Center for Independent Experts (CIE) Independent Peer Review Report

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

Woods Hole, Massachusetts. December 3-7, 2012

Dr. Noel Cadigan

St. John's, NL, Canada

Executive Summary

The Review Panel could not recommend a single model as the basis for advice for the ***Gulf of Maine cod*** stock. The Stock Assessment Workshop (SAW) Working Group proposed several alternatives that greatly differed in terms of reference points and stock status. Two assessment "packages" were also used for this stock, but these packages gave essentially the same results when they were configured using similar assumptions. The Review Panel was able to narrow the range to essentially two scenarios involving different assumptions about natural mortality. Both assessment scenarios (i.e. models) indicated that the Gulf of Maine cod stock is overfished and overfishing is occurring. This is consistent with the status evaluation based on the previously used assessment model and associated reference points. Spawning stock biomass (SSB) in 2011 is estimated to be 18% or 13% of the SSBmsy proxy, depending on the model. The 2011 fully selected fishing mortality (F) is estimated to be 0.86 or 0.90 which is about 4 or 5 times the Fmsy proxy (0.18 for both models).

A single assessment model was recommended for the ***Georges Bank cod*** stock. However, diagnostics for this model indicated serious model mis-specification and the Review Panel agreed with the SAW Working Group proposal to adjust for this mis-specification (i.e. a retrospective correction). This procedure indicated that the Georges Bank cod stock is overfished and overfishing is occurring. SSB in 2011 is estimated to be 7% of the SSBmsy. The 2011 fully selected F is estimated to be 0.43 which is 2.4 times the Fmsy proxy (0.18).

Background

The 55th Stock Assessment Review Committee (SARC) convened at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA, from December 3-7, 2012. The purpose of the meeting was to provide an external peer review of the stock assessments for Georges Bank and Gulf of Maine cod (*Gadus morhua*). 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 was composed of three independently appointed Center for Independent Experts (CIE) reviewers (Dr. N. Cadigan, Canada; Dr. J. Casey, UK; Dr. S. Holmes, UK), and an independent chair from the Science and Statistical Committee (SSC; Dr. P. Sullivan, USA) of the New England Fishery Management Council. 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 Dr. Robert O'Boyle (New England Fisheries Management Council's Scientific and Statistical Committee). The Georges Bank assessment was presented by Dr. Loretta O'Brien (NEFSC). The Gulf of Maine assessment was presented by Dr. 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.

The CIE reviewers were tasked with conducting an impartial and independent peer review in accordance with the Statement of Work (SoW) and Review Workshop (RW) Terms of Reference (ToRs; Appendix 2). The agenda of the panel review meeting is attached in Annex 3 of Appendix 2. The CIE reviewers were independent, with working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise included statistical catch-at-age, state-space and index methods. Reviewers also had experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers had experience in development of Biological Reference Points that included an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 55 addressed 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 was desirable.

Role of reviewer

About two weeks before the meeting, assessment documents and supporting materials were made available to the Panel via an ftp server. These are listed in Appendix 1. On the morning of the meeting, the Panel met with Drs. Weinberg and Rago to discuss the meeting agenda, reporting requirements, and meeting logistics.

I reviewed the backgrounds documents I was provided. I attended SARC55 held in Woods Hole, MA from December 3-7, 2012. I reviewed presentations and reports and participated in the discussion of these documents, in accordance with the SoW and ToRs (see Appendix 2). This report is structured according to my interpretation of the required format and content described in Annex 1 of Appendix 2. After the meeting, I participated in email discussions dealing with the review panel report and CIE reports.

Summary of findings

The CIE Statement of Work for SAW/SARC55 required that in my CIE report I "should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification". Also, my 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". However, the SARC summary report was not finalized in time for the submission deadline for this report. Hence, I could not report on additional points of view or provide clarifications to the Review Panel consensus opinions on each term of reference (ToR). The best I could do is present my independent views for each ToR, but I could not indicate if these views are divergent with the Review Panel. In what follows some text may simply repeat text I wrote for the summary report.

A. **Gulf of Maine cod stock**

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.

All elements of this ToR were thoroughly addressed.

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

I found 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 ToR. 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 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 *I 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 less than 200 mt per 100 lengths 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 replace 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 catch*.

The revised MRIP recreational catch estimates were approximately 25% lower than the MRFSS estimates pre-2003. Potential sources of this systematic difference should be discussed. A ratio method was used to adjust MRFSS estimates pre-2003 to MRIP equivalents. It does not seem that this uncertainty was accounted for.

I conclude that all elements of this ToR were thoroughly addressed. However, it is clear that the quality of catch information has improved with time. This uncertainty has been adequately characterized.

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.

All aspects of this ToR were addressed.

I was generally satisfied with the way the survey and LPUE data were presented.

However, spatial plots of survey catches by year, similar to Figure A.104 in the GoM cod assessment document but for the two stock areas combined (NEFSC Spring and Fall Surveys), would be helpful to see trans-boundary 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.

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 substantial source of uncertainty. The confidence intervals for ρ (ratio of q's) in Figure A.94 of the assessment WP (#1) were not narrow enough to ignore completely. Even for the relatively well sampled lengths, the range (i.e. $\rho \in$ 1-2.5) could have substantial impact on the assessment. There are a couple of ways I can think of to account for this uncertainty. This may be a useful area for future research, although hopefully the time-series will soon be long enough that direct calibration and adjustment of indices will not be required.

The GLM modeling of survey data was briefly discussed. The modeling was superficial with little motivation. I suggest a GLM approach could be used to combine NEFSC and Massachusetts Department of Marine Fisheries (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. There will be a difference in catchabilities for the two surveys and this could be estimated using a stratum-effects GLM if there are strata in which both surveys sampled in some years.

A GLM model could also be used to address anomalous survey catches – although there are no standard methods for this.

Consideration was given to using LPUE as a measure of relative abundance. There have been changes in the spatial distribution of fishing effort that complicates the interpretation of LPUE in terms of the stock as a whole. This could be accounted for using an appropriate stratified modeling approach. However, a number of regulatory changes were identified (e.g. seasonal closures, trip limits, etc.) which also affect LPUE as an index of stock size. It may also be possible to standardize for some of these regulatory changes but I am not sure the "gain" will be worth the effort.

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 addressed this ToR. This is a work in progress, and I have no additional views.

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 ToR was addressed. The results of considerable investigation were presented to the Review Panel. I generally agreed with the interpretations of results provided by the Working Group (WG). Unfortunately, the information was equivocal.

Various estimates of natural mortality based on life history characteristics were considered. The WG concluded that the evidence available with respect to Gulf of Maine cod life history parameters suggested that an assumption of M=0.2 is reasonable. My sense is that the life history approaches provide only a crude approximation of M, and these approaches may be better at defining the range of possible values. Experience with other Northwest Atlantic cod stocks suggests that M can vary over short time scales (several years), and the life-history approaches do not seem directly useful for detecting such changes in M.

Tagging estimates of M were also considered. The WG identified several concerns with these data. They focused on the implications of assumptions about the high-reward tag return rate on estimates of M. Another problem is short-term tagging mortality. This is a problem in other cod tagging studies when water temperatures are too warm or there is too much thermal stratification (Brattey and Cadigan, 2003). Depth of capture is also an issue. Hence, the WG did not find the high estimates of M based on tagging to be conclusive, and they recommended that this evidence only provided motivation to investigate a change in M in the assessment model.

The WG considered predator field information for cod and concluded it did not provide evidence for a change in M. The WG did not provide any analysis on changes in preferred environment/habitat for cod and potential implications on M. I wondered if the Gulf of Maine is becoming too warm for cod? Changes in temperature have been implicated, perhaps via changes in prey distribution and abundance, in changes in M for northern cod. The WG also provided little information on changes in cod diet and potential implications on M; however, the WG did provide annual estimates of cod condition which did not suggest strong changes over time.

Assumptions about M can have substantial implications for the assessment and management of the stock. Telemetry tagging may provide a more direct way to measure natural mortality, particularly if there are local cod populations with high site fidelity. An acoustic array could be constructed around an over-wintering or feeding aggregation in which some fish are tagged with acoustic tags. The proportion of fish over time that do not return to the “site” can indicate total mortality rates, and if really high rewards are offered for captured tags then it may be possible to partition total mortality into F and M. Those tags retuned by fishermen indicate F, and the rest indicate M. If the fish do not have high site fidelity then the problem is more complex, because fish may simply move to an area without an acoustic array and therefore be undetected but not dead. Pop-up satellite tags are another option for large cod (> 90cm) and this approach is currently being used on the Grand Bank off Newfoundland. However, this would only give information on mortality rates for the large cod that are tagged.

Brattey, J. and Cadigan, N.G. 2003. Estimation of Short-term Tagging Mortality among Adult Atlantic Cod (*Gadus morhua*). Fish. Res. 66: 223-233.

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.

This aspect of the ToR was addressed.

The previous assessment of Gulf of Maine cod was conducted using the statistical catch-at-age model (ASAP) that incorporated commercial and recreational landings and discards. The updated assessment included updates to the recreational catch estimates, revised discard mortality estimates and minor modifications to the Massachusetts Department of Marine Fisheries (MADMF) spring survey. Discard mortality was assumed to vary by gear type and fishery (commercial, recreational). This represents a change from the previous assessment that assumed 100% mortality of all discarded fish. The revised discard estimates range from 20-80% depending on gear type and fishery. The revision to the recreational discard mortality assumption had the largest impact on the assessment with respect to the data changes made since the previous assessment. The recreational catch estimation was changed from the previous assessment from the MRFSS to the MRIP.

