55rd Northeast Regional Stock Assessment Review Committee

December 3 – December 7, 2012

Northeast Fisheries Science Center

Woods Hole, MA

SARC 55

SUMMARY REPORT

December 31, 2012

Review Committee

Patrick J Sullivan (Chair)
Noel Cadigan
John Casey
Steven Holmes

INTRODUCTION

Background

The 55th Stock Assessment Review Committee (SARC) convened at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA from December 3rd – December 7th, 2012 to review the stock assessments of Gulf of Maine and Georges Bank cod (*Gadus morhua*).

The SARC Review Panel (hereafter referred to as the Review Panel) consisted of Dr. Patrick J. Sullivan (Chair of the Committee, New England Fisheries Management Council's Scientific and Statistical Committee and Cornell University, Ithaca, New York, USA) and three scientists appointed by the Center for Independent Experts: Dr. Noel G. Cadigan (Centre for Fisheries Ecosystems Research, Fisheries and Marine Institute of Memorial University, St. John's Newfoundland, Canada), Dr. John Casey (CEFAS, Lowestoft, Suffolk, United Kingdom), and Dr. Steven Homes (Marine Scotland, Aberdeen, Scotland).

The SARC was supported and assisted by Dr. Jim Weinberg (NEFSC SAW Chairman), Dr. Paul Rago (Branch Chief of the NEFSC's Population Dynamics Branch) and NEFSC staff. The assessment documents for the Gulf of Maine and Georges Bank cod assessments were prepared by Stock Assessment Workshop Working Group (hereafter referred to as the Working Group), chaired by Robert O'Boyle (New England Fisheries Management Council's Scientific and Statistical Committee). The Georges Bank assessment was presented by Loretta O'Brien (NEFSC). The Gulf of Maine assessment was presented by Michael Palmer (NEFSC) and Dr. Doug Butterworth (Univ. of Cape Town). The support of all of these scientists and staff to the SARC process is gratefully acknowledged.

Review Activities

About two weeks before the meeting, assessment documents and supporting materials were made available to the Review Panel via an ftp server. At that time, the Review Panel was alerted that alternative models for the Gulf of Maine cod assessment were going to be presented for review as the SAW was unable to reach consensus on which model represented the best available science. On the morning of the meeting, the Review Panel met with Drs. Weinberg and Rago to discuss the meeting agenda, reporting requirements, and meeting logistics.

The SARC meeting started on Monday afternoon (December 3rd) with a welcome and introductions by Drs. Weinberg and Sullivan (See Appendix 1 for the detailed agenda). The Georges Bank cod assessment was presented over the remainder of the afternoon. The Gulf of Maine cod assessment was presented on December 4th (Day 2). Questions of clarification were raised and discussion pursued the following days on both assessments. All meetings of the SARC on Days 1-3 and part of 4 were held in open session. Towards the end of the week, the Review Panel prepared Assessment Summary Reports for both stocks in open session. Rapporteurs provided detailed records of all open sessions. For the remaining parts of December 6th - 7th (Day 4-5), the Panel met in closed session to work on its SARC Summary Report. There was insufficient time to finish the report during the last day and a half of the meeting and the SARC Summary Report was completed by correspondence.

SARC Process and General Conclusions

The Review Panel agreed unanimously regarding its conclusions on all the Terms of Reference it was charged to address for both Gulf of Maine cod and Georges Bank cod assessments. It acknowledges the significant work that the Working Group undertook in preparing and presenting the assessments. It also appreciates the professionalism and cooperation of NEFSC staff and all participants at the SARC meeting, which significantly assisted the peer review. Here we identify some overall conclusions pertinent to the SARC process for both assessments and highlight the principal conclusions for each assessment. We expand on the principal conclusions for each assessment in the subsequent sections.

General Conclusions

The Review Panel found the work presented by all participants to be thorough and thoughtful. The Panel recognizes the challenges for both scientists and stakeholders regarding the development of the current assessment, but believes that the information presented in both the Gulf of Maine and Georges Bank cod assessments represents the best available science.

While acknowledging that the level of work that has gone into these assessments has been extraordinary, some issues remain unresolved, in particular with regard to several assumptions associated with the Gulf of Maine assessment. These matters are discussed in greater detail below. In brief, after three meetings and many months of work, no single preferred model was put forward by the Working Group for the Gulf of Maine cod stock. Several models, with their associated assumptions, remained on the table and were presented to the Review Panel. The proponents of each of these analytical approaches may have believed that their set of assumptions held the greater weight, but in the end, although several models eventually were removed from consideration, two variations for Gulf of Maine cod remained, which basically represented different methods for accounting for what appear to be unknown sources of change in the system. At the end of the meeting and with the time and information available there appeared to be no clear way forward for providing a single preferred model for guiding management. Consequently, the Review Panel, in conjunction with participating members of the Working Group, outlined the consequences associated with using or disregarding the remaining approaches so that managers might act accordingly.

With regard to Georges Bank cod, the Review Panel was able to reach consensus on a single assessment model.

Summary of Gulf of Maine Cod

The SAW Working Group was unable to reach a consensus on a preferred model for the Gulf of Maine cod assessment prior to submitting it to the SARC Review Panel for peer review. Although this was somewhat understandable given the history of the assessment process for this stock and the multiple perspectives in the Working Group on how model assumptions and inputs might be altered to best account for disparities in the assessment (e.g. differences in biological and ecological interpretations of observations, retrospective patterns, inconsistent indices, and residual patterns), this resulted in the need for a much more in-depth discussion of the technical content, justification for, and merits of the multiple models. And, while several of the Working Group models presented and discussed were, in the end, eliminated for further consideration by the Review Panel, two of the remaining models still had to

be moved forward. We recognize that this will likely complicate the management process by requiring multiple reference points, possibly alternative conclusions about stock status, and multiple methods for deriving stock projections. However, this was determined to be the best method for conveying our conclusions in the time available and under the rules of order that needed to be established for the meeting to facilitate at least some conclusions being drawn. Several of these points will be expanded upon below.

Multiple Models

It has been a standard recommended practice for many years (see for example, NRC Report, 1998), to consider multiple models when conducting an assessment. Alternative models can be useful in determining the likelihood and consequences of alternative assumptions about the data, the fishery, and the state of nature and how assessment models can be used to represent them. However, in New England, and in the United States in general, the practice has been to put forward a preferred assessment among the multiple available that best characterizes the state of the system, while perhaps carrying forward the results of alternative models to represent the range of assessment uncertainties. This gives decision makers a clear approach to use, while understanding the consequences of alternative decisions. One alternative to such an approach, used in other parts of the world, is model averaging, whereby the results of alternative models are averaged for the purposes of developing management actions. There are pros and cons to both approaches, however, the process of discussion and consensus building around a single model of choice often can lead to an improved understanding of the alternatives and their associated consequences. This part of the process can be undercut if the option of model averaging is relied on too quickly. Furthermore, model averaging can sometimes hide the consequences associated with alternative assumptions about the state of the system. Consequently, here we put forward several alternative assessments, but would suggest that they be viewed separately in terms of their assumptions, outcomes, and consequences rather than being averaged for decision making.

The Review Process

The SARC meeting is generally set up to provide an independent peer review of the outcome of the SAW Working Group deliberations and to determine if it is the best available science as defined under the Magnuson-Stevens Fishery Conservation and Management Act. The process usually expects the Working Group, conducting several week-long workshops over several months, not to mention months of work in parallel by individual assessment authors, to have a preferred assessment with perhaps alternative models to show sensitivity to and the consequences of alternative assumptions. The process also typically entails re-examination of model runs and outcomes to better understand the science behind the choices made, but also sometimes to correct for misstatements, misunderstandings, or misinterpretations in the Working Group assessment reports. And while, it is usually the case that the majority of stock assessment scientists participating in the SAW are NMFS scientists, the SAW process has increasingly created opportunities for participation by other scientists outside of NOAA, such as those from state, academic and non-governmental organizations. This move towards broader participation has to be good for the process as challenges and insights brought forth from different perspectives should increase the quality, understanding, and transparency of the assessment product, but it can also increase the amount of preparation and review time involved as the different ideas are worked through.

