SAW/SARC 55 Working Group
Report of the Models Issues Meeting
Aquarium Building, S.H. Clark Conference Room, Woods Hole
15 – 19 October 2012

Introduction

Benchmark assessments of the Gulf of Maine (GOM) and Georges Bank (GB) cod stocks are to be reviewed at the SAW/SARC 55 meeting of 3 – 7 December 2012. The Terms of Reference for this review are at http://www.nefsc.noaa.gov/saw/saw55/SAW-SARC-55-FINAL_TORs-2012-05-30-1.pdf. The last benchmark review of the Gulf of Maine cod assessment was conducted at the SARC 53 meeting of 29 Nov – 2 Dec 2011 while that of Georges Bank Cod was conducted at the GARM III meeting in August 2008. In preparation for the December review, the SAW/SARC 55 Working Group (WG) of the NMFS Northeast Fisheries Science Centre (NEFSC), Woods Hole, conducted a review of data issues during 27 – 31 August 2012, the report of which is at http://www.nefsc.noaa.gov/saw/. The WG is to discuss and formulate the assessment models and reference points for the GOM and GB cod during two meetings:

- Modeling issues: 15 – 19 October 2012
- Modeling and Reference point issues: 29 October – 2 November 2012

This document is the report of the Models issues meeting.

The task of the WG is to prepare a draft Assessment Report and Assessment Summary Report for each stock by 16 November 2012 at the latest. These reports are to address the SAW/SARC 55 TOR, taking into consideration the issues confronting each stock and will be peer reviewed at the December SARC 55.

Models Meetings

A number of issues confront the assessments of Gulf of Maine and Georges Bank cod, some of which are common to both. For GOM cod, these include but not exclusively:

- Implications for use of landings per unit effort (LPUE) and use of offshore NEFSC survey strata
- Fishing fleets to be included in the models
- Treatment of uncertainty in survey calibration coefficients
- Assumptions on flat vs. domed survey and fishery selectivity
- Pre – 1982 recruitment and implications for stock – recruitment relationship and reference points
- Implications for possible changes in natural mortality

For GB cod, a major issue is the source of the strong retrospective pattern observed in past assessments. This possibly implicates some of the following:

- Unreported catch
- Fishing fleets to be included in the models
- Treatment of uncertainty in survey calibration coefficients
- Assumptions on flat vs. domed survey and fishery selectivity

- Assumptions on the stock – recruitment relationship and implications for reference points
- Implications for possible changes in natural mortality

The draft Models Issues meeting agenda is provided in Appendix 1. A number of changes were made to the agenda to meet contingences which arose during the meeting.

A list of background documents and working papers (WP) considered during the meeting are provided in Table 4. All papers and analyses conducted during WG meetings are preliminary and have no official status with the agency.

The participant list is provided in Table 5.

The reports by the rapporteurs (J. Blaylock, T. Chute, J. Nieland, M. Traver, S. Wigley, A. Wood) greatly assisted the drafting of this report and were much appreciated.

Gulf of Maine Cod

TOR 1: Estimation of Catch

Work for this term of reference was undertaken at the Data meeting of 27 – 31 August 2012. There was no discussion during the Models meeting relevant to this term of reference (TOR).

TOR 2: Survey and Commercial Indices of Abundance

Most of the discussion on this term of reference was undertaken at the Data meeting of 27 – 31 August 2012.

An issue arose relevant to this term of reference (TOR) that required resolution. Initial runs of the ASAP model indicated that the NEFSC spring survey catchability for the Albatross was about 0.9. With the adjustment for the Bigelow, the catchability would rise above, implying a hyper efficient survey.

When the GOM cod survey biomass was re-estimated excluding strata which were outside the stock area, the re-adjusted (for Bigelow) survey catchability was close to one (WP 41). Notwithstanding this, the WG concurred with the recommendation of the data meeting to include the NEFSC strata which extend beyond the GOM cod management area (strata 29, 30, and 36) in the assessment datasets.

Analyses were also undertaken to compare ASAP model biomass to the time series of NEFSC spring and fall survey biomass under assumptions of low and high survey catchability (WP 41). These indicated that the model estimated biomass was generally within the envelope of the survey estimated biomass.

In summary, it was considered that the model estimates of cod biomass were not inconsistent with the NEFSC survey catchabilities.

