hypothesis, although estimates of current spawning stock biomass (SSB) were somewhat lower for the temporary movement assumption. Compared with the isolated stock hypothesis, catch projections were somewhat higher under the permanent movement hypothesis and were similar for the temporary movement hypothesis.

Given the robustness of assessment results to stock mixing assumptions, the low estimate of mixing rates, and the potential risk to the GoM stock if mixing rates are overestimated, the final ASAP model (ASAP_final_temp10) recommended by the SAW 59 as the preferred model for the assessment is appropriate.

5. *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.*

This ToR was met.

The existing reference points are based on *MSY* proxies. That is, overfishing is defined as fully-recruited fishing mortality rates greater than the F_{MSY} proxy of $F_{40\%}$= 0.46, and overfished is based on B_{MSY} which is calculated as the median *SSB* in long-term stock projections fishing at $F_{40\%}$. Overfished is defined as *SSB* less than 50% B_{MSY} (2,452 mt).

New reference points were calculated because of changes in data inputs and the assessment model, as well as new estimates of age-specific selectivity and weights-at-age. The recent selectivity pattern and average weights-at-age from 2009 – 2012 were assumed. The existing F_{MSY} proxy, $F_{40\%}$, was maintained but different assumptions about recruitment were made for the long-term stock projections. The previous (GARM III) method for simulating recruitment excluded very large and very small historic year-classes in the simulations. For this benchmark assessment, the SAW 59 WG decided to use the entire series of

recruitment estimates (1977-2011), excluding only the more recent and less well determined 2012 and 2013 recruitments. This approach is appropriate, given the previous decision of year-classes to include in the reference point estimation was ad hoc.

The SAW 59 WG also explored an alternate approach to estimating biological reference points, based on a Ricker stock-recruitment relationship with time-varying productivity. The WG decided that it was premature to adopt this approach as a broader range of stock-recruitment relationships should first be explored and an understanding of the processes driving a productivity shift would need to be developed.

The new BRPs proposed by the SAW 59 are appropriate.

6. *Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model 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-5).*

This ToR was essentially met, although the existing model was modified prior to the update with revised and new data. The stock is not currently in a rebuilding plan.

One of the key assumptions of the VPA model used for previous GoM haddock assessments was modified prior to updating the model with data through 2013. The assumption relates to how the terminal year plus group abundance is estimated, changed from a "backward" computation to a "combined" computation. When this change was applied to the data used for the 2012 assessment update, it resulted in approximately 10% increase in abundance and a reduction in the terminal year full-recruited F from 0.82 to 0.52. This modified VPA model was then updated with the revised and updated data used for the 2013 assessment. While the change from the "backward" to "combined" computation method is warranted, it is not clear if reference points, in particular the F_{MSY} proxy would change with this alternative model assumption. The stock status determination for this revised VPA model is that the stock is not overfished, but overfishing is occurring.

Using the new ASAP model results and proposed biological reference points, the GoM haddock stock is not overfished ($F_{40\%}$=0.46 and $F_{Current}$=0.39) and overfishing is not occurring ($B_{Threshold}$=2,054 mt and B_{curren}t=4,108 mt).

The existing $F_{40\%}$ reference point is based on the average fishing mortality rate over ages 6 to 8 while the proposed $F_{40\%}$ reference point is based on the age 7 fishing mortality rate (the fully-selected age). Small differences in the exploitation pattern can have a large effect on the $F_{40\%}$ reference point when measured as "fully selected" F, in particular when only a small component of the catch is comprised of the "fully selected" age class(es). The difference in stock status determination between the existing and proposed F reference point may be the result of this sensitivity. Alternative fishing intensity reference points should be explored (e.g., exploitation rate or F averaged over all selected ages) to see if there are more stable ones that fulfill the intent of determining overfishing status.

7. *Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).*
    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, migration from Georges Bank).*
    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.*

This ToR was mostly met, though annual probabilities of exceeding the threshold *F* or falling below the threshold biomass were not provided.

Assumptions for stock projections followed those used in estimating reference points: using the current selectivity ogive; average weights-at-age from 2009 – 2013; and recruitment estimates from the 1977 to 2011. Stock projections were made for the base case model (ASAP_final_temp10) and for a sensitivity run that constrained the size of the 2012 year class (ASAP_final_temp11) which is considered to be the largest uncertainty in the assessment. Uncertainty in catch and SSB estimates was based on their marginal posterior distributions (from projections for each of the 1000 MCMC samples).

Annual probabilities for exceeding the *F* threshold and for falling below the biomass threshold could have been calculated by fixing the projection catches (say, at mean catch projected while fishing a target *F*), and evaluating the probabilities across the MCMC scenarios. This would be consistent with the ToR, and may be more useful to managers than the presentation of uncertainty in catch and SSB which assumes application of the target *F* with perfect knowledge of current abundance.