For this assessment review, because a consensus was not reached by the SAW Working Group, much more time was needed in the review to try to find a consensus (at least among the reviewers), and furthermore, barriers developed at the review for bringing forward additional analyses to help clarify or

support findings, which also inhibited consensus. In the past, the opportunity for additional model runs, or additional data analyses, in support of assessment findings, although limited by time, often clarified points of contention and promoted more unified findings. It seemed that because opinions remained so divided among contributors at the review, that any work that could have been done in the time allotted to help clarify points of contention was not often (though not always) supported by all participants present at the meeting. While we have no suggestions as a Review Panel for resolving these two problems, some thought should be put into establishing protocols and mechanisms for facilitating consensuses in both stages of the process (that is the SAW and the SARC).

Broadly, there were two issues in terms of the science that arose in the Working Group that resulted in significant differences in interpretation of the Gulf of Maine cod assessment, different assessment results, and consequently led to lack of consensus. Issue 1 involved the use of data prior to 1982 in conducting the assessment and in determining the stock recruitment relationship based on an assessment using this data. Issue 2 involved whether or not natural mortality (M) was changing in the Gulf of Maine system. As a consensus was not reached by the SAW Working Group, two models were put forward for review: 1) an ASAP model with proxy based reference points using data from 1982-present and with M = 0.2 over the entire assessment period; and 2) an SCAA model with stock-recruit based reference points using data from 1932-present and M that transitions from 0.2 to 0.4; specifically setting M to 0.2 prior to 1989, linearly transitioning from 0.2 to 0.4 from 1989-2002, and being set to 0.4 from 2003-2011.

Issue 1: Including the Longer Data Series and Estimating Parametric Stock-Recruitment Relationships

While using information in the earlier part of the time series to help define a stock-recruitment relationship is laudable, it can be tricky. A number of concerns were raised and discussed regarding the use of the pre-1982 data (which was not of the same detail and quality as the post-1982 series) and the results from fitting the stock-recruitment curves to these data. Any one concern, by itself, might not have been enough to preclude the use of these methods in the assessment, but together these concerns led the Review Panel to discount the results and consequently the approach was eliminated from further consideration. These concerns can be examined from the point of view of the two parametric stock-recruitment models (Ricker and Beverton-Holt) and then from the point of view of the data. These concerns are outlined below:

- The F_{MSY} reference point derived from the Ricker model based on the longer data series was sometimes higher than total mortality derived from surveys suggesting that F_{MSY} estimated in this way is higher than would make sense as the stock decreased at these mortality levels. The Review Panel acknowledges that the criterion for determining survey total mortality integrates selectivity as well, but believes the above argument still holds.
- Although the Ricker model fit the longer data series better than other models (neither the Ricker or Beverton-Holt could be reasonably fit without including some other information, as that derived from the longer data series or some other external piece of prior information), the fit was clearly influenced by low recruitments in earlier years associated with high spawning stock biomass (SSB). The Review Panel could not decide if this was a period with low recruitment productivity driven by external forces or if it was a low recruitment period because of high SSB. If the low productivity had been estimated at two or more periods of high SSB then the Review Panel would have had put more consideration into the Ricker model. There was also no evidence of density dependent effects on recruitment rate such as cannibalism.

- The Beverton-Holt stock-recruitment model was similarly rejected because these low recruitment points also inflated the steepness parameter to values beyond what seemed reasonable.
- Including the earlier catch series was necessary to fit a stock recruit relationship, however, because of the above arguments and concerns about the quality and the less detailed information available in earlier part of the data series, the Review Panel concluded that these relationships were too unreliable to provide MSY reference points for characterizing assessment advice and so all model formulations (either ASAP and SCAA) that included a stock recruitment relationship were not considered further.
- Regarding the low recruitment values of the 1960s, it looked like there were other avenues that could be pursued to help validate whether or not they should be included in determining stock-recruitment model fits and associated reference point calculations. For example, examining evidence of ecosystem drivers would help determine if these recruitments were more likely to be evidence of density dependence or alternatively an environmental regime shift or a change in predation by other species. A general concern about the quality of the data in the earlier part of the series provides further motivation for examining the credibility of these influential points.
- As no standard stock-recruitment relationship could be found, the use of proxy reference points for this stock was supported.
- One other important related issue should be noted when using the Ricker or the Beverton-Holt relationships for data like these. The two models result in very different SSB_{MSY} and F_{MSY} reference points although the resulting recruitment levels at these points may be close to indistinguishable. Basing overfishing thresholds on such a volatile criterion may not be the best approach for establishing stable and sustainable management actions for stocks with this type of recruitment history.

Issue 2: Has Natural Mortality Changed Over the Years? If so, Will It Continue to Change?

One of the motivations for examining how or if M might be changing was the idea that such a change might lead to a resolution of the retrospective pattern observed in the assessment. While there seems to be more than enough information present to try to tease out this problem (e.g. tagging studies, environmental drivers, goodness of fit analyses) there remained considerable uncertainty in the estimates of M. The finding that including a changing M provides a better fit, is generally not sufficient to justify using such a model modification without other ecologically directed information to back it up. Consequently, in a way similar to what was found for the Georges Bank stock, it was unclear as to whether this factor or some other factor was what was influencing the retrospective pattern. Unlike the conclusions drawn for the Georges Bank stock, the Review Panel decided to put forward for consideration a time varying M model as this model improved the retrospective pattern, but, for the reasons stated above, the model without a time varying M seemed equally valid.

Going forward with two versions of the model has obvious consequences for reference points and the interpretation of stock status. There are a similar set of issues when making projections based on these models. If natural mortality can be assumed constant, then the usual things can be done: determine reference points and projections under steady state assumptions. If M is changing, however, then what does one use for projections? And does that mean different environmental regimes and therefore different reference points? All these issues were discussed in depth as well. For the purposes of specifying longer term projections the standard approach was recommended when M=0.2, and two alternative approaches when the mortality is ramped up. One assumes that M will remain high at M=0.4, while the other assumes that M will drop back to a base level of M=0.2 in the long term. In the short

term (~ 5 years), M values are assumed to remain the same as those for the current year in the assessment chosen (M=0.2 or 0.4 respectively).

Summary of Georges Bank Cod

Two variants of the Georges Bank cod stock assessment were presented, one assuming a natural mortality M equaling 0.2 at all ages in all years, the second assuming a 'ramping' of M through the intermediate years of the time series to a higher averaged level of 0.4 at all ages in the recent years. The second model exhibited less retrospective bias and there was some circumstantial data that it could be seen as pointing to changes in natural mortality but the evidence was inconclusive. The Review Panel believed that a bias correction using Mohn's Rho applied to the constant M assessment would allow for bias correction without giving unmerited credence to the idea that the suspected change in the system was totally due to a change in M over time. The Review Panel was keen to stress however that the decision did not indicate it believed an unchanging M was certain. Instead, it suggests additional focused research in this area (e.g. analysis of tagging data over the full time period, and studies of changes to condition factor, stomach content and predation or other factors that are likely to influence survivorship). The assessment assuming M = 0.2 with bias correction is recommended as the preferred means of assessment at this time.

Reference points were set on the assumption of M = 0.2 at all ages. It seemed that even if there had been a change in M in recent times, there was little evidence of a permanent regime shift. Consequently the long term equilibrium values of M should adequately reflect circumstances for the future. No stock-recruit relationship has been identified for this stock and so a proxy F_{MSY} is used. It was not clear to the Review Panel on what basis the F_{MSY} proxy value of $F_{40\%SPR}$ was chosen, but equally it had no basis on which to choose an alternative value.

GULF OF MAINE COD

TOR1: Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data and take into account the recommendations and subsequent work from the March 2012 MRIP workshop. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch.

A general assumption was made that the control systems to monitor catches were adequate and no concerns were raised either by the analysts or by the members of the public attending the meeting that led the Review Panel to question the validity of the catch reports.

There was no indication that important sources of catches were not accounted for.