The Review Panel could not recommend a single model as the basis for advice for the Gulf of Maine cod stock. The SAW Working Group proposed several alternatives that greatly differed in terms of reference points and stock status. The differences were largely dependent on (1) whether MSY reference points were derived directly using a parametric stock-recruit model applied to stock-recruit estimates derived from assessment models that incorporated historical data prior to 1982, (2) whether proxy methods were used, and (3) whether an M=0.2 or M=0.4 assumption was used to derive reference points. Two assessment "packages" (ASAP and SCAA) were also used for this stock but these packages gave essentially the same results when they were configured using the same the same assumptions.

The Review Panel was able to narrow the range to essentially two scenarios involving different assumptions about natural mortality (M). In one model, M was assumed to be 0.2 for all years. The other model was implemented with M=0.2 from 1982 to 1988, and M=0.4 between 2003 and 2011, with a linear ramp between 1989 and 2002 (denoted Mramp). These values of M are applied to all ages.

Swept-area estimates of abundance from the NEFSC spring and autumn surveys (1982-2011), and the Massachusetts Department of Marine Fisheries (MADMF) spring survey (1982-2011) were used in both ASAP models along with associated estimates of uncertainty and annual age composition. Three fishery selectivity blocks were used instead of the two used in the previous model. Both models assumed flat-topped selectivity for both the catch and NEFSC survey indices whereas the previous model only assumed flat-topped selectivity for the NEFSC surveys. All catch sources were combined into a single fleet.

Feasibility of survey catchability

This aspect of the ToR was addressed.

The MADMF spring survey swept-area q for fully selected ages was approximately 0.25 which makes some sense because this survey only covers part of the stock distribution. The NEFSC spring survey q was close to one whereas in the fall the q is close to 0.6 (Fig. A.177 in WG report). I could not find in the WG report rationale for the differences in q for the spring and fall surveys. It was suggested from the floor during the review workshop that there is seasonal movement of Gulf of Maine cod from the inshore in the spring to the offshore, although the timing of this was not described. Perhaps the fall survey q is lower because cod have moved outside of the survey area?. However, if this explanation is correct then it raises further issues about the stock composition of commercial catches. Also, this suggests that there may be reasons to expect a dome in the fall survey age-pattern in q. This requires further investigation.

Higher catchability of younger ages in the fall survey compared to the spring survey made sense because of increasing recruitment to the survey gear due to growth between the spring and fall.

Starting year for the assessment and estimation of the stock recruitment curve

This aspect of the ToR was addressed.

The start year of the assessment is particularly an issue when deriving MSY reference points directly using a stock-recruit relationship. I appreciated the desire to extend the assessment time-series as long as possible. It seemed that, although earlier data were more uncertain, this was adequately accounted for in the extended assessment models (ASAP and SCAA). A complication with longer time-series is that the various aspects of stock productivity relevant for MSY calculations (i.e. stock-recruit relationship, spawner and yield per recruit relationship) may not be constant. This was a problem for the stock-recruit relationship. There appeared to be evidence of a Ricker-type relationship which has fairly large implications on MSY reference points. However, the low recruitments at high SSB's mostly all occurred during a period in the 1960's. It may be that this was a period of low recruitment productivity for some other reason than high SSB. The stock-recruit data were noisy and, in conjunction with time-trends in residuals, I felt that there were a wide range of plausible stock-recruit relationships that may be changing over time because of changes in the ecosystem. It has been suggested in some of the background material that portions of the population complex of cod are suffering reduced reproductive productivity due to thermally induced changes in zooplankton abundance. Hence, in the end I felt that the WG made the right decision in not directly using a stock-recruit relationship to derive Fmsy. The WG used an %FSPR proxy.

Inclusion of multiple fleets

This aspect of the ToR was addressed.

This was not given much consideration at the review workshop. However, the assessment model used three time-blocks to model fishery selectivity, and these blocks were chosen in part to be consistent with changes in the fishery and regulations. Some other assessment models are using 'smoother' approaches (e.g. random walks) to deal with changes in selectivity and this may be more useful but will require additional research.

Domed or flat selectivity-at–age for the NEFSC surveys

This aspect of the ToR was addressed.

There was no strong external evidence for domed-shaped survey selectivities, and little difference in fit for the assessment model with domed or flat-topped selectivity. Hence the WG recommended the assumption of flat-topped survey selectivity. This seemed reasonable to me, although after the review meeting I was still left with some uncertainty about this because of my concerns about potential movement of older fish outside the fall survey area (see ***survey catchability*** section above).

I conducted a diagnostic analysis using a survey assessment model (SURBA; described below) to get a better understanding of possible causes for the retrospective pattern for this stock. It is noteworthy that the SURBA estimated a domed mortality selectivity pattern when this was freely estimated.

Provide a summary of steps in the model building process

This aspect of the ToR was addressed. This was done very well.

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 aspect of the ToR was addressed.

Historical retrospective analyses were provided, including projections with respect to stock size and fishing mortality. Retrospective performance on catch projections was not provided, and it is not clear to me how this should be done.

The WG also provided many model-retrospective analyses, focusing on effects on SSB, F, recruitment and survey q. There was a tendency to underestimate F and overestimate SSB, although in the very last retrospective run the reverse occurred – SSB was underestimated and F overestimated. The WG could not agree to criteria to adjust for the retrospective pattern observed in the ASAP base model. The WG agreed that there should be no adjustment for the observed retrospective pattern in the base ASAP model because (1) the retrospective pattern is small, (2) it is of an opposite direction to previous patterns, (3) it may be transient, and (4) because of the precedent set in SAW 53 which applied no retrospective adjustment.

Retrospective patterns often indicate that there is a time-trend in residuals, although it is possible to have a retrospective pattern when there is no time-trend. The log-residual time-series plots in Figure A.164 (NEFSC Spring), Figure A.168 (NEFSC Fall), and Figure A.172 (MADMF spring) all had somewhat similar patterns: usually negative early in the time series, followed by a period of positive residuals, but the three survey series all have negative residuals in 2010-2011. The ASAP SAW55_3BLOCK_BASE model tended to over-estimate survey indices early in the assessment time-series and then under-estimate survey indices in the 2000's, except 2010-2011. This is curious, and to get such consistent patterns in residuals there must be some other information in the commercial catches and survey age-compositions that conflicts with the age-aggregated survey indices.

It can be informative to look at the potential conflict between survey indices and catch information using a survey-only stock assessment model. I used the SURBA (e.g. Needle, 2008; Cadigan, 2010) approach (R code donated by C Needle) with a few modifications to predict the age-composition of the catch and trends in total catch for those ages in the assessment model. **Note that this model is only presented for diagnostic purposes**. SURBA is based on the separable mortality model developed by Cook (1997). SURBA is a simple cohort model in which annual age-specific total mortality rates are decomposed into age and year effects. These mortality rates can be cumulated along a cohort and applied to estimates of recruitment to provide age-based estimates of stock size. Most surveys only provide relative measures of stock size, and consequently SURBA can only provide relative estimates of stock size. However, SURBA can provide absolute estimates of total mortality rates. SURBA is probably a better approach to estimate survey Z's than more simple catch-curve methods, although SURBA is only a somewhat sophisticated catch-curve model.

To predict trends in catches based only on surveys, I modified the SURBA model as

$$N_{a,y} = N_{a-1,y-1} \exp(-M - F_{a-1,y-1}). \tag{1}$$

where

$$\log(F_{ay}) = s_a + f_y, \tag{2}$$

and M = 0.2 for all ages and years. Once F's were estimated I used the Baranov catch equations to estimate catch. Surveys are used to estimate model parameters. Let I_{ay} denote the survey index for age a fish in year y. Let t be the midpoint of the survey dates, which I express in a fraction of the year. I used the usual observation equation to estimate parameters,

$$E\{I_{ay}(t)\} = q_a N_{ay} \exp(-tZ_{ay}). \tag{3}$$

For convenience I only used the spring and fall NEFSC indices. I assumed that the indices had a lognormal distribution and I use the maximum likelihood method to estimate parameters. All indices were equally weighted. The q_a 's are confounded with the s_a 's in SURBA and the usual approach around this problem is to fix the q_a 's. I used the values from the ASAP SAW55_3BLOCK_BASE model (Table A.84 in WG report). I assumed the fully selected q for the spring survey was 1 (which defines the scale of the SURBA stock size estimates) and the fully selected q for the fall survey was 0.63.

I used ages 1-9 indices from 1970-2012, but with no plus group. The SURBA code I used did not include a plus group option. ASAP SAW55_3BLOCK_BASE started in 1982 because of issues with the catch data prior to 1982; however, I presumed that there were no issues with the comparability of survey data prior and post 1982. In a preliminary run of this model the estimated s_a 's has a substantially domed pattern. To avoid complications this can cause I decided to also fix the s_a 's at the values produced by the ASAP SAW55_3BLOCK_BASE model (Table A.80) for Block 2.

Parameter estimates are shown in Figure 1, although the selectivity values were inputted and not estimated. Other stock size estimates are shown in Figure 2. Survey indices were in mean number per tow; therefore, the biomass estimates are in kg/tow but adjusted for survey catchability. SURBA residuals (Figures 3-6) look reasonable which indicates that the model has reasonably captured the stock size information in the surveys.

The stock total mortality signal is decomposed into an assumed natural mortality component (M=0.2) and F = Z – M. If M=0.2, and if survey and fishery selectivity have been correctly specified by the values I used from the ASAP SAW55_3BLOCK_BASE model, then the implied trend in catches at ages 1-8 is somewhat different than the WG estimates of catches (Figure 7), with considerable variability since 2000 and prior to 1990. The "SURBA-implied" catch trend is not as steep as the reported trend. There are several explanations for these discrepancies. For example, catches or M prior to 1990 could be too high, or catches or M after 2000 could be too low. Some of the inter-annual variability in "SURBA-implied" catch trends may be related to year effects or other sampling variability in the survey indices. This variability is not accounted for in the SURBA model.