The SAW 59 WG determined that projections from the base case model (ASAP_final_temp10) were the most realistic. This decision is appropriate given that it had been demonstrated that recruitments estimated with an unconstrained (on recruitment deviations) model produced the most reliable recruitment estimates. The alternative projection (ASAP_final_temp11) addresses the major uncertainty in the stock assessment.

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

This ToR was met.

The SAW 59 WG reviewed the status of previous research recommendations and proposed new ones to address issues raised during their meetings, some of which are generic to many northeast US groundfish stocks.

I have two research suggestions, in addition to those proposed by the WG.

A common procedure for northeast US groundfish stock assessments is to adjust the NEFSC bottom trawl survey abundance indices to standardize for the effects of changes in vessels and doors. The adjustments are based on calibration coefficients estimated from paired trawl experiments, which have a large associated uncertainty. Probability distributions could be calculated for the calibration coefficients, and

these used as priors in Bayesian assessment models that fit to the unadjusted index data. This would ensure that uncertainty from the calibration coefficients are included in the assessment models.

Ageing data are not available for the GoM haddock recreational fishery and commercial discards, so age-length keys are borrowed from commercial fisheries and surveys. This has the potential to introduce bias in the age composition estimates and should be investigated. Given the recreational fishery is becoming a major component of the GoM haddock catch, direct ageing of samples from this fishery may be warranted.

4.3 SEA SCALLOPS

SARC 59 reviewed a benchmark stock assessment for the US sea scallop resource. Sea scallop stock assessment and management occur at a range of geographical scales. Stock status, in terms of overfished and overfishing declaration, is based on the entire east coast scallop resource, while actual management is at a relatively fine scale with numerous closed and rotational fishing areas. The SARC 59 stock assessment was conducted using separate models for two Georges Bank (GB) areas (open and closed) and for a single Mid-Atlantic (MA) area. For the purpose of stock status determination, results from the three assessment models are combined, while stock projections are conducted at the scale of fisheries management (i.e. explicit modelling of closed and rotational fishing areas). Given the disconnect between the scale at which actual fisheries management occurs and the requirement to determine stock status for the entire stock, and the inability to disaggregate catch at a finer level, the geographical units used for the assessments are appropriate.

The data sources for the sea scallop assessment were adequate and used correctly. The catch is almost entirely from the commercial fishery and is well documented, though there is uncertainty in the magnitude of dredge fishery-induced incidental mortality. A long-term survey (dredge) provides solid information on stock trends and a drop-video survey provides estimates of near-absolute abundance, resulting in relatively high precision in the stock reconstructions. A length-based statistical model (CASA), developed explicitly for the sea scallop assessment, is appropriate for the available data. Sensitivity analyses were conducted to investigate alternative data and model assumptions, and the model selected for determining stock status was appropriate.

Biological reference points, based on F_{MSY} and B_{MSY}, were updated using results from the new benchmark assessment. The SAW conclusion that the stock is not overfished, and overfishing is not occurring is consistent with the analyses presented.

Findings relative to each of the Terms of Reference follow:

1. *Estimate removals from all sources including landings, discards, incidental mortality, and natural mortality. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these assumptions and sources of data. If possible using sensitivity analyses, consider the potential effects that changes in fishing gear, fishing behavior, and management may have on the assumptions.*

This ToR was met.

Estimates of commercial sea scallop landings, measured in meat weight, are likely fairly precise and accurate, with some uncertainty arising from the proration procedure used to allocate landings to regions. When fitted in the stock assessment model, there is additional uncertainty due to predicting catch in meat weight because of uncertainty in the shell height-meat weight relationship. The shell height-meat weight

relationships are from research surveys but are adjusted for fishing practices and seasonal variation based on observer data.

Discard estimates are available from observer observations and vessel trip reports since 1989, though not all gear types were surveyed in all years. Discards have tended to be a small proportion of landings, with somewhat higher rates of about 10% in the MA between 2001 and 2004. The average mortality rate for discarded scallops is uncertain and likely variable. Size distributions for landed and discarded scallops show that primarily smaller scallops are discarded, though the assessment model assumes the same size distribution for landings and discards.

Incidental scallop mortality occurs when dredges do not capture all scallops encountered and these are damaged and killed. Encounter of scallops not captured by the dredge gear, is accounted for in the assessment model through the length-dependent fishery selectivity, and an assumed mortality rate applied to these. Experimental studies suggest incidental mortality rates of about 10% to 20%, and these are likely substrate dependent.

The size distribution of commercial landing and discard length frequencies (LFs, measured as shell height) are available from observer sampling since 1992. Prior to that, length frequency data are available from port samples but comparison of 1992 to 1994 port samples and observer LF data indicate bias in the port samples. The port sample LF data are only fitted in the model prior to 1985, because those samples are not thought to be biased.