The Review Panel considered that the documentation of results and procedures to estimate catch and their uncertainty was exceptional and very helpful.

Timeframe is an important consideration for this Term of Reference. Modern catch monitoring began in 1964. Total species landings are derived from weighout reports of commercial seafood dealers and these data are generally considered a census of total landings. While un-reported landings are possible, no estimates exist to evaluate their magnitude. No indication was given to the Review Panel that important sources of catches were not accounted for. Landings statistics for area 5 (Gulf of Maine and part of Georges Bank stocks) exist back to 1893. The methods used to apportion landings to individual stock complexes are not well documented and these early stock landings are considered less certain. Prior to 1994, port agents partitioned total cod landings to stocks through a port-interview process (< 40% of landings) or other local knowledge. Starting in 1994, the area of catch and effort information was inferred directly from vessel-reported VTRs. While there is still a potential to mis-report the area where catch was taken, since 2006 the magnitude of this error was estimated to be ≤ 2%; however, prior to 1994 the Review Panel assumed there is a greater potential error of mis-allocation of landings between the Gulf of Maine and Georges Bank stocks.

Biological sampling (length and age) of Gulf of Maine cod prior to 1982 was poor. Sampling intensity has generally increased over time and has exceeded the unofficial NAFO/ICNAF standard of 100 lengths per 200 mt since 1996. Age sampling intensity followed a similar trend. There is sufficient information to estimate the age and composition of catches from 1982 onward, and the uncertainty in these estimates (1984-2011) was derived by a bootstrap procedure and was included in the stock assessment models.

Since 1999 commercial discards (due to restrictive trip limits during 1999-2004) and recreational landings and discards have accounted for a much larger portion (25%-50%) of Gulf of Maine catches. Recreational landings peaked in 1987, but prior to 1999 they constituted only approximately 13% of the overall catch. Direct sampling of the commercial fishery for discards has been conducted by fisheries observers since 1989. Biological sampling during this period was considered to be good. The main reason for discarding was small size and this information was used when estimating the age composition of discards. Discards were hindcasted prior to 1989.

The recreational fishery has accounted for 20%-30% of the catch during 1990-2011. In this assessment, the Marine Recreational Fisheries Statistical Survey (MRFSS) data were re-estimated using revised

methodologies consistent with the new Marine Recreational Information Program (MRIP) which has replaced the MRFSS program. The MRFSS data collection program began in 1979, though estimates of recreationally caught cod are not available until 1981. The numbers-based estimates of recreational landings were converted to numbers-at-age using ALKs borrowed from the NEFSC survey which include age information collected from the inshore strata where the majority of recreational fishing occurs. Beginning in 2005 direct sampling of cod discards from party boats began in the Gulf of Maine. The length and age-distribution of discards was hindcasted prior to 2005. Recreational discard mortality was taken to be 30% and, although the discard mortality rate is highly uncertain, it is not considered to be a large assessment uncertainty because of the relatively small contribution of discards to total landings.

In previous cod assessments, discard mortality was set at 100%. In 2012, a Working Group was convened to evaluate the information available on discard mortality. At the meeting, a Delphi method was used to gather and summarize discard mortality values derived from what participants of the Working Group believed were reasonable values. These values, by gear type, are what are used in the current cod assessments. While the Review Panel felt these values were not unreasonable, we encourage further research to validate these values.

The Review Panel concluded that all elements of this Terms of Reference were thoroughly addressed. However, it is clear that the quality of catch information has improved with time. This uncertainty has been adequately characterized.

Thus, the Review Panel concludes that this term of reference was addressed adequately for the purpose of assessment.

TOR 2: Present the survey data and calibration information being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Consider model-based (e.g. GLM) as well as design-based analyses of the survey data in developing trends in relative abundance. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

Overall, the Review Panel was satisfied with the way the data were examined. However, spatial plots of survey landings by year similar to Figure A.104 in the Gulf of Maine cod assessment document but for the two stock areas combined (NEFSC Spring and Fall Surveys) would be helpful to see transboundary distributions and what effects that may have on the interpretation of stock structure, survey coverage of the stock, and the appropriate specification of stock strata to be included. Stock management boundaries on these plots should be clearly identified. Something like this was given in the GBK cod presentation, but was not provided in the report.

Both model-based and design-based analyses were considered for the surveys. While a deeper model based analysis using, for example, a GLM or a mixed-effects model GLM might be considered to address changes in spatial distribution and time trends by area, perhaps including interactions, the design-based methods, if the right strata (area, gear, season, etc.) are chosen, should address most of the issues of concern.

Commercial and recreational LPUE were explored as potential indices of abundance to be used in the assessment, however, the Working Group clearly demonstrated a number of reasons why these time series are not indicative of trends in the stock as a whole. We recognize that information potentially can still be found in this type of information and work should continue in this area to evaluate this. Some examples include, using the recreational landings to provide early recruitment indicators, examining time trends of individual boats in Coops, and comparing what the fishing fleet is "seeing" relative to what the surveys and assessments show.

The latest survey indices (Spring 2012 NEFSC survey) were the among the lowest on record, but were not used in the ASAP assessment model because of software design, although they were presented and considered in the review.

The Review Panel concluded that this term of reference was addressed adequately for the purpose of assessment.

As was similarly stated in the SARC 53 review, the SARC 55 Review Panel recommends that ongoing inspection and analysis of survey data be conducted prior to inclusion into the model. Examples of such analyses include:

- Routine internal estimates of variance of annual survey estimates.
- Inspection of relationships between age i and age i+1 within individual surveys to ensure cohorts are tracked – such analyses may help identify appropriate designation of plus groups.
- Inspection of correlations among different surveys to examine information content of individual surveys.

The Review Panel similarly notes that the Albatross IV – Henry B. Bigelow conversion factors have important consequences for the interpretation of survey data and for the assessment model. Given the high uncertainty in these conversions, we recommend that methods that do not rely on these conversion factors be implemented as soon as the length of the Bigelow time series permits.

TOR 3: Summarize the findings of recent workshops on stock structure of cod of the Northeastern US and Atlantic Canada.

The summarization of the workshop findings was thorough and met the Terms of Reference. However, as stated earlier, spatial plots of survey landings by year similar to Figure A.104 in the Gulf of Maine cod assessment document but for all hauls from the NEFSC Spring and Fall Surveys would be helpful to see transboundary distributions and what effects that may have on the interpretation of stock structure, survey coverage of the stock, and the appropriate specification of stock strata to be included. Stock management boundaries on these plots should be clearly identified. Something like this was given in the Georges Bank cod presentation, but was not provided in the report.

TOR 4: Investigate the evidence for natural mortality rates which are time- and/or age-specific. If appropriate, integrate these into the stock assessment (TOR 5).

This Term of Reference was addressed, but the evidence provided appeared to be equivocal. The pros and cons were presented on the quality and values of the natural mortality estimates and arguments were given for different natural mortality regimes over the years. As with the Working Group, the Review Panel was unable to reach a decision on which natural mortality values or time varying scenarios best characterized this system. In three days, the Review Panel did not have the time to sort through all the possible evidence supporting these issues and apparently the Working Group didn't either.

The information provided on M was based on tagging, life history information, and on total mortality from survey catch curve analysis.

It looked like there were other avenues that could be pursued including:

- Evidence of environmental drivers;
- Changes in the diet of the cod that might lead to a change in condition or spawning potential;
- Temperature preferences;
- Mechanisms influencing juvenile mortality;
- Reexamination of tagging data collected in earlier years;

TOR 5: Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Consider feasibility of survey catchability estimates, the starting year for the assessment, estimation of the stock recruitment curve, inclusion of multiple fleets, and whether to use domed or flat selectivity-at–age for the NEFSC surveys. Provide a summary of steps in the model building process. Include a historical retrospective analysis to allow a comparison with previous assessment results. Review the performance of historical projections with respect to stock size, catch recruitment and fishing mortality.

Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty.