The SURBA-implied age-compositions of catches at ages 1-8 were usually fairly consistent with the assessment estimates. However, periods where fishery selectivity or survey catchability have changed are apparent: 1982-1986/7 may have higher selectivity for ages 1-2; 2005-2011 may have lower selectivity for ages 1-3. These periods correspond well with the selectivity blocks in the SAW55_3BLOCK_BASE model.

My conclusion from this analysis is that the F signal from the NEFSC surveys (assuming M=0.2 and the ASAP SAW55_3BLOCK_BASE model values for survey and fishery selectivity are correct) is substantially different from the reported catches. The surveys are noisy but over all they suggest the catches implied by F=Z-M have not declined as much as reported. The Mramp option for M improves the situation (Figure 9).

Interestingly SURBA has a retrospective pattern in SSB at the beginning of the model (Figure 10). I am not sure why this is. I have not found retrospective patterns with SURBA applications to 2 other cod socks.

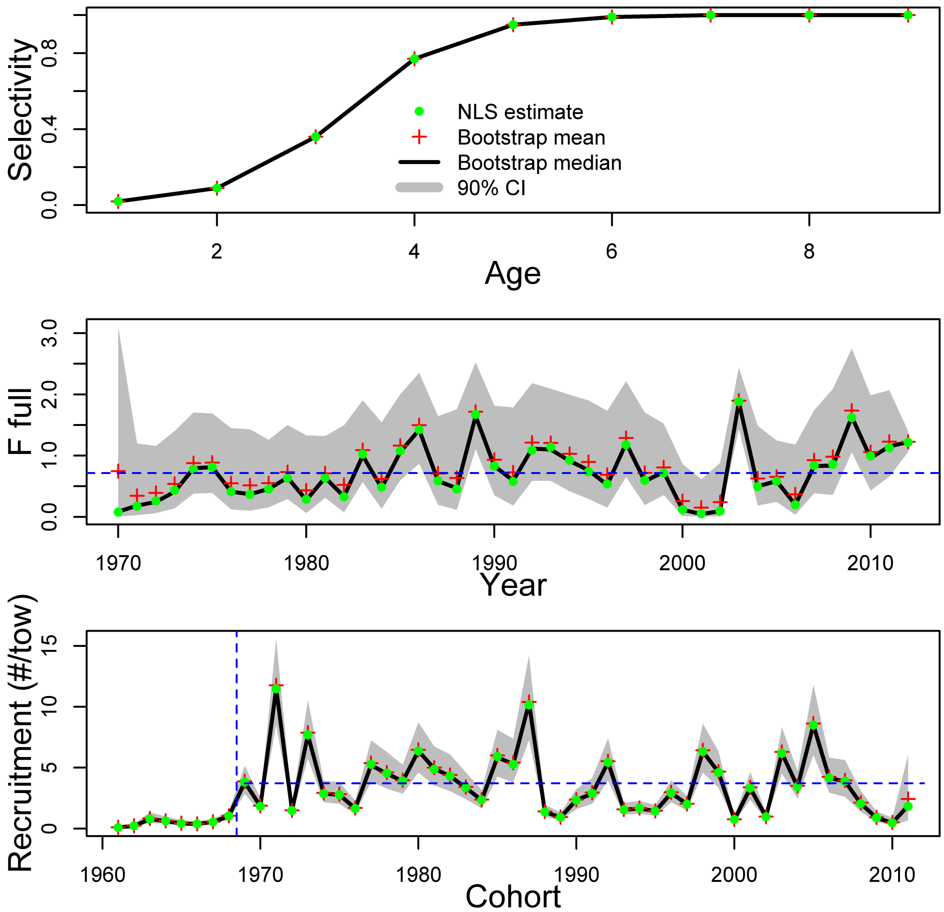

Figure 1. SURBA parameter estimates. Age pattern in fishing mortality (top panel) was fixed and not estimated.

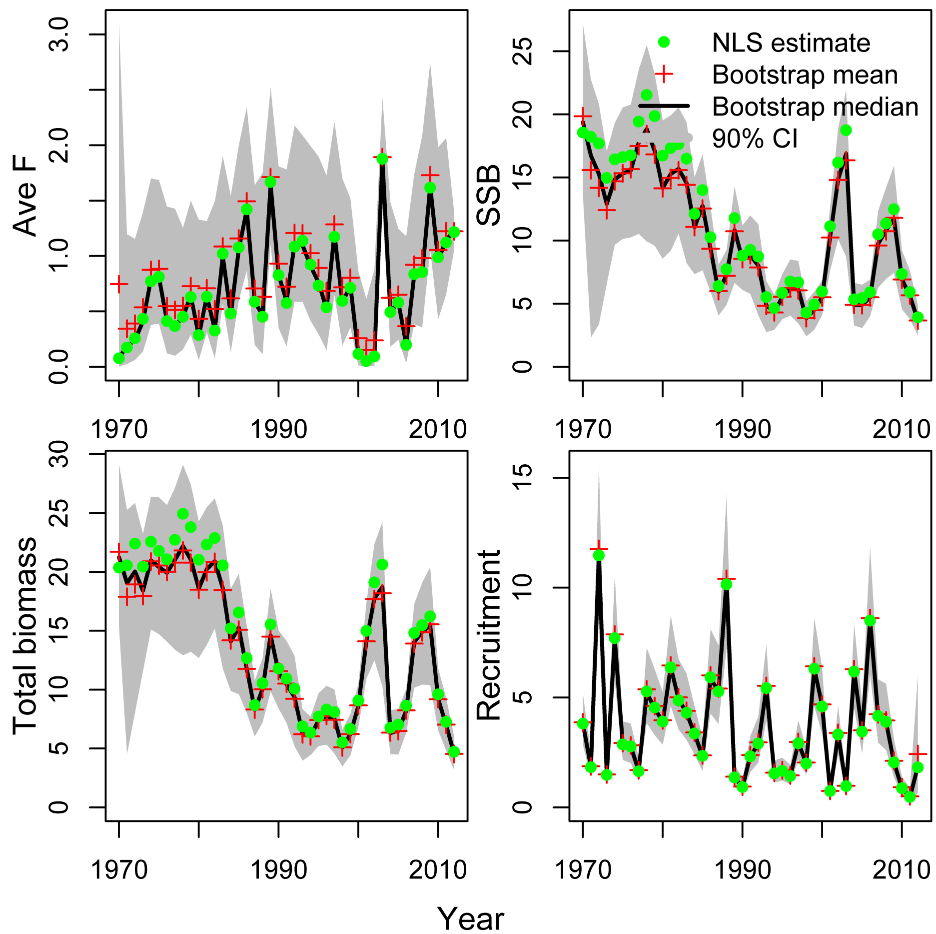

Figure 2. SURBA estimates of stock size and fishing mortality. Average F was for ages 6-9.

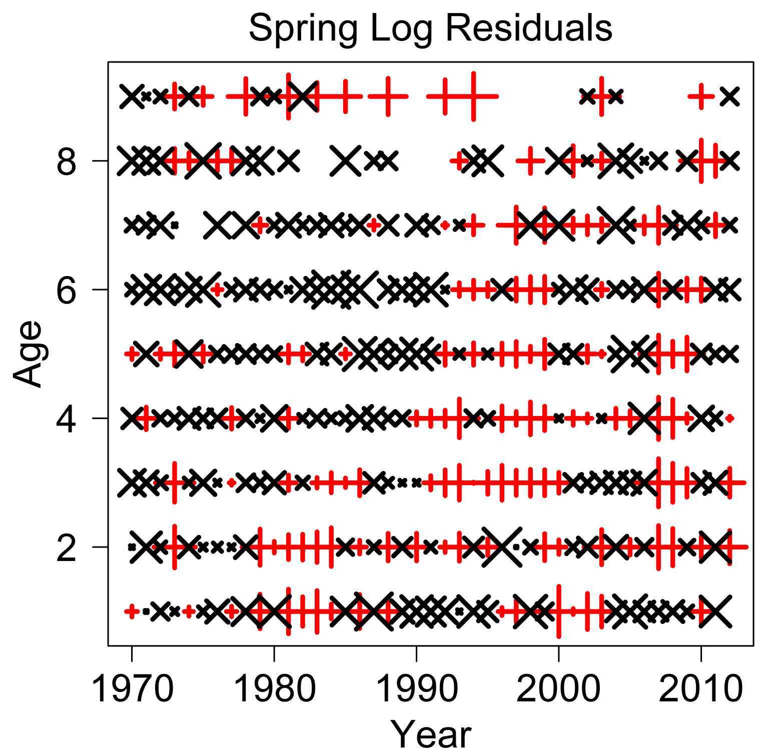

Figure 3. Matrix plot of residuals. Red +'s are positive and black ×'s are negative. The sizes of plotting are proportional to the absolute value of the residuals. Blanks indicate values with zero estimation weights.