Recreational catch is negligible, and is ignored in the assessment.

The natural mortality rate (*M*) assumed for GB scallops (0.16), which had been estimated in the CASA model previously, is somewhat higher than the estimate from clapper count and clapper separation time (0.12) used in previous stock assessments. The natural mortality rate is adjusted for the MA region based on the assumption that the ratio of the growth coefficient *K* to *M* is invariant, resulting in an estimate of *M* for the MA area of 0.20. Mortality for the "plus" length class is assumed to be 1.5 times that of smaller scallops.

The compilation of the sea scallop catch data appears to have accounted for all significant sources of mortality. Landings (meat weights) estimates appear to be relatively accurate and precise. Discards are generally a small fraction of landings, so uncertainty in discard mortality is not likely a major source of uncertainty for the assessment. Uncertainty in incidental and natural mortality rates are a greater source of uncertainty for the assessment, in particular potential size and density effects on natural mortality.

2. *Present the survey data being used in the assessment (e.g., regional indices of relative or absolute abundance, recruitment, size data, etc.). Characterize the uncertainty and any bias in these sources of data.*

Survey data used in the assessment include: a stratified random dredge survey that has been conducted annually since 1975; a drop camera survey (SMAST) conducted from 2003 to 2012; a towed video survey (Habcam) conducted from 2011 to 2013 in GB and in 2012 and 2013 in MA; and a bottom trawl survey (MA only). All the surveys are wide-area, encompassing the majority of the MA and GB regions.

The dredge survey used a different dredge for the first 4 years, and in the assessment model, these data are appropriately treated as a separate survey. Changes since then, a vessel change and use of rock excluder chains on hard bottom, have affected the consistency of the time series but are accounted for by adjusting the abundance estimates to account for their effects. These changes are unlikely to introduce

bias to the survey time series, and the additional uncertainty due to the use of the rock excluder chains is included in the variance calculations.

For this assessment, the dredge survey area has been modified to remove marginal scallop habitat. Then, survey abundance estimates are increased to account for the areas not surveyed. While this approach decreases the variance of the abundance estimates, there is the potential that it could result in survey bias if abundance in the currently marginal habitat were to increase in the future.

Swept-area estimates from the dredge survey are adjusted to account for dredge efficiency to generate estimates of absolute abundance. The substrate-dependent dredge efficiency estimates, from paired video (Habcam) and dredge data, assume that all scallops greater than 40mm shell height are available to both gears. If the region-specific dredge efficiency estimates are incorrect, this will lead to biased estimates of absolute abundance (see ToR 4 for further discussion of this).

Considerable work has gone into developing the survey design for the Habcam drop video survey, and to analyzing results and estimating the uncertainty in density estimates. The survey has the potential to generate a relatively accurate and unbiased estimate of absolute abundance. The primary issue that could potentially bias abundance estimates from this survey is difficulty in identifying "dead" shells.

The SMAST drop camera survey uses a "large" camera system which does not fully sample smaller scallops. Results from comparison work, using both a "small" and "large" camera, are used to estimate the size-dependent detectability of the SMAST "large" camera, and these are used to adjust the survey results to absolute abundance estimates. This approach does not account for uncertainty in the SMAST abundance estimates, and has the potential to introduce bias to the time-series (see ToR 4 for further discussion of this).

The assessment document description of the SMAST survey is incomplete and requires some work so that readers can understand how this survey is conducted and how the data are analyzed.

3. *Investigate the role of environmental and ecological factors in determining recruitment success. If possible, integrate the results into the stock assessment.*

This ToR was met.

Preliminary analyses of remote sensing and scallop dredge data suggest that recruitment may be influenced by spring and summer phytoplankton bloom activity in the MA. Further research is required to develop techniques for predicting regional recruitment patterns, but this work has the potential to provide results at spatial scales of interest to managers.

An additional study looked at the relationship between scallop recruitment and abundance of the sea star *Astropecten americanus*, a predator of juvenile sea scallops. A stock-predator-recruitment model for the southern MA was developed, with promising results that warrant further research.

4. *Estimate annual fishing mortality, recruitment and stock biomass for the time series, and estimate their uncertainty. Report these elements for both the combined resource and by sub-region. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.*

This ToR was met.

Stock reconstructions were conducted using a statistical length-based model (CASA) that has been used for US scallop assessments since 2007. Separate analyses were conducted for the open and closed areas of Georges Bank (GB) and for the Mid-Atlantic Bight (MAB), and results combined to assess the entire stock. CASA models growth using growth transition matrices estimated from shell ring increment data, and incidental mortality can be modelled explicitly. The models are fitted to abundance estimates from a relatively long time-series of dredge survey data (1979-2013), a camera survey (SMAST), a video survey (HABCAM), and a NEFSC bottom trawl survey (MAB only).