The Review Panel concluded that while this Term of Reference was addressed in detail, a number of issues remained that complicated the derivation of a consensus assessment. In the end, multiple models were presented by the Working Group to the Review Panel and while several of these were eventually rejected for use in the development of management advice, two important variations remained. A constant M=0.2 scenario and a ramped M scenario (see Assessment Summary and SAW reports for descriptions). Given the time available, the Review Panel was unable to identify, based on best available science, which of the two should move forward, so both are being considered.

Consider

- *feasibility of survey catchability estimates,*
- *the starting year for the assessment,*
- *estimation of the stock recruitment curve,*
- *inclusion of multiple fleets*
- *whether to use domed or flat selectivity-at–age for the NEFSC surveys.*

The Review Panel concluded that all these issues were thoroughly addressed and clearly a significant amount of work went into these analyses. Nevertheless, for starting year and estimation of stock recruitment curves no consensus was reached by the Working Group. As stated above, the Review Panel decided that fitting parametric stock-recruitment curves to data that included the pre-1982 catch information remained open to question. The Working Group did arrive at findings regarding catchability, selectivity and fleet inclusion that proved reasonable to the Review Panel.

Provide a summary of steps in the model building process.

The Working Group provided a detailed overview showing the connection between the previous ASAP assessment (NEFMC 2012) and the current version. In addition a detailed overview was provided on the contrasts between the ASAP and SCAA assessments so that the magnitude of changes and consequences of differing assumptions could be seen and documented.

Include a historical retrospective analysis to allow a comparison with previous assessment results.

Retrospective analyses were provided and clearly documented.

Review the performance of historical projections with respect to stock size, catch, recruitment and fishing mortality.

Results of the performance of historical projections were provided, except for catch, but it is not clear how one might evaluate catch projections.

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

The stock is overfished and overfishing is occurring under both assessment models considered. Reference points were recalculated for $F_{40\%\ SPR}$ as a basis for stock determination. The Review Panel, in conjunction with Working Group members present, revised the Mramp SSB_{MSY} reference points during the meeting. The other reference points presented by the Working Group were accepted by the Review Panel.

The Review Panel emphasizes that the recommendation to maintain an $F_{40\%}$ basis for reference point determination was based on the lack of a consistent stock recruit relationship. We do not suggest that $F_{40\%}$ is necessarily the best proxy to use, rather there has yet to be compelling reasons to abandon it.

Two alternative formulations were put forward, and the Mramp formulation itself had several alternative configurations for projections. We recognize that this greatly complicates the question of making projections and identifying Frebuild and the rebuilding schedules. But, it is difficult to expect that a Review Panel will be able to sort out what a Working Group and assessment scientists were not able to sort out through months if not years of work.

TOR 7: Evaluate stock status with respect to the existing model (from the most recent accepted peer reviewed assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt. 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 (from Cod TOR-6).

Based on the previous reference points, the previous assessment indicated that the Gulf of Maine cod stock was overfished and overfishing was occurring.

Based on the reference points derived from the two new models, the new assessments indicate that the Gulf of Maine cod stock is still assessed to be overfished and overfishing is occurring.

The Review Panel notes a long history of this stock experiencing overfishing.

TOR 8: Develop and apply analytical approaches to conduct single and multi-year stock projections to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. **Provide numerical annual projections (3-5 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).**

Short-term projections were provided using the same stochastic projection method used for the reference point calculations (i.e. same biological parameters, MCMC for survivors in 2011, and resampling of recruitment from an empirical CDF based on historic values, with a ramp to zero for SSB's below the minimum in the assessment time-series). This procedure accounts for uncertainties in terminal year abundance and variability in recruitment. However, only projection medians were provided. Annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass, were not provided although presumably these results exist. A sensitivity analysis to assumptions about M (i.e. M=0.2 or M ramp from 0.2 to 0.4) was provided, and for the Mramp scenario the projections were provided assuming that M remained at 0.4 or that M returns to 0.2 in the projection period.

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.**

The Review Panel concluded that the M=0.2 projections and the Mramp projections with M remaining at 0.4 in the short-term were equally realistic. Little evidence was presented to the Review Panel to favor

one scenario over the other. The Working Group could not decide which option was more plausible and neither could the Review Panel. The Panel concluded that if M is currently 0.4 then it seemed more reasonable to assume that in the short-term M would remain at 0.4 rather than reduce to 0.2. Note that for long-term projections that Review Panel decided that M should be 0.2, because the longer-term historical evidence seems to indicate that M=0.2 is more plausible.

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

The Review Panel appreciated the description of the stock's vulnerability to becoming overfished. We emphasize that since the mid-2000s, the fishery has become particularly concentrated in a small region of the western Gulf. The most recent survey indices are at or near the lowest values in their time series and there are concerns the industry will not be able catch their full quota. The available information points to a stock at a low level and with a concentration of the remaining stock into a relatively small region of the western Gulf, the vulnerability of the stock is likely to be increased.

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

The Review Panel appreciated both the amount of progress and the reporting of progress on previous research recommendations. Of the nine research recommendations brought forward from SARC 53, six were either partially or fully addressed. A single recommendation was carried forward from GARM III which was addressed in the Working Group report. The GARM III research recommendation involved using historical data to hindcast recruitments as far back in time as possible for use in the estimation of reference points and projections. This was addressed in both the SCAA and ASAP models. However, the Review Panel notes that there are additional complications due to temporal non-stationarity that can occur when using long time-series of stock-recruitment estimates or estimates of other components of stock productivity. Many factors in addition to parental stock size can influence how much recruitment is produced, and these factors can change over time which introduces non-stationarity. This is an additional complication that should be accounted for when estimating reference points and projections.

Assumptions about natural mortality have important implications on the stock assessment and management advice. A SARC53 research recommendation involved evaluating the level, schedule and variability of natural mortality.

The Review Panel recommends in addition:

1. Provide analysis on changes in the location and quality of preferred environment and habitats for cod and potential implications on M (adult and juvenile) and spawning potential.

2. Telemetry tagging may provide a more direct way to measure natural mortality, particularly if there are local cod populations with high site fidelity.

3. Consider other assessment models that include 'smoothing' approaches (e.g. penalized random walks) to deal with changes in fishery selectivity and natural mortality.

4. Consider accounting for residual patterns and retrospective patterns using process errors. A rationale for this is that process errors can be projected into the future to potentially better account for the model/process uncertainty (indicated by residual and retrospective patterns) in projections and MSY reference points. The current approach of retrospective correcting for process error does not seem sufficient particularly in long-term projections for rebuilding analyses and reference point calculations. Uncertainty in calibrations to standardize survey time series for changes in vessels and fishing gear (i.e. doors) was not accounted for in the stock size indices. This may be a useful area for future research, although hopefully the time-series will soon be long enough that direct calibration will not be required.

5. A GLM approach could be used to combine NEFSC and MADMF survey indices into two more complete indices for the Spring and Fall. The NEFSC surveys have better coverage in offshore strata, and the MADMF surveys had better coverage in inshore strata. Combining surveys would result in better coverage of the whole stock and hopefully better stock size indices.

6. As part of the model building exercise, consider summarizing the information about mortality rates and trends in stock size using a survey-only assessment model such as SURBA. This could replace catch-curve estimation of Z's. It can also be used to explore conflict (or lack thereof) between surveys and catches

7. When stock-recruit data are uncertain but the time-series is long, consider constraining Rmax to be some reasonable value (e.g. maximum of historic assessment values) and derive MSY reference points using the constrained stock-recruit curve. There are nonparametric approaches that could be used to address sensitivity of MSY reference points to simple parametric assumptions about stock-recruitment relationships.

Georges Bank Cod

TOR 1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data and take into account the recommendations and subsequent work from the March 2012 MRIP workshop. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch.

A general assumption was made that the control systems to monitor catches were adequate and no concerns were raised either by the analysts or by the members of the public attending the meeting that led the Review Panel to question the validity of the catch reports. There was no indication that important sources of catches were not accounted for.

The Review Panel considered that the documentation of results and procedures to estimate catch and their uncertainty was exceptional and very helpful.