TOR 3: Stock Structure

Most of the discussion on this TOR occurred during the Data meeting. At this meeting, the WG considered two WPs which would inform future studies of the impact of stock structure on the GOM cod assessment.

WP 24 explored the GOM cod assessment sensitivity to two spatial formulations. The first undertook an assessment of the western GOM cod as a separate management unit. The aggregate NEFSC survey trends of the wGOM were essentially identical to those of the full GOM due to the dominance of the western Gulf signal. The MADMF survey operates entirely in the wGOM and exhibited similar trends to those in the full Gulf. A wGOM cod dataset

consisting of estimated wGOM catch and survey information specific to the western Gulf was prepared and an ASAP assessment model conducted. A comparison of the trend in wGOM to full GOM cod spawning stock biomass (SSB) indicated that whereas SSB in the western Gulf has varied without a significant long-term trend, that of the entire Gulf has declined. While this could imply large declines in the eGOM biomass, it was noted that given non-linearities in the models, eGOM biomass does not necessarily equal the total minus that the wGOM.

The second sensitivity analysis (WP 24) considered GOM and Georges Bank (GB) cod as a unit stock. Again, a number of adjustments were made to the catch and survey datasets to allow assessment of the unit stock. The SSB, recruitment and fishing mortality trends in the unit stock were compared to those based on the sum of the GOM and GB cod assessments. Overall, the unit stock appeared to agree with the sum of the parts. This was an exploratory analysis and a number of issues would need to be dealt with before undertaking future such analyses of a unit GOM + GB cod stock.

WP 22 undertook a simulation study which examined the impact of mixing amongst GOM, GB and NAFO Div. 4X cod on VPA and ASAP assessments of GOM cod. The uncertainties introduced into the estimates of SSB and F indicated that the ASAP model was more robust than the VPA. Overall, the results indicated that the lack of consideration of inter-stock mixing had little impact on the GOM cod assessment results. The importance of the quality of the catch information was highlighted.

There appeared to be substantial bias in F in the VPA runs (WP fig. 11) which required explanation. It was also noted that the inter-stock mixing rates were those of Miller and Tallack (2007) which had been updated (upwards) at this meeting. Further, the instantaneous mixing rates estimated by Miller and Tallack (2007) were annual whereas the study used these only for half a year, implying that mixing in the model was underestimated by 50%. It was further noted that the mixing rates were representative of the relative biomass of GOM, GB and 4X cod during 2003 – 2006, the period of the tagging study (see TOR 4). The importance of the mixing rates relative to natural mortality (assumed as 0.2) as well as their sensitivity to tag reporting rates was highlighted.

While the study is a work in progress with many assumptions and issues to be resolved, it highlighted the value of undertaking modeling to explore complex spatial processes influencing cod in the Gulf of Maine. It was suggested that future assessment models might more formally explore incorporating the tagging analyses into their formulations.

TOR 4: Natural Mortality

The WG considered two working papers on estimates of natural mortality for the assessment models.

The first paper (WP 31) summarized an analysis of tagging data collected during 2003 – 2006 by the GMRI. The model estimated natural (M) and fishing mortality (F) for 50 cm plus GOM, GB and 4X cod along with estimates of tag shedding, reporting rates and tag-induced mortality. A sensitivity analysis of mortality to the reporting rate of high reward tags was also undertaken. The model assumed that the reporting rates of high reward tags were constant across fleets with a region but different across regions. The model estimated M in the order of 0.56 – 0.57 for GOM and GB cod and 0.44 for 4X cod, compared to the estimate of 0.2 currently used in the assessment models. It was noted that estimates of migration between some areas (e.g. about 0.3 from GOM to GB) was high which had implications for the assessment of stock biomass (see TOR 3). There was discussion on the age groups of cod represented by the study. GOM cod of 50 cm of about 2.5 – 3 years old, implying that the estimates of M are for ages 2.5 – 3 plus with it weighted towards the younger ages. The M estimates are highly sensitive to the

reporting rate of high reward tags, with M rising from 0.3 at 50% reporting to almost 0.7 at 100%.

The second paper (WP 28) reported on a catch curve analytical approach to tagging information on GB cod during 2003 – 2008, again for 50 cm plus cod. The analysis provided estimates of M in the order of 0.74 – 0.93 with a sharp drop in these estimates when the assumed reporting rate dropped below 40%. Overall, the results of this study were broadly consistent with the first.