The methods used to reconstruct the scallop resource and estimate fishing mortality, recruitment, and stock biomass are sound and provide a scientifically credible basis for developing fishery management advice. Uncertainty in the assessment is underestimated and some suggestions that may improve this are identified below.

The HABCAM video survey theoretically provides an estimate of absolute abundance (for scallops of 40mm shell height and greater), and the assumption that "catchability" (q) for this survey has a mean of 1 is probably reasonable. However, the assumed c.v. on this survey q (0.1) likely underestimates it's true uncertainty, given adjustments made to the data for areas not surveyed and difficulty in identifying dead shells in the video images.

The dredge survey and SMAST camera surveys do not provide estimates of absolute abundance (for the 40mm+ scallops), however the data are adjusted to account for scallops not observed in the surveys and then treated as if they were absolute abundance in the model (i.e. assumption that the q mean is 1). For the dredge survey, the data adjustments are based on capture efficiency estimates from experimental work comparing dredge and HABCAM observations. Thus, the model q's for HABCAM and the dredge survey are not independent. An approach that is more consistent with how the data were collected and reflects the lack of independence between the HABCAM and dredge survey q's would be to use the unadjusted dredge survey observations in the model with a prior for the ratio of the dredge q to the HABCAM q. The prior for that ratio would be based on the efficiency estimates from the HABCAM/dredge comparisons and some idea of the uncertainty in the substrate types in each area.

The SMAST camera survey does not fully "observe" smaller-sized scallops, and comparison data using "small" and "large" cameras were used to estimate size-selectivity for the SMAST survey. The actual SMAST observations are adjusted using this size-selectivity to estimate total (40mm+ scallops) abundance for use in the CASA model. A more appropriate approach would be to use the "small" versus "large" camera comparison data directly in CASA to estimate the SMAST size-selectivity and fit to the actual SMAST observations. This would appropriately account for the uncertainty in the fraction of the population sampled by the SMAST survey.

For all three assessment regions at least two of the point estimates for the survey q's are at the extremes of their priors. The treatment of the survey time series (dredge survey, HABCAM and SMAST) as independent estimates of absolute abundance will result in uncertainty in stock biomass being underestimated and potentially biased.

Certain aspects of the data suggest that juvenile scallop mortality may be density-dependent, with higher natural mortality on large year-classes. Survey abundance increases dramatically when a large year-class recruits but over the following 1 to 2 years declines much quicker than expected from a constant natural mortality assumption. The model cannot fit the large increases in survey abundance, and retrospective patterns occur when large year classes recruit. A formulation of the CASA model that included a juvenile density dependent mortality assumption was fitted to the GB open and GB closed data. This formulation improved the fits to the survey data, and resulted in greater consistency between the point estimates for survey q's (i.e. modes of the posteriors) and their priors. Due to insufficient time to fully explore the

density dependent mortality models, they were presented only as exploratory runs. This work is promising and evaluation of the density dependent juvenile mortality assumption should be continued.

The implementation of ASAP used for the scallop assessment assumed the size composition of commercial landings and discards were the same, although this assumption is not supported by the landings and discard length frequency data. This model assumption should be changed and the model fitted to the discard length frequency data. It is unlikely that the current assumption of equal selectivity would lead to any significant bias in the assessment because discards have tended to be a relatively small component of the catch.

There is some indication in the profiles on dredge survey *q* that some of the runs minimized at local, rather than global, minima. That is, the relationship between dredge *q* and the likelihoods for some model components are not smooth (e.g., see the dredge survey likelihood component in Table 6.2), as would be expected. Methods for testing for local minima should be developed. Example approaches are to initialize the model at a range of initial values, or when doing profiles to restart each run at the parameter set from the previous minima.

Model fits to the MA data were unsatisfactory when a standard data weighting approach was used (LF data weighted to be consistent with the magnitude of the resulting residuals), which was resolved by weighting the dredge survey LF data much higher than suggested by the fits to the data. The lack-of-fit without overweighting the dredge survey data is indicative of conflict among some of the data or some form of model misspecification. This issue should be investigated further. One potential approach is to remove data sets one at a time, to see if there is conflict among them.

Two aspects of the stock assessment suggest that the CASA models may be overestimating recent stock abundance: 1) there are relatively strong retrospective patterns for all 3 regions, with stock biomass tending to decrease as additional years of data are added; and 2) there is a lack of fit to recent survey trends, with surveys suggesting the total stock abundance has decreased over the last 3 years while the assessment models suggest abundance has remained at historic high levels. These patterns may be related to incorrect model specification of juvenile natural mortality.

5. *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. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.*

This ToR was met.