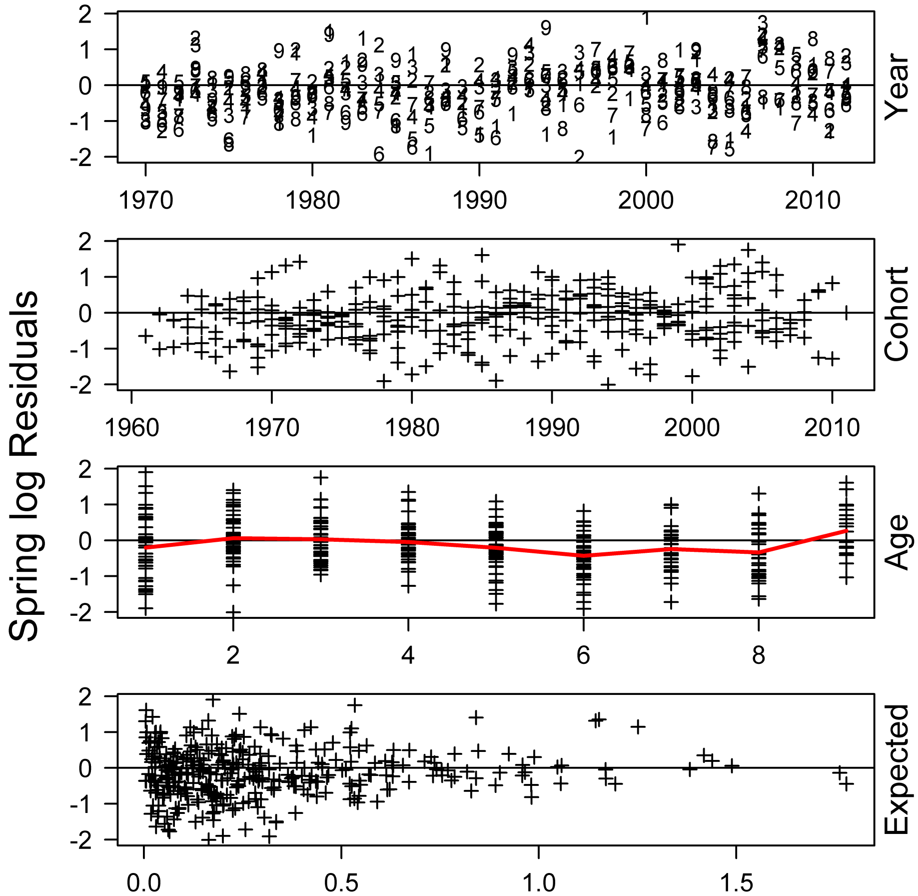

Figure 4. Residuals versus year, age, cohort, and predicted value. The plotting symbols in the top panel indicate age. The red line in the third panel indicates the average residual for each age.

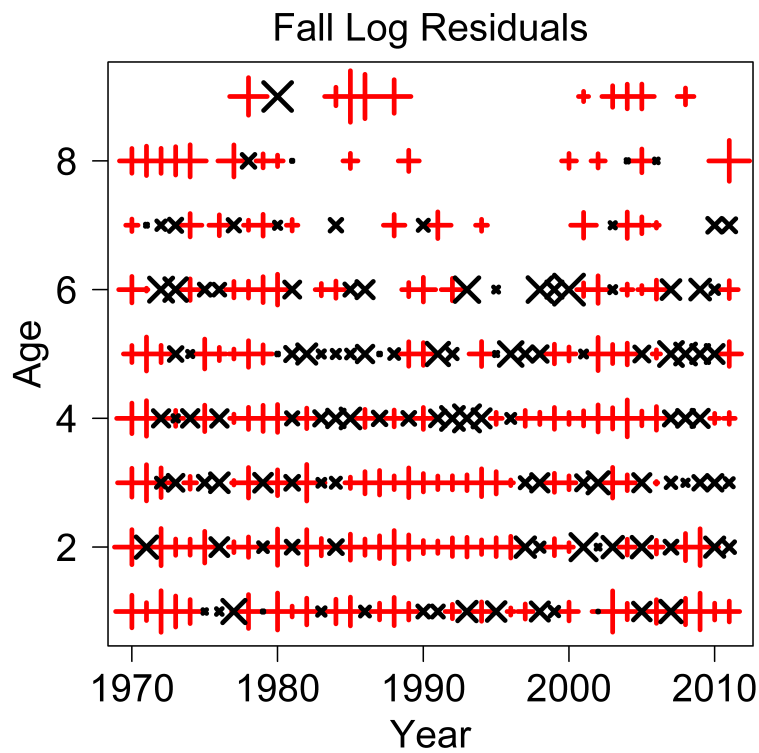

Figure 5. Matrix plot of residuals. Red +'s are positive and black ×'s are negative. The sizes of plotting are proportional to the absolute value of the residuals. Blanks indicate values with zero estimation weights.

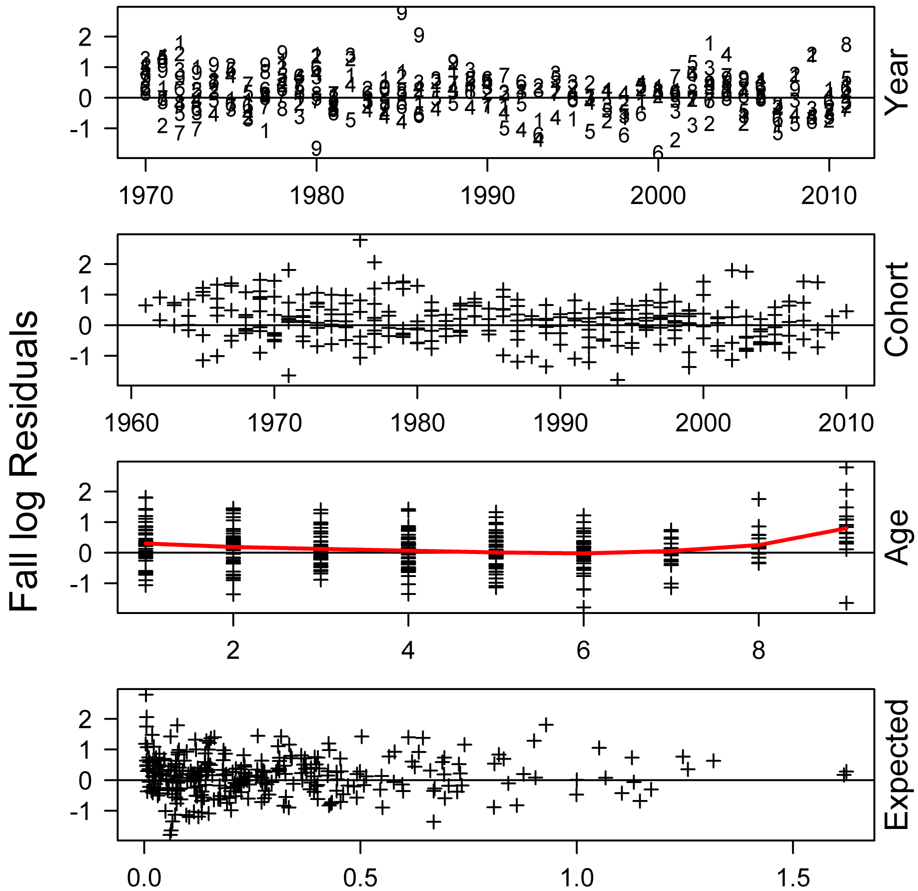

Figure 6. Residuals versus year, age, cohort, and predicted value. The plotting symbols in the top panel indicate age. The red line in the third panel indicates the average residual for each age.

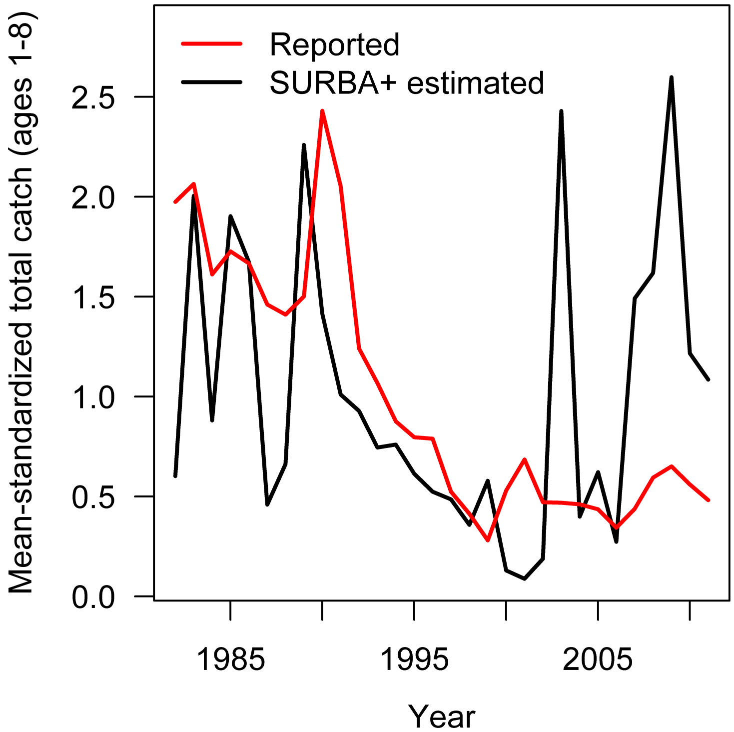

Figure 7. SURBA estimates of trends in catch numbers at ages 1-8 compared to assessment values (i.e. reported). Based on F = Z – M with M=0.2 for all ages and year. Each series is standardized to have a mean of one.