In the discussion on both WPs, it was noted that while the results were sensitive to a number of assumptions, there were strong indications that M during the 2003 – 2006 period could have been as high as 0.5 – 0.7 for both GOM and GB cod although the WG considered that estimates as high as 0.7 were unlikely. There were some concerns that the GOM tagging analysis was of data which had not been designed to estimate M although it was countered that many processes had been included in the model to adjust for possible sampling issues.

The WG discussed the availability of historical tagging to which the current estimates could be compared. It was reported that tagging work conducted in the Gulf of Maine area (primarily in Canadian waters including Southwest Nova Scotia, Georges Bank and Gulf of Maine) during the 1970s and 1980s suggested M estimates in the order of 0.2 – 0.3 whereas tagging in the 1990s was suggestive of M similar to the more recent results. These observations are based upon unpublished work that could not be corroborated at the meeting. Much of the historical work (e.g. Hunt et al, 1998) had been focused on cod movements and did not provide estimates of natural, fishing or total mortality. Further, concerns were raised that there was no obvious mechanism (e.g. predation) that could explain a recent increase in M, although it was countered that no specific biological mechanisms has been identified to explain the current M estimate of 0.2. The current estimate of M=0.2 for GOM and GB cod stocks is based on analyses conducted by Paloheimo (1961).

The WG considered assessment profile runs of M during the historical and more recent period to inform the discussion as to whether or not there has been a long-term change in M. When M profiling was conducted on the GOM cod data set restricted to the historical time period, an M of between 0.1 and 0.2 was determined whereas profiling conducted on the more recent time period suggested an M between 0.1 and 0.6. These profiles were consistent with the tagging evidence for M being greater than 0.2 in the 2000s and a change in M over the longer term.

The WG agreed that an option with an M change should be considered as an alternate to a Base model which would assume no change in M (i.e. M = 0.2). In this option, M during 1982 – 88 would be set equal to 0.2, during 2003 – 2011 at 0.4, with a linear ramping of M during 1989 – 2002 from 0.2 to 0.4. The implications of this M change model for stock status and the reference points would be compared and contrasted to the Base model.

TOR 5: Estimation of Fishing Mortality, Recruitment and Stock Biomass

ASAP Model

The initial ASAP formulation of the NEFSC was summarized in WP 24. This described the SAW 53 model, noting the changes that had been made in the long-term maturity at age, stock weight, discard mortality, recreational catch, and MADMF age-length keys. Additionally several sensitivity runs were explored including the investigation of varying levels of natural mortality, the inclusion/exclusion of offshore survey strata, NEFSC calibration coefficients, the assessment starting period (1982 vs. 1970, 1964) and age range (1 – 9+ vs. 11+) and survey selectivity assumptions (domed vs. flat-topped). It was noted that the SAW 55 WG data meeting had recommended that the fishery and recreational LPUE not be used in the Base model due to

changes in the distribution of the stock and the regulations since the mid-2000s. A catch curve analysis and examination of proportions at age in the fishery and surveys suggested no strong doming in either fishery or survey selectivity with a suggestion that surveys tend to be less domed that the fishery.

A SAW 53 – SAW 55 bridging analysis was presented which described the effect of sequentially making eight adjustments to the ASAP model. Most of these adjustments related to fishery and survey data updates. This was followed by SAW 55 developments on a number of model elements as well as sensitivity analyses. Highlights of these explorations are provided below.

The sensitivity of the model to the inclusion of the MADMF spring and fall surveys highlighted the lack of fit of the fall survey and the low information content in the older ages of the spring survey. However, it was noted by the WG that the process of doing survey by survey model sensitivities was not very informative.

It was reported that use of either survey aggregate numbers or biomass in the model fitting resulted in slightly different trends, with the biomass time series producing less stock biomass in more recent years. Discussion ensued on how the model weights each series differently in the fits which could account for the difference. The WG pursued this topic later in the meeting.

Retrospective patterns of survey (NEFSC spring and fall and MADMF spring) catchability (q) changes were examined. While q has remained relatively stable for the NEFSC fall and MADMF spring surveys, such is not the case for the NEFSC spring survey which exhibited a sharp increase in q during the 1990s. The WG discussed potential causes for this pattern which could be included in the model by splitting this series pre and post 2000 but this was not pursued further during the meeting.