Per recruit reference points (F_{MAX} and B_{MAX}) were used as proxies for F_{MSY} and B_{MSY} in assessments prior to 2010, however as selectivity shifted to larger scallops, yield per recruit curves become increasing flat, making per-recruit reference points unstable. For SARC 50 and this benchmark stock assessment, a stochastic model (SYM) was used to calculate biological reference points (F_{MSY} and B_{MSY}) and estimate their uncertainty. F_{MSY} and B_{MSY} were estimated rather than using proxies because a stock-recruitment relationship was assumed in the estimation. The stock-recruitment relationship was not a feature of the stock assessment.

Parameters for the SYM model were similar to those for the assessment. Uncertainty was incorporated in the assumptions about growth, natural mortality, discard and incidental mortality, fishery selectivity, shell height- meat weight conversions, and the stock-recruitment relationship. A 10% trimmed mean was used

to obtain the central tendency of per recruit and yield curves. F_{MSY} was taken as the fishing mortality that maximized the trimmed mean curve.

The maximum F investigated in the SYM analysis, 1, is too low given F_{MSY} was restricted by this bound for some of the parameter combinations. Although a higher maximum F should be modelled, this will likely have only a minimal effect on the F_{MSY} and B_{MSY} estimates.

The SYM analysis was run separately for the GB and MA regions. F_{MSY} was estimated at 0.30 for GB, 0.74 for MA, and 0.48 for the combined regions. The differences between GB and MA estimates are consistent with the higher observed growth and estimated natural mortality for MA. While stock status determination (overfished and overfishing) is done at the combined stock level, actual stock management is at a finer spatial scale with rotational open/closed areas.

6. *Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model or model formulation developed for this peer review.*
    a. *Update the existing model 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-5).*

This ToR was met.

Sea scallops are considered overfished when the stock falls below the biomass threshold, defined as $0.5B_{MSY}$. The current biomass threshold is 62,679 mt and the recommended value from the SARC 59 benchmark assessment is 48,250 mt. The estimated combined area biomass in 2013, 132,561 mt, is above both biomass threshold levels, so the stock is not overfished based on either definition. Results suggest the probability that the stock is overfished is very small, likely less than 1%.

The current F_{MSY} is 0.38 and the proposed F_{MSY} is 0.48. The estimated fully-recruited fishing mortality rate for the combined stock in 2013 is 0.32, so overfishing is not occurring under either criterion.

7. *Evaluate the realism of stock and catch projections and compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level).*
    a. *Provide numerical annual projections (through 2016). 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 the realism of the projections. 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.*

This ToR was met.

Example stock projections were carried out using the SAMS model. SAMS is spatially explicit, reflecting ten sub-areas in GB and seven sub-areas in the MA. The sub-areas reflect the spatial structure used in fisheries management, which encompasses closed and rotational fishing areas. Actual projections used to inform fisheries management are carried out by the Scallop Plan Development Team to evaluate potential management measures.

The SAMS model is an appropriate tool for evaluating alternative management measures. It more realistically captures the biologically heterogeneity of the scallop resource than the assessment model, and allows explicit evaluation of alternative management approaches (open/closed areas). Uncertainty in the projections is included through stochastic initial conditions, recruitment and natural mortality.

Under the example stock projection, catches are forecast to increase over the next 3 years due to reopening rotational fishing areas. Projections suggest there is almost no chance of overfishing or the stock becoming overfished in the near future under the assumed management scenario.

SAMS is believed to overestimate stock abundance. This may be related to incorrect assumptions about juvenile mortality or misspecification of survey *q*'s, and requires further investigation.

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

This ToR was met.

A majority of the research recommendations from SARC 50 have been addressed or currently have experimental studies underway. A recommendation to investigate seasonal patterns of growth and incorporate this into the assessment model was not pursued. I concur with the assessment team's conclusion that that work is not a high priority. SARC 50 also recommended the CASA code be modified to fit to scallop discard data. This has not been done, however I feel it should be made a priority for future assessments. Discard size distribution data are available for at least some years, so it would be possible to model discard mortality directly.

The SAW 59 Invertebrate Subcommittee made a number of recommendations for future work, though they did not prioritize their list. I feel that some of their recommendations are worth highlighting as they potentially would have greater impact on future stock assessments. 1) Improve procedures for identifying dead scallops in optical surveys and evaluate the effect of uncertainty in assessments. This could include a comparison of Habcam and SMAST estimates of the proportion dead. 2) Quantify and improve accuracy of the SAMS projection model which appears to overestimate stock size. 3) Continue work on density dependent natural mortality for small scallops.

In addition to the SAW 59 Invertebrate Subcommittee recommendations, I recommend that priors for the survey *q*'s (absolute or relative) be developed and incorporated in the assessment model. If a fully Bayesian assessment were implemented, this would allow evaluating the uncertainty in assessment results more fully.