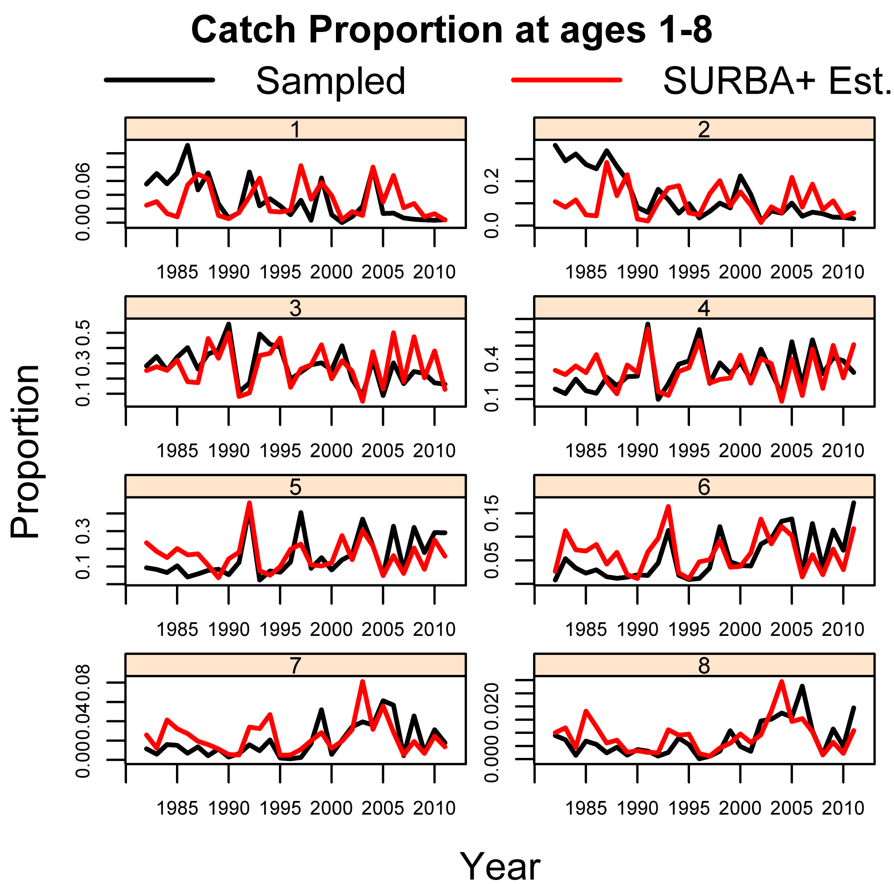

Figure 8. SURBA estimates of the age composition of catches at ages 1-8, versus assessment values (i.e. sampled).

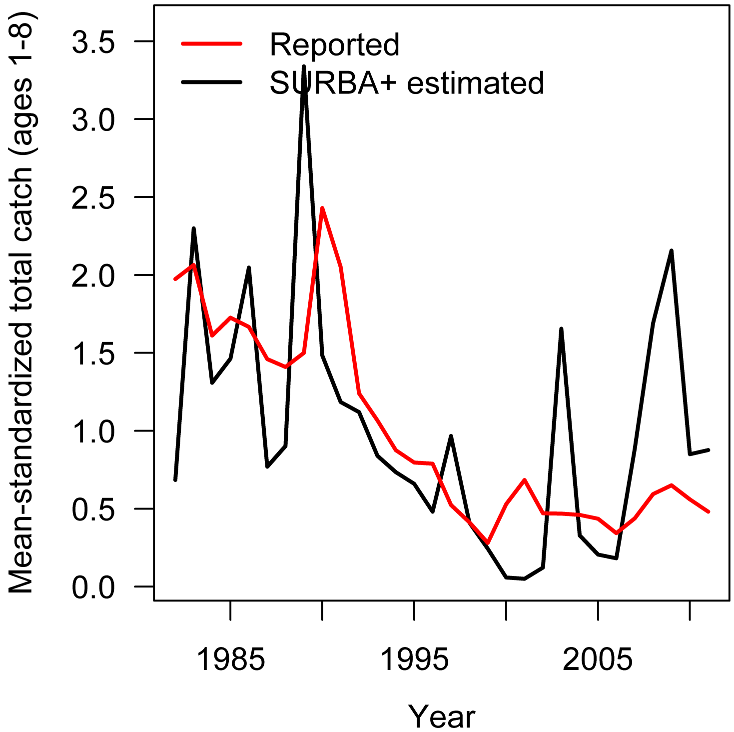

Figure 9. SURBA estimates of trends in catch numbers at ages 1-8 compared to assessment values (i.e. reported). Based on F = Z – M with Mramp option for M. Each series is standardized to have a mean of one

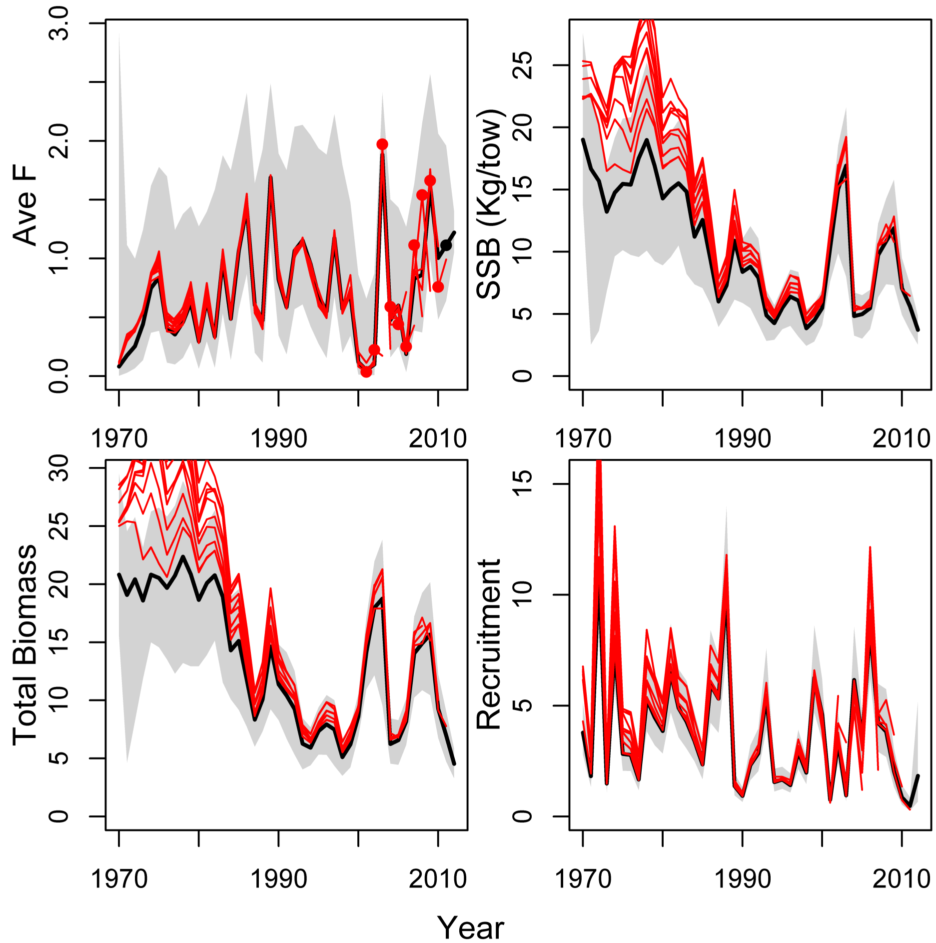

Figure 10. SURBA retrospective (2002-2012) estimates of stock size and fishing mortality. Average F was for ages 6-9.

Cadigan, N. 2010. Trends in Northwest Atlantic Fisheries Organization (NAFO) Subdivision 3Ps Cod (Gadus morhua) stock size based on a separable total mortality model and the Fisheries and Oceans Canada Research Vessel survey index. DFO Can. Sci. Advis. Sec. Res. Doc. 2010/015. iv + 43 p.

Cook, R.M. 1997. Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. ICES J. Mar. Sci. 54: 924–933.

Needle, C.L. 2008. Survey-based fish stock assessment with SURBA. Fisheries Research Services Marine Laboratory. Aberdeen, Scotland.

ToR 6: State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY, and MSY) and provide estimates of their uncertainty. 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.

This ToR was addressed.

Both assessment scenarios (i.e. models) indicated that the Gulf of Maine cod stock is overfished (Bcurrent < 0.5Bmsy) and overfishing is occurring (Fcurrent > Fmsy).

The biological reference points estimated in the last assessment (NEFSC 2012) using a constant M=0.2 were FMSY=F40%=0.20, SSBMSY=61,218 mt, and MSY=10,392 mt.

An MSY could not be derived directly from the two recommended assessments. The Review Panel recommended that proxy methods be used for MSY reference points based on M=0.2

I acknowledge that reference points are sensitive to the value of M used in their calculations. There were no compelling reasons to deviate from the usual assumption of M=0.2 for long-term projections to derive MSY reference points.

F40% is the proxy used for the overfishing threshold (Fmsy). This is consistent with the choice of proxy in the previous assessment and the SAW 55 working group's recommendation. A deterministic value of F40% was calculated from a spawner per recruit analysis using 2007-2011 average SSB weights, catch weights, selectivity and maturity. Expressed as a fully selected fishing mortality, F40% is 0.18.

Stochastic projections at F40% were used to determine new recommended biomass-related reference points (proxies for both SSBMSY and MSY). The projection method involved recruitment sampled from a cumulative density function derived from ASAP estimated age-1 recruitment between 1982 and 2009. No retrospective adjustment was applied in the projections. Projected recruitment was adjusted when SSB fell below the lowest observed SSB estimate (6.3 kmt or 7.9 kmt) based on a linear function that declined to zero when SSB=0. This depended on the assessment model formulation. This projection method was recommended in SARC53.

The SSBMSY proxies were 54,473 mt or 80,200 mt in the M=0.2 or Mramp models, respectively.

For reasons I indicated under ToR5, I conclude that the WG made the right decision in not directly using a parametric stock-recruit relationship to derive Fmsy. However, the choice of the proxy seemed subjective and it is difficult for me to evaluate the appropriateness of this choice.