The sensitivity analysis of the ASAP model to the Albatross to Bigelow calibration indicated that the main effect was an increase in the uncertainty in recent biomass rather than adding bias.

In relation to the inclusion of LPUE indices, it was noted that splitting the commercial series pre and post 2006 would produce the same stock trends although it was suggested that splitting the time series could increase model precision. It was considered useful to include the LPUE sensitivity runs in the final report.

A number of catch and fleet related explorations were reported to the WG. These included fleet structure (one fleet recommended), fishery selectivity (fishery tends to dome more than survey) and assessment start data (explored 1964, 1970 and 1982 each having different data inputs and model assumptions). It was noted that the model was not very sensitive to assumptions on the error in the catch with CVs of 0.01 to 0.3 providing similar results. In relation to fleet selectivity, both non-parametric (i.e. by age) and parametric (e.g. single and double logistic models) were attempted with the non-parametric approach used in the recommended ASAP Base model. A number of temporal blocks were explored with a 3 – block model finally recommended.

A significant feature of these ASAP formulations was that a retrospective pattern had become evident, which had not been prominent in past assessments. During the bridge building process from the SAW 53 model to an updated SAW 55 version, it was evident that the application of alternate discard mortality rates resulted in retrospective patterns. A sensitivity run showed that assuming 100% discard mortality could reduce the retrospective patterns of the updated SAW 55 model.

SCAA Model

The initial SCAA[1] formulation by Butterworth and Rademeyer was summarized in WP 23. While having several similarities to the ASAP model, it had a number of important differences, which are highlighted below.

The SCAA formulation started in 1964 and ended in 2012. Before 1982, 1964 – 69 NEFSC survey length frequency (with growth assumptions) and 1970 – 81 NEFSC survey at age data were used as indices of abundance. In contrast to the ASAP model, which ended in 2011, the SCAA model ended in 2012. The 2012 spring surveys were used and the catch was assumed equal to that in 2011. Catch was used as an input (negligible error assumed). The Bigelow/Albatross calibration was incorporated into the formulation to allow re-estimation based on cross – cohort and annual processes. Survey aggregate biomass rather than numbers was used.

The 1964 numbers at age were fit using three parameters for ages 1 – 3 and a phi depletion parameter for the older age groups. The weighting of the survey data in the objective function was composed of the observed variance plus variance estimated within the model (sum of squares approach). This is in contrast with the ASAP formulation in which the survey weighting was composed of the square of observed and model estimated deviation (square of sums approach). Also, the proportions at age were fit with an adjusted lognormal model to take into account variation at age, in contrast to the multinomial model used in ASAP.

A domed selectivity was allowed in both the survey and the fishery with one fleet and two temporal blocks employed.

The SCAA results indicated that, as with the ASAP, the fishery selectivity was more domed than the surveys. There was considerable discussion on the adjusted lognormal weighting of the proportions data as well as the use of the aggregate vs. number surveys indices, which is discussed further below. In relation to the 1964 – 69 survey length frequencies, it was noted that the spring fits exhibited over-representation of small cod compared to under-representation of large cod, which could be due to the six months of growth between the two surveys. The treatment of the Bigelow / Albatross calibration was not a large source of difference between the two models. Regarding the starting year, use of 1970 resulted in similar historical trends in spawning stock biomass (SSB) although use of 1964 exhibited appreciable differences from historical reconstructions employed in ASAP. Specifically, the SCAA formulation estimated higher SSB during 1964 – 69 than the ASAP formulation. The WG noted the need to understand the differences in the model formulations during 1964 – 69 as these are important to the reference point discussion. This topic is discussed further below.

Model Comparisons

There were numerous differences between the ASAP and SCAA formulations, the relative impacts of which needed to be explored. Further, WP 27 noted that many of the SCAA model runs did not meet ADMB-published criteria (e.g. magnitude of gradients at solution) generally used to evaluate assessment model convergence and stability. While the two models were providing similar overall results, there was concern that the SCAA model was displaying characteristics of over-parameterization. There were also initial concerns of compiler – based differences although these subsequently proved unfounded. Further investigation indicated that the gradient issue was a feature of the manner in which the SCAA effected exact solutions to the Baranov catch equation; when some catch uncertainty was introduced, gradients generally satisfied the requirements for convergence.