5. SUMMARY AND CONCLUSIONS

The SARC 59 review process for Gulf of Maine haddock and sea scallops was thorough and effective, and resulted in a comprehensive review of the data and analytical methods used to assess the stocks. The working group process (SAW) used to develop the assessments ensured that assumptions and uncertainties in the assessments had been thoroughly investigated. Preparation of draft assessment documents and summary reports prior to review meeting and limiting the review to two assessments ensured there was adequate time to become familiar with the assessment data and methods and to

thoroughly investigate implications of modelling assumptions. The terms of reference for the review process are clear and explicit, and provide a useful guideline for the content of stock assessment reports.

The Review Panel members agreed all substantive issues and the SARC 59 Summary Report represents consensus opinion. Stock status Summary Statements, agreed during the review, reflect the best possible use of the available information and conclude that Gulf of Maine haddock and sea scallops are not overfished and overfishing is not occurring.

Appendix 1: SARC 59 Panel Members

Jean-Jacques Maguire, New England Fishery Management Council SSC (Chair)
Yiota Apostolaki, Center for Independent Experts reviewer
Vivian Haist, Center for Independent Experts reviewer
Coby Needle, Center for Independent Experts reviewer

Appendix 2: Bibliography of materials provided for review

Working Papers

Working Group, Stock Assessment Workshop (SAW 59) 2014. Stock Assessment Report of Gulf of Maine haddock. Working Paper #1. SAW/SARC 59. July 15-18, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 59) 2014. Stock Assessment Report of Sea scallops. Working Paper #1.SAW/SARC 59. July 15-18, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 59) 2014. Stock Assessment Summary Report of Gulf of Maine haddock. Working Paper #2. SAW/SARC 59. July 15-18, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 59) 2014. Stock Assessment Summary Report of Sea scallops. Working Paper #2. SAW/SARC 59. July 15-18, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Background Documents

Gulf of Maine Haddock

Bell R., Hare J. 2014. Gulf of Maine Haddock Stock Recruitment Model: Time Varying Parameters. Northeast Fisheries Science Center, National Marine Fisheries Service, Naragansett, RI. 6 p.

Brodziak J.K.T., Col L., Palmer M., Brooks L. 2008. Northeast Consortium Cooperative Haddock Tagging Project: Summary of Reported Haddock Tag Recaptures Through September, 2008. NEFSC unpublished manuscript. 31 p.

Cape Cod Commercial Hook Fishermen's Association.2009.Haddock Migration in New England Waters: Year 1 and Year 2 Analysis of Closed Area and Stock Boundaries. 61 p.

Miller T. and Palmer M. 2014. Estimates of mortality and migration from Gulf of Maine and Georges Bank haddock tag-recovery data, 2005-2010. SARC 59 Working Paper. 15 p.

New England Fishery Management Council. 2013a. GB haddock stock spillover to GOM haddock stock. Groundfish PDT Memo, August 8, 2013. 46 p.

New England Fishery Management Council. 2013b. Spillover of haddock between the Georges Bank and Gulf of Maine stocks. 12. SSC - September 24-26, 2013 – M. 2 p.

NOAA Fisheries Toolbox (NFT).September 2012. Technical Documentation for ASAP Version 3.0. 71 p.

Northeast Fisheries Science Center (NEFSC). 2010. VPA/ADAPT Version 3.0 Reference Manual. NEFSC. 29 p.

Northeast Fisheries Science Center (NEFSC). 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p.

Palmer M. Gulf of Maine haddock 2-756 R. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 8, 2008, by Northeast Fisheries Science Center. 2-756 – 2-823, 68 p.

Palmer M., Correia S., Nitschke P., Estimating the year-class size of terminal year cohorts in stock assessment models: the Gulf of Maine haddock (Melanogrammus aeglefinus) example. SARC 59 Working Paper. Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water St., Woods Hole, MA, 02543. 33 p.

Palmer M., Nitschke P., Wigley S., Rago P. 2014. Estimation of haddock bycatch in the northeast United States midwater trawl fishery. Northeast Fisheries Science Center, 166 Water St., Woods Hole, MA 02543.10 p

Atlantic Sea Scallops

Hart D.R. 2013. Quantifying the tradeoff between precaution and yield in fishery reference points. ICES Journal of Marine Science; doi:10.1093/icesjms/fss204. 13 p

Hart, D.R., and Chute, A. S. 2009. Estimating von Bertalanffy growth parameters from growth increment data using a linear mixed-effects model, with an application to the sea scallop Placopecten magellanicus. ICES Journal of Marine Science, 66: 2165–2175. 11 p.

Hart D.R., Jacobson L.D., Tang J. 2013. To split or not to split: Assessment of Georges Bank sea scallops in the presence of marine protected areas. Fisheries Research, 144, 74– 83.

Hennen D.R. and Hart D.R. 2012. Shell height-to-weight relationships for Atlantic sea scallops (Placopecten magellanicus) in offshore US waters. Journal of Shellfish Research, Vol. 31, No. 4, 1133–1144. 13 p.