One could use the stochastic projection method to directly derive an Fmsy, but I suspect it will be similar to Fmax. However, this should be given additional consideration in the future.

There were three problems in deriving MSY reference points: (1) the stock-recruit data did not clearly indicate the level of SSB in which density-dependent processes reduced recruitment rates, (2) there were also some temporal dependencies in recruitment rates indicating other factors are at play in addition to stock size, and (3) there was uncertainty in the value of M to use when computing reference points. These are common problems. For the first problem, I think another reasonable approach is to simply constrain Rmax at the historic maximum observed value. A stock-recruit model that indicates otherwise will usually be highly speculative. This is a data-based solution, whereas the choice of Fx% proxy is not completely data-based (i.e. the x% is somewhat subjective). However, one may still find substantial sensitivity to the assumed form of the Rmax-constrained stock-recruit model. There are nonparametric alternatives that could be considered (i.e. Cadigan, 2013) although the message there is that there is substantial uncertainty about values for MSY reference points due to uncertainty in the stock-recruit relationship. Problem (2) can be addressed through stochastic simulations including temporal dependencies in recruitment rates. Problem (3) can be addressed in the same way, but our knowledge about M is more limited and I think the appropriate values of M relevant for MSY reference point calculations will remain speculative for some time.

Related to problems 2) and 3), provide retrospective analysis of proposed reference points. For example, retrospective F proxy's should be provided, based on retrospective average weights, maturities, and fishery selectivity. Stochastic projection Bmsy's could be provided based on retrospective F proxy's, retrospective stock-recruitment relationships, and retrospective averages for biological parameters (or whatever procedure is used). If there is substantial retrospective variation then this will need to be further investigated. My understanding is that MSY reference points should be based on conditions that are thought to prevail in the future. It may not be reasonable to assume recent short-term (i.e. 5 year) averages will prevail for the time frame relevant for MSY calculations (i.e. many years). However, if there is little retrospective variation in reference points then this may not be an important issue. This is my purpose for this recommendation.

Cadigan, N.G. 2013. Fitting a nonparametric stock-recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty. ICES J. Mar. Sci. 70: 56–67

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).

This was addressed.

Both SARC55 recommended assessment scenarios (i.e. models: M=0.2 and Mramp) indicated that the Gulf of Maine cod stock is overfished and overfishing is occurring. This is consistent with the status evaluation based on the previously used SARC53 assessment model and associated reference points. Spawning stock biomass (SSB) in 2011 is estimated to be 18% or 13% of the SSBmsy proxy, depending on the model. The 2011 fully selected fishing mortality (F) is estimated to be 0.86 or 0.90 which is about 4 or 5 times the Fmsy proxy (0.18 for both models).

It is odd that this ToR does not ask for uncertainty to be accounted for. We can never be absolutely certain of stock status. While there is uncertainty in the best stock assessment model formulation and in the best values for reference points, I am confident in the statement: There is a high (i.e. > 0.5) probability that the stock is overfished and overfishing is occurring. This is my independent point of view and I recognize that there are different approaches to measure the probability of stock status.

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.***

This ToR was addressed.

For ToR8a, 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.

The review panel concluded for ToR8b that the M=0.2 projections and the Mramp projections with M remaining at 0.4 in the short-term were equally plausible. Little evidence was presented to the panel to favor one scenario over the other. The WG 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 that in the short-term M would remain at 0.4 rather than reduce to 0.2. Note that for long-term projections that panel decided that M should be 0.2, because the longer-term historical evidence seems to indicate that M=0.2 is more plausible.

The review panel appreciated the description of the stock's vulnerability to becoming overfished (ToR8c). 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. All of this points to a stock at a low level and with a concentrated distribution that is vulnerable to overfishing.

Potential future variability in maturities, catch and stock weights should be considered in medium to long-term projections.

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.

This ToR was addressed.

The review panel appreciated both the amount of progress and the reporting of progress on previous research recommendations. A single recommendation was carried forward from GARM III, which was addressed in the WG report. Of the nine research recommendations brought forward from SARC 53, six were either partially or fully addressed.

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 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.

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

Additional research recommendations:

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 ‘smoother’ approaches (e.g. random walks like in ICES-SSAM) to deal with changes in fishery selectivity and natural mortality.
4. 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, as illustrated in this report.
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.
8. Provide retrospective analysis of proposed reference points

B. Georges Bank cod stock

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.

This ToR was addressed.

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

I found that the documentation of results and procedures to estimate catch and their uncertainty were very helpful. I found no reason to disagree with the conclusions of the Working Group, and these are highlighted below, and include some justification.

Prior to 1994, information on 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 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 length 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 TOR 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 Panel concludes that this term of reference was addressed adequately for the purpose of assessment.

Figures like B2a in the WG report are useful. In addition, it would be useful when assessing how important are the recreational and discard estimates to the assessment if figures like B2a could be provided in numbers and by age. Mortality is based on numbers that die and not their weight.

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 ToR was addressed.

I was generally satisfied with the way the survey and LPUE data were presented.

However, similar to the recommendation for Gulf of Maine cod, spatial plots of survey catches by year would be helpful to see trans-boundary 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.

Similar to the Gulf of Maine cod assessment, 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 substantial source of uncertainty. The confidence intervals for ρ (ratio of q's) in Table B13 of the assessment WP were not narrow enough to ignore completely. For example, for lengths > 55cm the range $\rho \in$ 1-2.5 could have substantial impact on the assessment. There are a couple of ways I can think of to account for this uncertainty. 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.

The GLM modeling of survey data was briefly discussed. The modeling was superficial with little motivation.

Consideration was given to using LPUE as a measure of relative abundance. There have been changes in the spatial distribution of fishing effort that complicates the interpretation of LPUE in terms of the stock as a whole. This could be accounted for using an appropriate stratified modeling approach. However, other regulatory changes were identified which also affect LPUE as an index of stock size. Canadian fishery contributes about an average 25% to the overall landings and they are not accounted for in the LPUE. It may be possible to standardize for some of these regulatory changes and include Canadian LPUE data but I am not sure the "gain" will be worth the effort.

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

This ToR was addressed.

The summarization of the workshop findings was thorough and met the Terms of Reference. This is a work in progress, and I have no additional views.

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 ToR was addressed.

The results of considerable investigation were presented to the Review Panel. I generally agreed with the interpretations of results provided by the Working Group. Unfortunately the information was equivocal.

Various estimates of natural mortality based on life history characteristics were considered. The WG concluded that the evidence available with respect to Gulf of Maine cod life history parameters suggested that an assumption of M=0.2 is reasonable. My sense is that the life history approaches provide only a crude approximation of M, and these approaches may be better at defining the range of possible values. Experience with other Northwest Atlantic cod stocks suggests that M can vary over short time scales (several years), and the life-history approaches do not seem directly useful for detecting such changes in M.

Similar to Gulf of Maine cod, tagging estimates of M were also considered. The WG identified several concerns with these data. They focused on the implications of assumptions about the high-reward tag return rate on estimates of M. Another problem is short-term tagging mortality. This is a problem in other cod tagging studies when water temperatures are too warm or there is too much thermal stratification (Brattey and Cadigan, 2003). Depth of capture is also an issue. Hence, the WG did not finding the high estimates of M based on tagging to be conclusive, and recommended that this evidence provided motivation to investigate a change in M in the assessment model.

The WG considered predator field information for cod and concluded it did not provide evidence for a change in M. The WG did not provide any analysis on changes in preferred environment/habitat for cod and potential implications on M. I wondered if Georges Bank is becoming too warm for cod? Changes in temperature have been implicated, perhaps via changes in prey distribution and abundance, in changes in M for northern cod. The WG also provided little information on changes in cod diet and potential implications on M; however, the WG did provide annual estimates of cod condition – and there has been a decrease in condition in the spring but not the autumn. This was considered to be a conflicting result. I am not sure what is going on here, but one interpretation is that the cod in poor condition in the spring do not survive till the fall.

Brattey, J. and Cadigan, N.G. 2003. Estimation of Short-term Tagging Mortality among Adult Atlantic Cod (Gadus morhua). Fish. Res. 66: 223-233.

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.

This aspect of the ToR was addressed.

The previous assessment of Georges Bank cod was conducted using virtual population analysis (VPA). This assessment used ASAP which can more fully account for the uncertainties in the catch and surveys. The ASAP model incorporates the total catch of USA and CDN commercial and recreational landings and discards. The model uses two fishery selectivity blocks assuming flat-topped selectivity. Discard mortality was assumed to vary by gear type, but only for USA fisheries. This represents a change from previous assessments that assumed 100% mortality of discarded fish.

Swept-area estimates of abundance from the NEFSC spring and autumn surveys (1978-2011), and the Department of Fisheries and Oceans (DFO) survey (1986-2011) were used in the ASAP model along with associated estimates of uncertainty and annual age composition. The model assumed flat-topped selectivity for survey indices.

The Working Group proposed three model scenarios for the Review Panel to consider, and other sensitivity runs were described. The BASE model assumed M=0.2 for all ages and years. In the Mramp model M=0.2 prior to 1990 but increased from 0.2 to 0.4 between 1990-2003 and remained at 0.4 since then. A catch multiplier (Cmult) scenario incorporated a three-fold increase in catch during the 1995-2011.

The Cmult scenario resulted in stock size estimates that were quite different than the BASE M=0.2 and Mramp scenarios. The WG did not present residual diagnostics for the catch multiplier model; however, I suspect there may be evidence of model-misspecification because the stock trends seem so different. The WG did note (pg 46 WG report) that this model was 29 log-likelihood points higher than the base formulation. This seems substantial. I did not considered the catch multiplier scenario to be as feasible as M=0.2 and Mramp scenarios.