Prior to 1994, information of the catch quantity was derived from reports of landings transactions submitted voluntarily by processors and dealers. More detailed data on fishing effort and location were obtained for a subset of trips from personal interviews of fishing captains conducted by port agents. This information was used to augment the total catch information obtained from dealers. Starting in 1994 the area of catch and effort information was inferred directly from vessel-reported VTRs. The uncertainty in allocation of landings to Georges Bank and Gulf of Maine cod stock areas is considered by the WG to be little to no consequence.

Atlantic cod discarded on Georges Bank by the USA commercial fisheries were estimated from 1989-2011 observer data and 2010-2011 at- sea monitoring data. Estimates of discards in the large mesh otter trawl fishery during 1978-1988 were hindcasted using a survey filter method. 'Delphi' determined mortality rates were to be applied to the final estimates of USA discards. Discards in Canadian fisheries have been estimated using various methods. Discards have represented about 5% of the USA commercial and 9% of the Canadian catch on average.

USA recreational landings and discards were estimated using MRFSS data from 1981-2003 and MRIP data from 2003-2011. Recreational catch accounts for 1%-10% of the total catch since 1981.

In the USA fishery, sampling intensity by market category has improved since 1978 and has been relatively high since 2003. Sampling intensity in the Canadian fishery has also been good since 2003. There is sufficient information to estimate the age and composition of catches from 1978 onward, and the uncertainty in estimates for 2003-2011 was derived by a bootstrap procedure and was included in the stock assessment models.

The age and size composition of cod discarded in the USA commercial fishery were estimated for 1989-2011 using combined survey and commercial age-length keys and observer length frequency data. The age and size composition of discards for 1978-1988 were estimated using hindcasted discards at length for large mesh otter trawls and autumn research survey proportions at age. Discards from the Canadian groundfish fishery were assumed to have the same size and age composition as the fishery landings. The size composition of discards from the Canadian scallop fishery was estimated using observer length

frequency and age data. The commercial discards are generally dominated by age 2 and age 3 fish during the time series.

The number of length samples taken in the recreational fishery was insufficient to estimate the landings at age. A combined commercial and survey age-length key and research survey length frequencies and length-weight were used to estimate recreational landings and discards at age for 1981-2011. Landings and discard length frequencies were differentiated by applying a length cutoff to the survey length frequency. The recreational catch estimates are dominated by ages 4-5 in the landings component and ages 2-3 in the discard component in recent years

The Review Panel concluded that all elements of this Term of Reference were thoroughly addressed. However, it is clear that the quality of catch information has improved with time. Uncertainty in age-compositions has been partially characterized, only for USA commercial landings during 2003-2011.

Thus, the Review Panel concludes that this term of reference was addressed adequately for the purpose of assessment.

TOR 2. Present the survey data and calibration information being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Consider model-based (e.g. GLM) as well as design-based analyses of the survey data in developing trends in relative abundance. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

This term of reference was addressed adequately for the purposes of the assessment.

The model-based analyses of the survey data were presented but mainly in the presentations to the SARC Review Panel and little was said about them in the assessment report. However, it was acknowledged that to use the estimates from the GLM model in the assessment without accounting for the uncertainty in the GLM would lead to over smoothing of the values and that to best reflect uncertainty in the survey data the design-based indices should be used.

The reasons for not using the commercial LPUE index make sense, namely lack of Canadian data, which make up approximately 25% of the landings, significant regulatory changes since 1994 and associated spatial shifts in the fishery, and implementation of sector-based management since May 2010.

The reasons for not using the recreational LPUE index in the assessment also make sense, namely the uncertainty over the unit of measurement in the early years, restricted geographical coverage of the sector and only a small number of vessels that have contributed to the time series consistently.

It would have been good to see scatter plots of logged survey indices and line plots of standardized survey numbers by age as suggested by one of the SARC 53 review panelists with respect to Gulf of Maine cod. These do give a quick visual means to assess the ability of surveys to track cohorts.

TOR 3. Summarize the findings of recent workshops on stock structure of cod of the Northeastern US and Atlantic Canada.

This term of reference was addressed adequately.

The presentation of findings of the workshop was thorough. However, this is an area where more research would be appropriate and indeed it was stated to the SARC 55 that work investigating stock structure is ongoing. Suggestions for possible investigations are

- Producing geospatial smoothes of the survey CPUE values to confirm (or refute) consistent segregation of Gulf of Maine and GB cod concentrations at time of spawning.
- Investigate evidence for asynchrony between the Gulf of Maine and GB areas through use of a pairwise test after Holmes et al. (2008). That is SSB and recruitment indices can be log-transformed and the resultant trends for the different areas compared by fitting a GAM to the ratio between them. The resulting smoother is then compared to a constant fit of this ratio by a standard F-test. Investigation of sub-stock (metapopulation) structure can be tested by forming indices for each putative sub-population, applying a smoother to all indices and comparing to a model in which the smoothers were constrained to be parallel.

TOR 4. Investigate the evidence for natural mortality rates which are time- and/or age-specific. If appropriate, integrate these into the stock assessment (TOR 5).

This term of reference was addressed adequately for the purposes of the assessment, however a number of areas of further work seemed apparent.

An explanation for the decrease through time of the condition index in spring may be relevant to this question. Has the reduction in Spring condition been enough to affect survivorship? There is also clear evidence for a reduction over time in mean weight at age in this stock. If the cause can be identified, it may shed light on possible mechanisms for a change in natural mortality or other factors affecting stock health such as the stock-recruit relationship (through poorer condition fish producing less viable eggs).

TOR 5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Consider feasibility of survey catchability estimates, the starting year for the assessment, estimation of the stock recruitment curve, inclusion of multiple fleets, and whether to use domed or flat selectivity-at–age for the NEFSC surveys. Provide a summary of steps in the model building process. Include a historical retrospective analysis to allow a

comparison with previous assessment results. Review the performance of historical projections with respect to stock size, catch recruitment and fishing mortality.

This term of reference was addressed adequately for the purposes of the assessment.

The steps from previous ADAPT model, through a VPA like ASAP model to the adoption and specification of the base case and Mramp ASAP models is well described. The rational for the selection of flat topped selectivity curves for the surveys and fishery is clear and it seems evident that a reliable stock-recruitment curve could not be estimated because of a lack of contrast in the SSB and recruitment data.

The retrospective analysis figures for the base case and Mramp runs were very useful for they show that although the Mohn's rho values for the Mramp run are much lower than for the base run, the Mramp model formulation has produced considerable retrospective patterning. Given the end points in the runs to 2009 and 2010 in the Mramp version it is questionable whether even a version employing the bias correction suggested would have handled the retrospective issue.

The Review Panel considered bias correction using Mohn's Rho on the model with constant M would correct the bias without giving unmerited credence to a change in M over time. The Review Panel was keen to stress however that the decision did not indicate it believed an unchanging M was certain or that the true value for M was M=0.2.

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

The Term of Reference was met, however, the existing definitions for "overfished" and "overfishing" are not stated in the executive summary under Term of Reference 6 which would be useful.

The Review Panel accepted that parametric stock recruitment relationships could not be obtained for this stock but no scatter plot of the stock-recruit relationship is included in the report. This together with the parametric fits normally associated with cod (Ricker and Beverton-Holt) should be included given the decision to use non-parametric biological reference points.

Reference points were calculated using $F_{40\%SPR}$ for stock status determination. The justification for choosing a given F%SPR was not entirely clear to the Review Panel from the documentation available, however, $F_{40\%SPR}$ appears to have based on the GARM III review of 19 groundfish stocks. There were no compelling reasons provided to choose a value that differed from $F_{40\%SPR}$.

TOR 7. Evaluate stock status with respect to the existing model (from the most recent accepted peer reviewed assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt.

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 (from Cod TOR-6).

The Term of Reference was met. Whether based on the previous assessment model, or the new assessment model, inclusion of the latest data did not change the conclusion that the stock is overfished and overfishing is taking place.