The chair and the lead scientist of each assessment prepared a table of what they considered were the key differences between the two models (columns 1 – 3 of Table 1). This was then discussed by WG and used to define a common model formulation (Comparison Run in

[1] SCAA is a generic term for statistical catch at age models. While ASAP is also a SCAA model, the terms ASAP and SCAA are used in this report to distinguish between the two models considered by the WG.

Table 1). The intent was to construct formulations of ASAP and SCAA which were as similar as possible, which would in itself lead to better understanding of the differences in each formulation. This common formulation would then be used as the basis for further model exploration. It was recognized that some differences between ASAP and SCAA would be difficult to resolve due to the need to make significant software coding changes.

Evaluation of the results of the ASAP and SCAA comparison runs indicated that there were still differences of note, the possible sources of which were the use of age zero data in the SCAA, the weighting on the survey indices and the weighting on the proportions data. Regarding the age zero data, after discussion, it was agreed that the SCAA formulation not use any age zero information. Regarding the weighting on the survey indices, further explorations which employed the ASAP model – estimated component of the variance in the SCAA comparison run indicated that the two weighting approaches were comparable and thus both could be used in the Base formulation.

The outstanding issue was the weighting on the survey and fishery proportions at age. It was noted that the age-dependence of the variance on the proportions at age was not consistent with either an adjusted lognormal or a multinomial assumption and that in order to properly account for this in the model fitting process, it was necessary to employ an age-dependent weighting, as the refined adjusted log-normal does (WP20). While use of the multinomial would not produce biased estimates, it would likely result in the variance being over-estimated. Further, the AIC criterion would not be valid in model selection.

In response to these observations, WP 21 described a simulation which argued that use of the multinomial assumption to weight the proportions at age fits was superior to that of the adjusted log-normal, as is used in the SCAA formulation. It noted that while use of both the multinomial and adjusted log-normal assumptions tended to result in domed selectivities being preferred, this tendency was more pronounced with the adjusted lognormal. Thus, both approaches have a tendency to estimate domes where none exist, or stronger domes than actually exist. This argued for greater reliance on external information for the existence of domes. Or as a corollary, more forcing of flat tops in the selectivity functions unless strong external evidence is available to support the presence of a domed selectivity.

The conclusions of this simulation were countered in WP 26. While there were detailed issues which required attention and needed follow-up (i.e. equation typos, using mean-unbiased simulated data to test a median-unbiased estimator was questioned and sum of proportions equaling 1 needed to be confirmed), the more fundamental issue was the validity of the simulation. It was contested, based upon WPs 33 and 34, that the simulations undertaken in WP 21 were only indicative, not definitive, in that they were not 'conditioned' to the GOM cod situation, instead employing 'cod-like' data. This led to discussion on whether or not these contentions were valid. The WG did not resolve these issues and recommended further follow-up.

Table 1. Synopsis of Issues and Specification of Comparison Run; N/A indicates that issue not relevant to comparison run; red indicates preference of Scientist 2

Chair	Scientist 1	Scientist 2	Comparison Run
General Issues			
Address convergence issues	Address convergence issues		Will assist in addressing issues
Data Issues			
		Include catches prior to 1964: start in 1945, don't use	N/A
Assessment start year: 1964, 1970, 1982	Assessment start year: 1964, 1970, 1982	Assessment start year: 1945, 1964, 1965,	1982