The Mramp model did not fit the data better; it was 2 log-likelihood points higher than the base formulation. It did provide a better fit to age-aggregated survey indices but provided a poorer fit to the survey age-compositions. While the Mramp model reduced the SSB rho to 0.053 (base SSB rho = 0.681) and the F rho to 0.088 (base F rho = 0.458), it introduced a retrospective pattern during 1994-2002. I do not consider that the specific Mramp scenario proposed by the WG was an appropriate correction for the model-misspecification evident in the M=0.2 base model.

Of the three ASAP model formulations provided, I concluded that the BASE M=0.2 model with retrospective correction was the best option for short-term catch advice.

Feasibility of survey catchability

This aspect of the ToR was partially addressed.

Swept-area q's for fully selected ages were not provided in the WG report but I did find them in results presented at the meeting (e.g. BASE.RESULTS.PLOTS.pdf). The fall survey q was a little less than 50% of the spring q, for the both the ASAP base and Mramp models. This is the same pattern as for Gulf of Maine cod, and rationale for this should be provided.

Higher catchability of younger ages in the fall survey compared to the spring made sense because of increasing recruitment to the survey gear due to growth between the spring and fall. However, the differences were much larger than in the Gulf of Maine ASAP models. Age 3 fish are fully recruited to the Georges Bank fall survey while for the Gulf of Maine, fish are not fully recruited until age 6. Mechanisms for this should be considered. I conclude that some of the patterns in survey catchability estimates are not intuitive.

Starting year for the assessment and estimation of the stock recruitment curve

This aspect of the ToR was addressed.

The start year of the assessment was 1978. No other alternatives were explored.

The WG concluded that the relationship between stock and recruitment during 1978-2011 did not provide support for use of either a Ricker or Beverton-Holt (BH) function. The stock-recruitment data were not shown in the WG report but I found them in files provided at the meeting (RAMPM.RESULTS.PLOTS.pdf; BASE.RESULTS.PLOTS.pdf). I agree that with both of these models there is no evidence of a Ricker stock-recruitment relationship. However, for the Mramp model I suggest there is evidence of a BH function. In many assessments it would be considered appropriate to fit a BH function to these data and derive MSY reference points this way. However, the Mramp model was not ultimately used for advice. There are ways to fit stock-recruit curves to these data and I illustrate this under ToR6 below.

Inclusion of multiple fleets

This aspect of the ToR was addressed.

This was not given much consideration at the review workshop. However, the assessment model explored various time-blocks to model fishery selectivity, and these blocks were chosen in part to be consistent with changes in the fishery and regulations. Other assessments are using 'smoother' approaches (e.g. random walks) to deal with changes in selectivity and this may be more useful but will require additional research.

Domed or flat selectivity-at–age for the NEFSC surveys

This aspect of the ToR was addressed.

The review panel was not provided with convincing evidence for domed-shaped survey selectivity. There was little difference in fit between assessment models with domed or flat-topped selectivity. When selectivity was freely estimated the curves were considered by the WG to be "essentially flat-topped". Hence the WG recommended the assumption of flat-topped survey selectivity. This seemed reasonable to me.

Provide a summary of steps in the model building process

This aspect of the ToR was addressed very well.

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 aspect of the ToR was partially addressed.

Historical retrospective analyses were not provided. It is not clear to me how retrospective performance on catch projections should be done.

The WG provided many within-model retrospective analyses, focusing on effects on SSB, F, recruitment and survey q. The BASE ASAP (M=0.2) model had a strong tendency to underestimate F and overestimate SSB. The WG agreed to address the retrospective bias in the BASE ASAP by adjusting the terminal year results by applying the 7-year rho factor. I felt this was a reasonable approach for short-term projections but probably does not adequately account for uncertainty due to model-misspecification in longer-term projections.

Retrospective patterns often indicate that there is a time-trend in residuals, although it is possible to have a retrospective pattern when there is no obvious time-trend in residuals. The log-residual time-series plots in Figure B39 (DFO), Figure B42 (NEFSC Fall), and Figure B48 (NEFSC spring, Yankee #36) all had somewhat similar patterns: usually negative early in the time series, followed by a period of positive residuals, but tending to have negative residuals in 2010-2011. The ASAP BASE model tended to over-estimate survey indices early in the assessment time-series and then under-estimate survey indices in the 2000's, except 2010-2011. This is curious, and to get such consistent patterns in residuals there must be some other information in the commercial catches and survey age-compositions that conflicts with the age-aggregated survey indices. It is even more curious that the Gulf of Maine stock has the same basic residual pattern. This suggests that migration is not the cause of these residual patterns.

ToR 6: State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY , and MSY) and provide estimates of their uncertainty. 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.

Stock status

This ToR was addressed.

The Georges Bank cod stock is overfished (Bcurrent < 0.5Bmsy) and overfishing is occurring (Fcurrent > Fmsy).

The biological reference points estimated in the last assessment (Groundfish Update, 2012) were F_{MSY}=$F_{40\%}$=0.23, SSB_{MSY}=140,424 mt, and MSY=28,774 mt.

The WG concluded that the "relationship between stock and recruitment does not support the use of a parametric model" and the same proxy for Fmsy used in the last assessment (i.e. F40%) was proposed for this assessment. I agree with this choice. A deterministic value of $F_{40\%}$ was estimated from a spawner per recruit analysis using 2007-2011 average SSB weights, catch weights, maturity and selectivity at age. Expressed as a fully recruited fishing mortality (ages 5+), $F_{40\%}$ is 0.18.

I acknowledge that reference points are sensitive to the value of M used in their calculations. There were no compelling reasons to deviate from the usual assumption of M=0.2 for long-term projections to derive MSY reference points.

Stochastic projections at $F_{40\%}$ were used to determine biomass reference points (proxies for both SSB_{MSY} and MSY). The projection method involved recruitment sampled from a 2-stage cumulative density function (CDF) of 1978-2011 ASAP estimated age-1 abundance associated with a SSB breakpoint of 41,500 mt. Recruitment is sampled from the low recruitment CDF when SSB is < 41,500 mt or from the high recruitment CDF when SSB > 41,500 mt. Age specific retrospective pattern adjustments to the abundance at age were be used to start the projections. This projection method was recommended in GARMIII and reviewed by the SAW 55 WG.

The proxy for SSB_{MSY} is estimated at 186,535 mt, the median of the stochastic projections. The proxy for MSY is 30,622 mt.

The choice of the F proxy seemed subjective and it is difficult for me to evaluate the appropriateness of this choice. I tried to fit a BH model to the stock-recruit data in Table B23 in an effort to directly derive Fmsy. The estimation did not converge because Rmax is not defined for these data. Hence, I conclude that the WG made the right decision in not directly using a parametric stock-recruit relationship to derive Fmsy.

Another reasonable approach is to simply constrain Rmax to be less than the historic maximum observed value in 1986. This is a data-based solution, whereas the choice of Fx% proxy is not completely data-based (i.e. the x% is somewhat subjective). I tried the constrained-approach **for illustration purposes only** and, using the biological data in Table B24, the resulting estimate of Fmsy is 0.14 with a 95% confidence interval (0.13, 0.16) and Bmsy is 340.2 Kt, (139.5, 344.7). Confidence intervals are based on bootstrapping the stock-recruit fits. The BH Fmsy is less than F40% = 0.18 and the BH Bmsy is greater than the stochastic projection Bmsy proxy of 186.5 Kt. These differences will not change conclusions from this assessment.

The constrained BH model did not appear to capture well the stock-recruit pattern when SSB was greater than 50 Kt (Figures 10 and 11). Most residuals were positive in this case. I fit the nonparametric model (NP) described in Cadigan (2013). It resulted in a slightly improved residual pattern. Note that although the fits are similar within the range of estimates SSB's (Figure 10) there are differences when extrapolating recruitment at large stock sizes (Figure 12). Unfortunately, such extrapolations are often required when directly deriving MSY reference points. The corresponding equilibrium yield curves are different (Figure 13) although the Fmsy's are less different, with NP Fmsy = 0.16 (0.13, 0.23). The NP Bmsy is 350.2 which is similar to the BH Bmsy of 340.2 Kt but greater than the WG Bmsy proxy of 186.5 Kt. The NP Bmsy has a wide confidence interval (90.9, 961.8). Bmsy is poorly defined. The NP Fmsy is close to F40% = 0.18, and well within the NP confidence interval.

An aside: The residual pattern in Figure 11 is similar to the basic pattern for 8 of 9 case studies in Cadigan (2013). I am not sure if this is important but it is baffling to me why these different stocks have similar residual patterns.

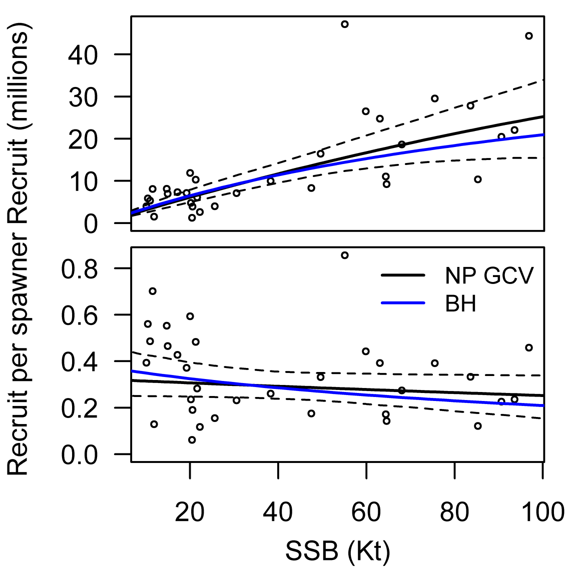

Figure 10. Fits to the BASE M=0.2 stock-recruit data based on a Beverton-Holt (BH) model and a nonparametric model (NP GCV). The bottom panel shows corresponding recruits per spawner. Dashed lines are 95% bootstrapped confidence intervals from the NP GCV model.