TOR 8. Develop and apply analytical approaches to conduct single and multi-year stock projections to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3-5 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

Term of Reference (a) was partially met. Stochastic projections were provided based on the final accepted model fit and the new reference points provided under Term of Reference 6. Only projection medians were provided. The annual probabilities of exceeding threshold BRPs for F and probabilities of falling below threshold BRPs for biomass were not provided as stock biomass is not likely to exceed the reference threshold biomass in the short term.

Term of Reference (b) was addressed by evaluating projections where M = 0.2 in both short and long term, M = 0.2 in the long term but with M=0.4 in the short term and M = 0.4 in both short and long term. The Review Panel concluded that in the short term, for the M=0.2 scenario projections should be based on an M=0.2, whereas in the scenario where M transitions from 0.2 to 0.4 short term projections should be based on an M=0.4. The Review Panel concluded that in the long term M should be 0.2 because the long term equilibrium should reflect circumstances for the majority of the historical time series. The projection based on M = 0.2 was chosen by the Review Panel because it

was not certain the restrospective bias seen in the assessment in recent years could be attributed to a higher M; a projection based on M = 0.2 and with starting values from the bias corrected base ASAP model would account for the bias without a need to accept higher M values representing the true state of nature and this approach could also accommodate a different, as yet unidentified and transitory affect.

Term of Reference (c) was met. A description of vulnerability issues, including current truncated age structure, evidence for low hatching rate for first and second time spawners and the two decades of poor recruitment for this stock was included in the assessment report. The considerations are in this case the vulnerability to continued overfishing as the stock is currently considered overfished.

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

There were three previous research recommendations (from GARM III) for Georges Bank cod. One was dealt with directly, one was examined using the synergies between the assessments of Georges Bank and Gulf of Maine cod, and the last was identified by the Working Group as under examination via an ongoing NOAA project.

Continued exploration of retrospective pattern and methods to account for it.

This Term of Reference was addressed through exploration of the Mramp version of the ASAP model and also one employing a catch multiplier.

Historical data be used to hindcast recruitment estimates as far back in time as possible for use in the estimation of reference points and projections.

The Working Group did not conduct the hindcasting citing the results and conclusions of the Gulf of Maine cod analysis (ASAP model) and the perceived quality of Georges Bank catch data prior to 1978. If two of the recommendations for new research from the Working Group (the last two recommendations given below) are completed it would seem relevant to re-visit the hindcasting idea.

Investigate the effect of uncertainty in maturity at age in the estimation of SSB_{MSY}. Research into incorporating trends in biological parameters (weights, maturity) into projections methodology.

The Working Group cited the NOAA funded FATE (Fisheries and the Environment) proposal as a method to address this area of research.

The research recommendations of the SAW 55 Working Group were as follows (those marked with a '*' are recommendations that also apply to Gulf of Maine cod):

- To further address the retrospective issue:

- Conduct 'forensic accounting' analysis of 'missing catch' i.e. lost/unreported VTRs, lost/unreported dealer data, underestimated discards.
- 100% observer coverage (for 3-5 years) of the fisheries that either target GB cod or have cod as bycatch to ascertain potential underestimation of GB cod discards.
- Conduct designed discard mortality study of cod that pass through the trawl via trouser trawl experiment, including blood analysis to determine stress levels compared to control group*.

These are all reasonable recommendations. Whether 100% observer coverage is financially possible is a question for managers.

- Inclusion of the tagging analysis formally within the stock assessment model*.

This line of research is described as a longer term project and it would seem there are more pressing issues to address for both the Georges Bank and Gulf of Maine cod assessments.

- Explore the appropriate weighting of the proportions at age data (constant versus age specific)*.

Differences in proportions at age could have a significant impact on estimates of total mortality and it is therefore important to ensure the best means possible are being used for their calculation.

- Incorporating the Bigelow/Albatross calibration coefficients within the assessment model so that coefficients can be re-estimated as data on year-classes is updated*.

As such calibration coefficients tend to be very uncertain it seems sensible to maximise the amount of data available for their estimation. Hopefully the Bigelow time-series will soon be long enough that direct calibration will not be required.

- Exploration of a random errors approach to the internal fitting of stock – recruitment relationships*.
- Simulations (conditioned on data) of the internal estimation of stock - recruitment functions to explore potential bias in the fitting of these relationships*.

Given the lack of an accepted parametric stock-recruit relationship and the debate over use of older catch data to help define such a relationship, this line of research would be very valuable.

In addition to the above it is further recommended:

- Re-visit the stomach data available to NEFSC to see if it contains enough information to explain
    1. The drop in spring condition over time.
    2. The reduction in mean weights at age over time.

The stomach data were used to consider predation on cod (in addressing Term of Reference 4) but may contain evidence to help explain these outcomes as well. Comparison between stomach data taken from Georges Bank and Gulf of Maine cod could help with ongoing work into defining stock structure

(dependant on the geographic uniqueness of prey items) and may point to factors that make one or other stock more vulnerable to changes in environmental factors such as water temperatures.

- Provide analyses on whether there have been changes in the location and quality of preferred habitat for cod on Georges Bank. If evidence is found to then consider the implications for straying (i.e. inter-stock mixing), M and spawning potential.

- Conduct telemetry tagging studies to obtain current estimates of M.

BIBLIOGRAPHY

The following material was provided to the Review Panel to assist it in meeting its terms of reference

References

Alexander KE et al. 2009. Gulf of Maine cod in 1861: historical analysis of fishery logbooks, with ecosystem implications. Fish and Fisheries, 2009, 10, 428-449. 22 p.

Ames EP. 2004. The Stock Structure of Atlantic cod in the Gulf of Maine. Fisheries, 29:10-28

Bell E. 2012. 53rd Northeast Regional Stock Assessment Review Committee (SARC 53) Report from Center of Independent Experts (CIE). SAW/SARC 53 Panelist Reports. 39 p.

Brander K, Mohn R. 2004. Effect of the North Atlantic Oscillation on recruitment of Atlantic cod (*Gadu morhua*). Abstract. Can J. Fish. Aquat. Sci. 61: 1558-1564

De Valpine P, Hastings A. 2002. Fitting Population Models Incorporating Process Noise and Observation Error. Ecological Monographs, 72(1), 2002, 57-76. 20 p.

Fogarty M, Incze K, Hayhoe K, Mountain D, Manning J. 2008. Potential climate change impacts on Atlantic cod (*Gadus morhua*) off the northeastern USA. Abstract. Mitig Adapt Strat Glob Change (2008) 13: 453-466.

Fogarty MJ, Myers RA, Bowen KG. 2001. Recruitment of cod and haddock in the North Atlantic: A comparative Analysis. Abstract. ICES Journal of Marine Science, 58: 952-961.

Gro"ger JP, Fogarty MJ. 2010. Broad –scale climate influences on cod (*Gadu morhua*) recruitment on Georges Bank. Abstract. ICES Journal of Marine Science; 10.1093

Gulf of Maine Research Institute (GMRI). 2012. Report of the Workshop on Stock Structure of Atlantic Cod in the Gulf of Maine Region, June 12-14, 2012. 26 p.

Haltuch MA, Hicks A, See K. National Marine Fisheries Service (NMFS). Report on the Status of the U.S. petrale sole resource in 2012. Northwest Fisheries Science Center. 2011. 389 p.

Hunt et al. 1999. Movement of *Gadus morhua* in the Gulf of Maine area. Fish. Bull. 97(4).19 p.

Mayo RK, Hesler TE, O'Brien L, Sosebee KA, Figuerido BF, Hayes DB. 1994. Estimation of standardized otter trawl effort, landings per unit effort, and landings-at-age for Gulf of Maine and Georges Bank cod. NMFS NEFSC Ref. Doc. 94-12. 22 p.

Meredith E. 2012. Utility of Catch and Landings per Unit of Fishing Effort (CPUE and LPUE) in Gulf of Maine and Georges Banks cod Stock Assessments. Report from workshop on August 21, 2012 in Gloucester, MA. Northeast Fisheries Science Center (NEFSC). 14 p.

Methot RD Jr, Taylor IG. 2011. Adjusting for bias due to variability of estimated recruitments in fishery

assessment models. Can. J. Fish. Aquat. Sci. 68: 1744-1760.17 p.