		1970, 1982	
Assessment end year: 2011, 2012	Assessment end year: 2011, 2012	Assessment end year: 2011, 2012	2011
		Start age (include age 0 catches: Yes, No	Preference to take out but lower priority; start at age 1
	End Age (plus group)		Age 9+
Fit of age-aggregated survey index: biomass, numbers	Fit of age-aggregated survey index: biomass, numbers	Fit of age-aggregated survey index: biomass, numbers	Numbers
Bigelow/Albatross calibration adjustment (re-estimation included in model)	Bigelow/Albatross calibration adjustment (re-estimation included in model); Other calibrations?	Bigelow/Albatross calibration adjustment (re-estimation included in model): Yes, No	Do not use re-estimation in model
		Use of 1960s survey length distributions: Yes, No	N/A
Process Issues			
		Start year number at ages estimated: Likelihood based estimation (e.g. 3 for 1964), all ages	All ages (1 – 9+)
Assumed CV in catch	Assumed CV in catch	Catch uncertainty CV post-1982 (unbiased): 0.05, 0.10, 0.20	CV = 0.05
		Catch uncertainty pre-1982: Bias from 1945 or from 1964: +25% +36%; CV from 1964: 0 0.05 0.10 0.20; CV from 1945 to 1963: 0.10 0.20 0.30 0.50	N/A
Treatment of error (multinomial vs. adjusted lognormal) in proportions at age	Treatment of error (multinomial vs. adjusted lognormal) in proportions at age	Treatment of error (multinomial vs. adjusted lognormal) in proportions at age: adjusted lognormal by age, sqrt(p) by age multinomial	ASAP: Multinomial SCAA: sqrt(p) at age
Survey and fishery selectivity assumptions	Survey and fishery selectivity assumptions	Survey and fishery selectivity assumptions: domed (estimated) , flat	Survey and fishery both flat
Fishery selectivity time blocks: preference for 3		Fishery selectivity time blocks: Two blocks (split in 1991), three blocks (splits in 1989 and 2004)	3 Blocks: 1982-1988, 1989-2004, 2005-2011
		Commercial selectivity prior to	

		1982: equal to first post-82 block, ascending limb moved 1 year to the left from first post-82 block, as last but moved 2 years	N/A
Natural mortality: Pre & post 2003 explorations		Natural mortality: 0.3, 0.4, 0.5, prior, higher from 2004	0.2
Stock – recruit relationship		Stock-recruit relationship: Internal fit, external fit, average over specified period, Beverton-Holt Adjusted B-H, Ricker	S-R turned off with sigma r being 0.5 for both models
		Additional variance for surveys: Estimate in model fit, fix on input (CV additive linearly or additive quadratically)	ASAP: $(sd + sd_{add})^2$ SCAA: $sd^2 + sd_{add}^2$

ASAP Base Formulations

Based on the WG discussion, two Base model formulations were agreed to. One would assume a long-term increase in natural mortality, as discussed in TOR 4 (herein termed M change model). The other would assume that M has remained at 0.2 over the long-term (herein termed the M constant model). In the case of the latter, any apparent retrospective pattern would be adjusted for in projections but it would probably not be necessary to adjust for this in the case of the former. It was noted that for both Base models, it would be useful to undertake a sensitivity analysis on the impact of employing 100% discard mortality. The WG left decisions on the most pertinent additional sensitivity runs, of which a number were discussed, to the discretion of the lead scientist.

Table 2 outlines the data and model specifications of the Base formulations agreed to by the WG. Observations on some of these are in order:

<u>Start year</u>

Table 2 indicates the Base formulation assuming a start year of 1982. It was agreed that modeling should explore population dynamics back to 1934 (changed to 1932 after the meeting), which has implications for a number of the model elements (e.g. pre-1982 catch uncertainty which is indicated in the table). While the WG had concerns with the accuracy of the catch information prior to 1963 and even 1970, it saw value in determining the consistency of historical population dynamics with those more recent.

For 1963 – 69, the contribution of the survey length frequencies needs to be evaluated. Discussion at the meeting indicated that these data may not be influential. This requires further exploration.

<u>End year</u>

WP 32 noted that the inclusion of the 2012 information had implications for recent stock trends, with 2011 SSB declining from about 14kt to 12kt when the 2012 data are included. A number of assumptions are needed to allow inclusion of the 2012 data. A 2012 catch has to be assumed as well as the weights at age. Further, the results are based on only two survey points. The WG concurred that the 2012 not be included in the Base models. However, comment on the low results of the 2012 surveys, plus the low catch by the fishery in 2012 to date, would be included in the narrative on the projections.

Biomass vs. Numbers Aggregate Indices

The WG considered the criterion on which to decide whether or not to use survey aggregate numbers or biomass in the model fits. Numbers would tend to weight the fit towards the younger age groups while biomass would tend to weight the fit towards the older age groups. The WG agreed that the index which produces the lower CV on the SSB is preferred. However there was a concern that smoothed weights are used for the biomass index which might lead to negatively biased variance estimates. Notwithstanding this, while the ASAP model uses aggregate numbers in its tuning, the WG agreed that sensitivity to the aggregate biomass numbers should be examined. Similarly, while the SCAA uses aggregate biomass, sensitivity to the aggregate numbers should be examined.