Northeast Fisheries Science Center (NEFSC). 2004. Essential Fish Habitat Source Document: Sea Scallop, Placopecten magellanicus, Life History and Habitat Characteristics (Second Edition). NOAA Technical Memorandum NMFS-NE-189. 32 p.

Northeast Fisheries Science Center (NEFSC). 2010. 50th Stock Assessment Workshop (SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 10-17. 106 p.

Northeast Fisheries Science Center (NEFSC). 2010. 50th Stock Assessment Workshop (SAW) Assessment Summary Report. Northeast Fisheries Science Center Reference Document 10-09. 12p.

Shank BV, Hart DR, Friedland KD. 2012. Post-settlement predation by sea stars and crabs on the sea scallop in the Mid-Atlantic Bight. Mar Ecol Prog Ser 468: 161–177. 17 p.

Appendix 3: Statement of Work

Attachment A: Statement of Work for Dr. Vivian Haist

59th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for Gulf of Maine haddock and sea scallops

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

BACKGROUND

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

SCOPE

Project Description: The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC peer review is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development and report preparation (which is done by SAW Working Groups or ASMFC technical committees), assessment peer review (by the SARC), public presentations, and document publication. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fishery management in the northeast region.

The purpose of this meeting will be to provide an external peer review of benchmark stock assessments for **Gulf of Maine haddock** and **sea scallops**.

OBJECTIVES

The SARC 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.

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 draft stock assessment Terms of Reference (ToRs) which are carried out by the SAW WGs are attached in **Annex 2**. The draft agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is described in **Annex 4.**

Requirements for the reviewers: Three reviewers shall conduct an impartial and independent peer review of the **Gulf of Maine haddock** and **sea scallop** stock assessments, and this review should be in accordance with this SoW and stock assessment ToRs herein. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include statistical catch-at-age, state-space and index models. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points that includes an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 59 will address fishery stock assessments of **Gulf of Maine haddock** and **sea scallop.** For scallops, knowledge of sessile invertebrates and spatial management would be desirable. For GOM haddock, understanding of fish movements and exchange between stocks would be desirable.

PERIOD OF PERFORMANCE

The contractor shall complete the tasks and deliverables as specified in the schedule of milestones within this statement of work. Each reviewer's duties shall not exceed a maximum of 16 days to complete all work tasks of the peer review described herein.

Not covered by the CIE, the SARC chair's duties should not exceed a maximum of 16 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation).

PLACE OF PERFORMANCE AND TRAVEL

Each reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during July 15-18, 2014.

STATEMENT OF TASKS

Charge to SARC panel: During the SARC meeting, the panel is to determine and write down whether each stock assessment Term of Reference (ToR) 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. **If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.** 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 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, <u>and</u> 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.

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

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

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (or by email if necessary) the requested information (e.g., 1.name [first, middle, and last], 2.contact information, 3.gender, 4.country of birth, 5.country of citizenship, 6.country of permanent residence, 7.whether there is dual citizenship, 8.country of current residence, 9.birth date [mo, day, year], 10.passport number, 11.country of passport) 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/.

Pre-review Background Documents and Working Papers: 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 SARC chair and 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 COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

Tasks during the panel review meeting: Each reviewer shall conduct the independent peer review of the stock assessments in accordance with the SoW and stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and contractor.** Each 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 discussions, 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. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

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. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

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.

Tasks after the panel review 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).

DELIVERY

Each reviewer shall complete an independent peer review report in accordance with the SoW. Each reviewer shall complete the independent peer review according to required format and content as

described in **Annex 1**. Each 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 scheduled during the tentative dates of July 15-18, 2014.
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than August 1, 2014, 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 Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment ToR in **Annex 2**.

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

June 10, 2014	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
July 1, 2014	NMFS Project Contact will attempt to provide reviewers the pre-review documents
July 15-18, 2014	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
July 18, 2014	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
August 1, 2014	Reviewers submit draft independent peer review reports to the contractor's technical team for independent review
August 1, 2014	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
August 8, 2014	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
August 15, 2014	Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements
August 22, 2014	The COR distributes the final 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 COR within 10 working days after receipt of all required information of the decision on substitutions. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the 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: The deliverables shall be the final peer review report from each reviewer that satisfies the requirements and terms of reference of this SoW. The contract shall be successfully completed upon the acceptance of the contract deliverables by the COR based on three performance standards:

(1) each report shall be completed with the format and content in accordance with **Annex 1**,
(2) each report shall address each stock assessment ToR listed in **Annex 2**,
(3) each report shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Upon the acceptance of each independent peer review report by the COR, the reports will be distributed to the NMFS Project Contact and pertinent NMFS science director, at which time the reports will be made publicly available through the government's website.