Miller TJ. Regional migration and mortality of Atlantic cod in the Northwest Atlantic Ocean from tag-recovery data, accounting for short-term tagging-induced fatalities. SARC 55 Model Meeting Working Paper 31. Northeast Fisheries Science Center (NEFSC). 21 p.

Miller TJ, Bell E, Patterson K, Trzcinski MK. 2012. 53rd Northeast Regional Stock Assessment Review Committee (SARC 53) Summary Report from Center of Independent Experts (CIE). SARC 53 Panelist Reports. 33 p.

Miller TJ, Das C, Politis PJ, Miller AS, Lucey SM, Legault CM, Brown RW, Rago PJ. 2010. Estimation of Albatross IV to Henry B. Bigelow Calibration Factors. NMFS NEFSC Ref Doc. 10-05; 237 p.

National Research Council, 1998. Improving Fish Stock Assessments. The National Academies Press.

NEFSC (Northeast Fisheries Science Center), 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 4-8, 2005. NMFS NEFSC Ref. Doc. 08-15. 884 p.

NEFSC (Northeast Fisheries Science Center), 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.

NEFSC (Northeast Fisheries Science Center), 2012. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Summary Report. NEFSC Ref Doc. 12-03; 39 p.

NEFSC (Northeast Fisheries Science Center), 2012. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Report. NEFSC Ref Doc. 12-05; 559 p.

Nitschke P. Georges Bank Cod Envelope Analysis. SARC 55 Working Paper 35. Northeast Fisheries Science Center (NEFSC). 1p.

O'Brien L, Worcester T, editors. 2009. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Transboundary Resources Assessment Committee Eastern Georges Bank cod benchmark assessment. TRAC Proceedings 2009/02.

Palmer, MC, O'Brien L, Wigley S, Mayo R, Rago P, Hendrickson L. 2008. A brief overview of discard estimation methods where observer coverage is unavailable. Groundfish Assessment Review Meeting (GARM) Reference Points Meeting. April 29 – May 2, 2008. Working Paper 4.5. 13 p.

Paloheimo JE. 1968. Analysis of the Southern Gulf of St. Lawrence Cod Population. J. Fish. Res. BD. Canada, 25(3) : 555-578, 1968. 24 p.

Paloheimo JE. 1961. Studies on Estimation of Mortalities I. Comparison of a Method Described by

Beverton and Holt and a New Linear Formula. J. Fish. Res. BD. Canada, 18(5), 1961. 18 p.

Patterson K. 2012. 53rd Northeast Regional Stock Assessment Review Committee (SARC 53) Report from Center of Independent Experts (CIE). SARC 53 Panelist Reports. 37 p.

Porter JM, O'Brien L, editors. 2011. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Transboundary Resources Assessment Committee for Eastern Georges Bank Cod. Report of Meeting held 21-24 June 2011. TRAC Pro 2011/01.

Puvanendran V, Laurel BJ, Brown JA. 2008. Cannibalism of Atlantic cod *Gadus morhua* larvae and juveniles on first-week larvae. Aquat. Biol. Vol 2: 113-118, 2008. 6 p.

Rademeyer RA, Butterworth DS, Plaganyi. 2008. African Journal of Marine Science. 30(2): 263-290. 28 p.

Rago P. Envelope Analyses for Georges Bank cod. SARC 55 Working Paper 36. 30 p.

Scientific and Statistical Committee (SSC). 2012. Response to New England Fishery Management Council. Topic: Gulf of Maine (GOM) cod Assessment. 4 p.

Scientific and Statistical Committee (SSC). 2012. Response to New England Fishery Management Council. Topic: Input to the Stock Assessment Review Committee (SARC) 53 Terms of Reference. 2 p.

Sheperd G. An alternative approach to GB cod tagging analysis. Northeast Fisheries Science Center (NEFSC) SARC 55 Working Paper 28. 12 p.

Singer L, Meredith E. 2012. Establishing Discard Mortality Rates for Atlantic Cod Stock Assessments Using a Modified Delphi Technique. Recommendations of Workshop on July 24, 2012 in Mansfield, MA. New England Fisheries Science Center (NEFSC). 22 p.

Smith BE, Link JS. 2010. The Trophic Dynamics of 50 Finfish and 2 Squid Species on the Northeast US Continental Shelf. National Marine Fisheries Service, Northeast Fish Sci Cent. NOAA Technical Memorandum NMFS-NE-216. 646 p.

Tallack S. 2006. Report of the Northeast Regional Cod Tagging Program to the National Marine Fisheries Service. Gulf of Maine Research Institute (GMRI). 100 p.

TRAC. 2011. Eastern Georges Bank Cod. TRAC Status Report 2011/02.

Trzcinski, MK. 2012. 53rd Northeast Regional Stock Assessment Review Committee (SARC 53) Report from Center of Independent Experts (CIE). SAW/SARC 53 Panelist Reports. 33 p.

Wang Y, O'Brien L, Clark K, Hatt B. 2011. Assessment of Eastern Georges Bank Atlantic Cod for 2011. TRAC Reference Document 2011/02. 9 p.

Wang Y, O'Brien L, Gavaris. 2009. Transboundary Resources Assessment Committee. Benchmark Assessment Review for Eastern Georges Bank Cod. TRAC Reference Document 2009/07. 114 p.

Wigley SE, Rago PJ, Sosebee KA, Palka DL. 2007. The Analytic Component to the Standardized Bycatch Reporting Methodology Omnibus Amendment: Sampling Design, and Estimation of Precision and Accuracy (2nd Edition). NMFS NEFSC Ref. Doc. 07-09. 156 p.

Wigley SE, Hersey P, Palmer JE. 2008. A description of the allocation procedure applied to the 1994 to 2007 commercial landings data NMFS NEFSC Ref. Doc. 08-18. 61 p.

Working Group on Re-Evaluation of Biological Reference Points for New England Groundfish. 2002. Re-evaluation of biological reference points for New England groundfish. Northeast Fish. Sci. Cent. Ref. Doc. 02-04; 395 p.

Working Papers

Working Group, Stock Assessment Workshop (SAW 55) 2012. Stock Assessment Report of Georges Bank (GBK) Cod. Working Paper #1. SAW/SARC 55. December 3-7, 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 55) 2012. Stock Assessment Report of Gulf of Maine (GOM) Cod. Working Paper #1.SAW/SARC 55. December 3-7, 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 55) 2012. Stock Assessment Summary Report of Georges Bank (GBK) Cod. Working Paper #2. SAW/SARC 55. December 3-7, 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Working Group, Stock Assessment Workshop (SAW 55) 2012. Stock Assessment Summary Report of Gulf of Maine (GOM) Cod. Working Paper #2.SAW/SARC 55. December 3-7, 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

Summary Reports from the Cod SAW55 WG meetings written by Robert O'Boyle (SAW WG Chair)

Report of the Data Meeting, Aquarium Building, S.H. Clark Conference Room, Woods Hole, 27 – 31 August 2012, 18 p.

Report of the Models Issues Meeting, Aquarium Building, S.H. Clark Conference Room, Woods Hole, 15 – 19 October 2012, 19 p.

Modeling and Reference Points Meeting, Aquarium Building, S.H. Clark Conference Room, Woods Hole, 29 October – 2 November 2012, 22 p.

APPENDIX 1

Task Order T36-01, final 03 October 2012 (revised 13 November 2012)

Statement of Work

55th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for Georges Bank cod and Gulf of Maine cod

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 is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication. The purpose of this panel review meeting will be to provide an external peer review of stock assessments for Georges Bank cod and Gulf of Maine cod. Atlantic cod, *Gadus morhua*, is a demersal gadoid species found on both sides of the North Atlantic. In U.S. waters, cod are assessed and managed as two stocks: Gulf of Maine, and Georges Bank and southward. Both stocks support important commercial and recreational fisheries. The last peer reviewed benchmark assessment of Gulf of Maine cod was in 2010 as part of SARC 53. The last peer reviewed assessment update of Georges Bank cod took place in 2012. The SARC 55 review panel will be composed of three independently appointed reviewers, and an independent chair from the Science and Statistical Committee (SSC) of the New England or MidAtlantic Fishery Management Council. The SARC panel will write the SARC Summary Report and each reviewer will write an individual independent review report. 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.