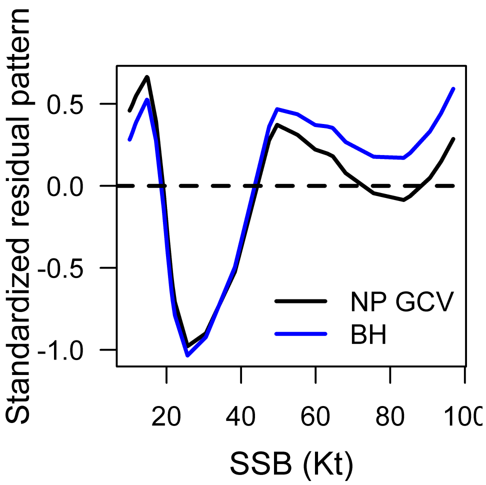

Figure 11. Loess smooths of the residuals from the Beverton-Holt (BH) and nonparametric (NP GCV) stock-recruit models.

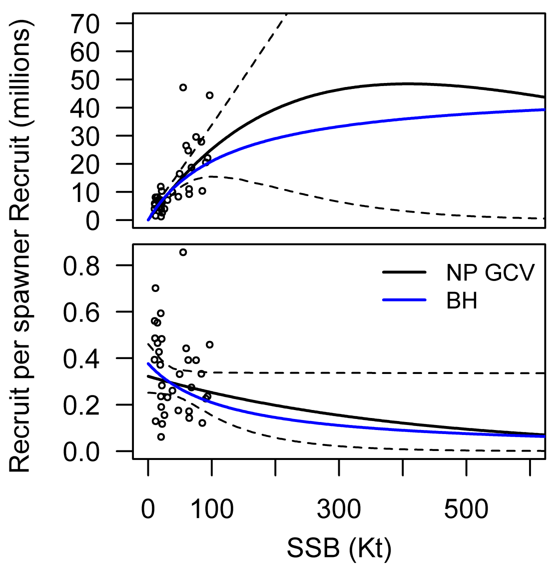

Figure 12. Stock-recruit predictions based on a Beverton-Holt (BH) model and a nonparametric model (NP GCV). The bottom panel shows corresponding recruits per spawner. Dashed lines are 95% bootstrapped confidence intervals from the NP GCV model.

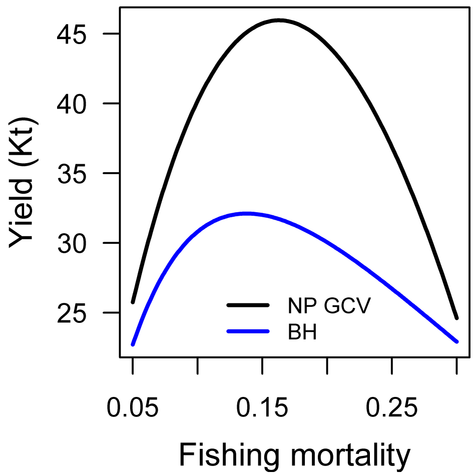

Figure 13. Equilibrium yield curves based on Beverton-Holt (BH) and nonparametric (NP GCV) stock-recruit models.

The stochastic projection proxy using F40% for MSY is 30,622 mt, with 10th and 90th percentiles spanning 25,450 – 36,302 mt. Historically, since the 1890's (Figure B2a in WG report), when catches have exceeded this MSY level they have been subsequently followed by periods of declining catches. Catches have increased for several years only following periods where catches were less than MSY. This historic data corroborated the MSY value for me.

This also suggests that a simple production model fitted with this longer catch time series and age-aggregated survey indices may provide some useful and corroborating information on the appropriateness of Fmsy and Bmsy for this stock. I suspect a production model with process errors will be required.

Cadigan, N.G. 2013. Fitting a nonparametric stock-recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty. ICES J. Mar. Sci. 70: 56–67.

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).

This ToR was addressed.

The Georges Bank cod stock is overfished and overfishing is occurring. SSB in 2011 is estimated to be 7% of the SSBmsy. The 2011 fully selected F is estimated to be 0.43 which is 2.4 times the Fmsy proxy (0.18). This is consistent with the status evaluation based on the previously used (Groundfish Update, 2012) assessment model and associated reference points.

It is odd that this ToR does not ask for uncertainty to be accounted for. We can never be absolutely certain of stock status. Nonetheless, I am confident in the statement: There is a high (i.e. > 0.5) probability that the stock is overfished and overfishing is occurring. This is my independent point of view and I recognize that there are different approaches to measure the probability of stock status.

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.***

This ToR was addressed.

For ToR8a, short-term projections were provided using the same stochastic projection method used for the reference point calculations. 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.

The review panel appreciated the description of the stock's vulnerability to becoming overfished (ToR8c). There has been poor recruitment in the last two decades and M may have increased recently.

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.

This ToR was addressed.

The review panel appreciated both the amount of progress and the reporting of progress on previous research recommendations.

The GARM III Panel recommended that 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 WG response was that was not productive due to issues in the catch information. I sort of agree for recruitment estimation; however, the historical catch information may provide some information on reference points (See recommendation c) below). It is possible to construct a model that can use just the historical catch data prior to age sampling (i.e. like SCAA), and this data may provide information on carrying capacity. For example, if you know historical catches were much larger than those in the assessment time-frame then this says something about carrying capacity.

Most of the Review Panel Gulf of Maine (GoM) research recommendations (RR) also apply to the Georges Bank stock. However, telemetry tagging in offshore areas (GoM RR #2) may be more problematic. GoM RR #5 does not apply.

In addition:

a) This stock had a retrospective pattern which was corrected for in short-term projections. Future assessments should 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 the starting population size for projections does not seem sufficient particularly in long-term projections for rebuilding analyses and reference point calculations.
b) Figures like B2a in the WG report are useful. In addition, it would be useful when assessing how important are the recreational and discard estimates to the assessment if figures like B2a could be provided in numbers and by age. Mortality is based on numbers that die and not their weight.
c) A simple production model fitted with this longer catch time series and age-aggregated survey indices may provide some useful and corroborating information on the appropriateness of Fmsy and Bmsy for this stock. I suspect a production model with process errors will be required.

Critique of the NMFS review process

There was insufficient time to provide a thorough review for these stocks. There were two main factors that contributed to this:

1. The documentation provided was extensive but not well organized. For GoM cod, the report was long and I was often frustrated with the time I had to spend finding particular tables and figures. Consider providing information in separate appendices or working papers, dealing with biological inputs, catch, surveys, and model results and perhaps other appropriate divisions. For Georges Bank, important information on model fits and stock-recruit relationships was not provided in the main document and I had to spend too much time searching through other files provided during the meeting.

2. There were several interventions from the audience during the review workshop that dealt with the review process. This consumed approximately 3-4 hours of valuable meeting time in total and also contributed to the review panel not being able to finalize the panel's summary bullets by the end of the meeting. It was difficult and time consuming to work on the summary report after the meeting. All panel members and other review group participants should have a clear understanding of the review process and the roles and responsibilities of all participants before the meeting.

A succinct assessment summary that describes the WG perceptions of critical issues should be provided before the review workshop. In SARC55 the SCAA presentation was a good example of this.

Appendix 1: Bibliography of materials provided for review

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 (*Gadus 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.

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 Hole27 – 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 2: CIE 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:

<u>SARC CIE reviewers</u>:

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 (NOTE: Order of Topics in the final agenda is likely to change)

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 – 3:30 Assessment Presentation (A. GOM cod)

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

3:30 – 3:45 Break

3:45 – 6 Assessment Presentation (A. GOM cod)

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

Tuesday, Dec. 4

9 – 10:45	SARC Discussion w/ presenters (A. GOM cod)	
	Patrick Sullivan, SARC Chair	**TBD**
10:45 – 11	Break	
11 – 12:15	Assessment Presentation (B. GBK COD)	
	Loretta O'Brien **TBD**	**TBD**
12:15 – 1:30	Lunch	
1:30– 3:45	(cont.) Assessment Presentation (B. GBK COD)	
	Loretta O'Brien **TBD**	**TBD**
3:45 – 4	Break	
4 – 5:45	SARC Discussion w/ presenters (B. GBK COD)	
	Patrick Sullivan, SARC Chair	**TBD**
7	social event **--locationTBD**	

Wednesday, Dec. 5

9 - 11	Revisit w/ presenters (A. GOM cod)	
	Patrick Sullivan, SARC Chair	**TBD**
11 – 11:15	Break	
11:15 – 12:30	Revisit w/ presenters (B. GBK COD)	
	Patrick Sullivan, SARC Chair	**TBD**
12:30 – 1:45	Lunch	
1:45 – 2:15	(cont.) Revisit w/ presenters (B. GBK COD)	
	Patrick Sullivan, SARC Chair	**TBD**
2:15 -2:30	Break	

2:30 – 5:30 Review/edit Assessment Summary Report (A. GOM cod)

Patrick Sullivan, SARC Chair **TBD**

<u>Thursday, Dec. 6</u>

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

Patrick Sullivan, SARC Chair **TBD**

12 – 1:15 **Lunch**

1:15 – 5 SARC Report writing. (closed meeting)

<u>Friday, Dec. 7</u>

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.