The contractor shall send the final reports in PDF format to the COR, designated to be William Michaels, via email William.Michaels@noaa.gov

Support Personnel:

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

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

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

Key Personnel:

Dr. James Weinberg, NEFSC SAW Chairman, NMFS Project Contact
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, NEFSC Science Director
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
william.karp@noaa.gov Phone: 508-495-2233

Annex 1: Format and Contents of Independent Peer Review Report

1. The independent peer review 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 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 ToR of the SAW was completed successfully. For each ToR, the Independent Review Report should state why that ToR was or was not completed successfully. To make this determination, the SARC chair and 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 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 independent report shall be an independent peer review of each ToR, 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 this Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: 59th SAW/SARC Stock Assessment Terms of Reference (file vers.: 1/17/2014)

A. Gulf of Maine (GOM) haddock

1. Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance.

2. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). If available, consider whether tagging information could be used in estimation of stock size or exploitation rate. Characterize the uncertainty and any bias in these sources of data.

3. Evaluate the hypothesis that haddock migration from Georges Bank influences dynamics of GOM stock. Consider role of potential causal factors such as density dependence and environmental conditions.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

5. 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.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model 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-5).

7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).
    d. 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, migration from Georges Bank).
    e. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
    f. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

Annex 2 (cont.):
B. Sea scallop

1. Estimate removals from all sources including landings, discards, incidental mortality, and natural mortality. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these assumptions and sources of data. If possible using sensitivity analyses, consider the potential effects that changes in fishing gear, fishing behavior, and management may have on the assumptions.

2. Present the survey data being used in the assessment (e.g., regional indices of relative or absolute abundance, recruitment, size data, etc.). Characterize the uncertainty and any bias in these sources of data.

3. Investigate the role of environmental and ecological factors in determining recruitment success. If possible, integrate the results into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass for the time series, and estimate their uncertainty. Report these elements for both the combined resource and by sub-region. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

5. 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. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model or model formulation developed for this peer review.
   a. Update the existing model 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-5).

7. Evaluate the realism of stock and catch projections and compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level).
   d. Provide numerical annual projections (through 2016). 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).
   e. Comment on the realism of the projections. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
   f. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

Annex 2 (cont.):

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 annual catch that is consistent with 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.

One model or alternative models:

The preferred outcome of the SAW/SARC is to identify a single "best" model and an accompanying set of assessment results and a stock status determination. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results.

Annex 3: Draft Agenda

59th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for A. Gulf of Maine haddock and **B. sea scallops**

July 15-18, 2014

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

DRAFT AGENDA* (version: Feb. 4, 2014)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR
Tuesday, July 15			
10 – 10:30 AM			
Welcome	**James Weinberg**, SAW Chair		
Introduction	**SARC Chair**		**TBD**
Agenda			
Conduct of Meeting			
10:30 – 12:30 PM	Assessment Presentation (Stock A.)		
	TBD	TBD	**TBD**
12:30 – 1:30 PM	Lunch		
1:30 – 3:30 PM	Assesssment Presentation (Stock A.)		
	TBD	TBD	**TBD**
3:30 – 3:45 PM	Break		
3:45 – 5:45 PM	SARC Discussion w/ Presenters (Stock A.)		
	SARC Chair		**TBD**
5:45 – 6 PM	Public Comments		

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Wednesday, July 16

9 – 10:45 AM	Assessment Presentation (Stock B.)		
	TBD	TBD	**TBD**
10:45 – 11 AM	Break		
11 – 12:30 PM	(cont.) Assessment Presentation (Stock B.)		
	TBD	TBD	**TBD**
12:30 – 1:45 PM	Lunch		
1:45 – 3:15 PM	SARC Discussion w/presenters (Stock B.)		
	SARC Chair		**TBD**
3:15 – 3:30 PM	Public Comments		
3:30 -3:45 PM	Break		
3:45 – 6 PM	Revisit with presenters (Stock A.)		
	SARC Chair		**TBD**
7 PM	**(**Social Gathering **)**		

Thursday, July 17

8:30 – 10:15	Revisit with presenter (Stock B.)		
	SARC Chair		**TBD**
10:15 – 10:30	Break		
10:30 – 12:30	Review/edit Assessment Summary Report (Stock B.)		
	SARC Chair		**TBD**
12:30 – 1:45 PM	Lunch		
1:45 – 2:15 PM	(cont.) edit Assessment Summary Report (Stock B.)		
	SARC Chair		**TBD**
2:15 – 2:30 PM	Break		
2:30 – 5 PM	Review/edit Assessment Summary Report (Stock A.)		
	SARC Chair		**TBD**

Friday, July 18

9:00 AM – 5:00 PM	SARC Report writing. (closed meeting)

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

The NMFS Project contact will provide the final agenda about four weeks before meeting. Reviewers must attend the entire meeting.

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 relevant 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.