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 MidAtlantic 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 stock assessment Terms of Reference (ToRs) are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is described in **Annex 4.**

Requirements for the reviewers: Three reviewers shall conduct an impartial and independent peer review of the Georges Bank cod and Gulf of Maine cod 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 methods. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points that includes an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 55 will address fishery stock assessments of Georges Bank cod and Gulf of Maine cod, therefore familiarity with forward projecting models and estimation used for North Atlantic stocks including cod stocks off North

America and Europe is desirable.

PERIOD OF PERFORMANCE

The period of performance begins on the award date, and the contractor shall complete the tasks and deliverables as specified in 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 December 3-7, 2012.

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, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.

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 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, and FAX number) 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, 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 (not by email) the requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, whether there is dual citizenship, passport number, 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 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 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. 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. 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 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 alternative assessment models and model assumptions were presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.

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. If alternative assessment models and model assumptions were presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. 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 during December 3-7, 2012 (Tuesday through Saturday).
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than December 21, 2012, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to Dr. David Die ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment ToR in **Annex 2**.

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

October 12, 2012	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
November 19, 2012	NMFS Project Contact will attempt to provide reviewers the pre-review documents
December 3-7, 2012	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
December 7, 2012	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
December 21, 2012	Reviewers submit draft independent peer review reports to the contractor's technical team for independent review
December 21, 2012	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
December 28, 2012	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
January 3, 2013	Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements
January 6, 2013	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

National Marine Fisheries Service, NOAA

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. If alternative assessment models and model assumptions were presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.

   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: 55th SAW/SARC Stock Assessment Terms of Reference

A. Gulf of Maine cod stock

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data and take into account the recommendations and subsequent work from the March 2012 MRIP workshop. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch.

2. Present the survey data and calibration information being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Consider model-based (e.g. GLM) as well as design-based analyses of the survey data in developing trends in relative abundance. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

3. Summarize the findings of recent workshops on stock structure of cod of the Northeastern US and Atlantic Canada.

4. Investigate the evidence for natural mortality rates which are time- and/or age-specific. If appropriate, integrate these into the stock assessment (TOR 5).

5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Consider feasibility of survey catchability estimates, the starting year for the assessment, estimation of the stock recruitment curve, inclusion of multiple fleets, and whether to use domed or flat selectivity-at–age for the NEFSC surveys. Provide a summary of steps in the model building process. Include a historical retrospective analysis to allow a comparison with previous assessment results. Review the performance of historical projections with respect to stock size, catch recruitment and fishing mortality.

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

7. Evaluate stock status with respect to the existing model (from the most recent accepted peer reviewed assessment) and with respect to a new model developed for this peer review. In both

cases, evaluate whether the stock is rebuilt.

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 (from Cod TOR-6).

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

    a. Provide numerical annual projections (3-5 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

B. Georges Bank cod stock

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data and take into account the recommendations and subsequent work from the March 2012 MRIP workshop. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch.

2. Present the survey data and calibration information being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Consider model-based (e.g. GLM) as well as design-based analyses of the survey data in developing trends in relative abundance. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

3. Summarize the findings of recent workshops on stock structure of cod of the Northeastern US and Atlantic Canada.

4. Investigate the evidence for natural mortality rates which are time- and/or age-specific. If appropriate, integrate these into the stock assessment (TOR 5).

5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Consider feasibility of survey catchability estimates, the starting year for the assessment, estimation of the stock recruitment curve, inclusion of multiple fleets, and whether to use domed or flat selectivity-at–age for the NEFSC surveys. Provide a summary of steps in the model building process. Include a historical retrospective analysis to allow a comparison with previous assessment results. Review the performance of historical projections with respect to stock size, catch recruitment and fishing mortality.

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

7. Evaluate stock status with respect to the existing model (from the most recent accepted peer reviewed assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt.

    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 (from Cod TOR-6).

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

    a. Provide numerical annual projections (3-5 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

Annex 2 (cont.):

Appendix to the Assessment TORs:

Explanation of "Acceptable Biological Catch" (DOC Natl. Standard Guidelines, Fed. Reg., vol. 74, no. 11, 1/16/2009):

> *Acceptable biological catch (ABC)* is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*
>
> *ABC for overfished stocks.* For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*
>
> NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)
>
> ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

Explanation of "Vulnerability" (DOC Natl. Standard Guidelines, Fed. Reg., vol. 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.

Annex 3: DRAFT Agenda

55th Northeast Regional Stock Assessment Workshop (SAW 55)

Stock Assessment Review Committee (SARC) Meeting

December 3-7, 2012

Stephen H. Clark Conference Room – Northeast Fisheries Science Center

Woods Hole, Massachusetts

Draft AGENDA* (version: 3 Oct. 2012)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR
Monday, Dec. 3			
1 – 1:30 PM			
Welcome	**James Weinberg**, SAW Chair		
Introduction	**Patrick Sullivan**, SARC Chair		
Agenda			
Conduct of Meeting			
1:30 – 3:15	Assessment Presentation (B. GBK COD)		
	Loretta O'Brien	**TBD**	**TBD**
3:15 – 3:30	Break		
3:30 – 4:45	Assessment Presentation (B. GBK cod)		
	Loretta O'Brien	**TBD**	**TBD**
4:45 – 6:00	SARC Discussion w/ Presenter (B. GBK cod)		
	Patrick Sullivan, SARC Chair	**TBD**	**TBD**

Tuesday, Dec. 4

8:30 – 9:30 (cont.) SARC Discussion w/ presenter (B. GBK COD)
Patrick Sullivan, SARC Chair

9:30 – 9:45 Break

9:45 – 10:15 Assessment Presentation (A. GOM COD)
Robert O'Boyle **TBD** **TBD**

10:15 – Noon Assessment Presentation (A. GOM COD)

Mike Palmer **TBD** **TBD**

Noon – 1:00 Lunch

1:00 – 2:15 Assessment Presentation (A. GOM COD)

Mike Palmer **TBD** **TBD**

2:15 – 3:15 Assessment Presentation (A. GOM COD)

Doug Butterworth **TBD** **TBD**

3:15 – 4:15 Assessment Presentation (A. GOM COD)
Robert O'Boyle **TBD** **TBD**

4:15 – 4:30 Break

4:30 – 6:15 SARC Discussion w/ presenters (A. GOM COD)

Patrick Sullivan, SARC Chair **TBD**

7:15 Social Event **--LocationTBD**

Wednesday, Dec. 5

9 - 11 Revisit w/ presenters (B. GBK COD)

Patrick Sullivan, SARC Chair **TBD**

11 – 11:15 Break

11:15 – 12:15 Revisit w/ presenters (A. GOM COD)

Patrick Sullivan, SARC Chair **TBD**

12:15 – 1:30 Lunch

1:30 – 2:45 (cont.) Revisit w/ presenters (A. GOM COD)

Patrick Sullivan, SARC Chair **TBD**

2:45 -3:00 Break

3:00 – 6:00 Review/edit Assessment Summary Report (A. GOM COD)

Patrick Sullivan, SARC Chair **TBD**

Thursday, Dec. 6

9 - 12 Review/edit Assessment Summary Report (B. GBK COD)

Patrick Sullivan, SARC Chair **TBD**

12 – 1:15 **Lunch**

1:15 – 5 Wrap up Assessment Summary Reports (A. and B as necessary)

Patrick Sullivan, SARC Chair **TBD**

3:00-3:15 **Break**

3:15 – 5:30 SARC Report Writing (Closed Meeting)

Friday, Dec. 7

9:00 - 3 PM (cont.) SARC Report writing. (closed meeting)

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

Annex 4: Contents of SARC Summary Report

1.

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

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

If alternative assessment models and model assumptions were presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.

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

2.

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

3.

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

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