Bigelow/Albatross Calibration

The WG considered that the effect of including re-estimation of the Bigelow/Albatross calibration in the ASAP model, as is done in the SCAA model, was small. Given the software coding implications of this modification, it was not considered essential to incorporate this in the ASAP Base models.

Start Year Numbers at Age

Prior to 1982, the SCAA formulation uses a phi term to estimate depletion rate of age groups not separately estimated. The WG noted that this could be an important difference between the two modeling approaches. It recommended that the ASAP formulations explore the sensitivity of pre-1982 stock determinations to the number of age groups included in the phi term.

Treatment of Error in Proportions at Age

The WG considered that the implications of use of the multinomial or adjusted log-normal were more related to precision of the current SSB rather than bias. Further work on this is recommended. For the ASAP Base models, it agreed that the multinomial be used.

Fishery and Survey Selectivities

The WG agreed that survey selectivity be assumed to be flat while the fishery selectivity would be estimated although the details of these analyses were left to the next meeting. In the case of the former, ages 6 plus would be considered fully recruited. In the case of the latter, three time blocks would be used with the assumed age of full recruitment moving from age 4 to 6 across blocks.

Survey Weightings

The explorations undertaken by the WG indicated that the differences in the survey weightings between the ASAP (square of sums approach) and the SCAA (sum of squares approach) were not consequential and thus each formulation was free to use its preferred approach.

Table 2. Specifications of Gulf of Maine Cod ASAP Base Assessment Models

Model Element	Base Case Formulation
Assessment start year	1982
Assessment end year	2011
Start age	Age 1
End Age (plus group)	Age 9+
Fit of age-aggregated survey index: biomass, numbers	Numbers with sensitivity to Biomass
Bigelow/Albatross calibration adjustment (re-estimation included in model)	No re-estimation in model
MADMF survey ages	Ages 1 - 6
Start year number at ages estimated	All ages (1 – 9+)
Assumed CV in catch	CV = 0.05
Catch uncertainty pre-1982	Bias (1945 – 64): 0.32 CV (1945-63): 0.4 CV (1964-81): 0.2
Treatment of error (multinomial vs. adjusted lognormal) in proportions at age	Multinomial
Survey and fishery selectivity assumptions	Survey: flat age 6+ = 1 Fishery: estimation process to be resolved
Fishery selectivity time blocks	Three blocks: 1982-1988, 1989-2004, 2005-2011
Commercial selectivity prior to 1982	Ascending limb moved 1 year to the left from first post-82 block
Natural mortality	Option 1) 0.2 for all years Option 2) 0.2 for 1982-88, 0.4 for 2003-2011, ramp during 1989-2002
Stock – recruit relationship	S-R internal & external (BH, Ricker) with $S_r = 0.6$ and associated sensitivity runs Proxy (Fx%) with consideration of percentage (see work on winter flounder)
Recruitment Lambda	0.2
Survey weighting	Square of sums

TOR 6: Reference Points

The WG held discussion related to this TOR to prepare for its next meeting.

It was agreed that the next meeting explore both proxy and analytical reference points (RPs) for GOM cod. Regarding the proxies, an evaluation of the current F40% proxy is in order, comparable to that done for SNE/MA Winter Flounder during SAW 52. An issue with the proxy

RPs is the recruitment time series from which to sample. The current approach should be a starting point for discussion.

Regarding analytical reference points, the choice of the stock – recruitment (SR) relationship will be important. The ASAP software allows use of the Beverton and Holt (BH) which should be used in the Base model comparisons. The SCAA formulation allows use of both the BH and Ricker SR functions as well as an adjusted BH to allow the upward and downward slopes of the relationship to be estimated independently. There were concerns raised about the implied underlying biology of this relationship, although the WG noted that this relationship may be performing like a smoother of the stock – recruit time series.

One of the most important considerations of the SR relationships is the sensitivity of the RPs (e.g. B_{MSY}) and starting SSB to data and model assumptions used in the 1963 – 1981 population constructions. The WG agreed that the sensitivity of RPs and starting SSB to the internal vs. external estimation of RP, the SR relationship, the number of age groups in the start year phi term and the starting year of the analysis be explored.

It was suggested that current ecosystem modeling efforts at the NEFSC might inform discussion on the productivity of cod in the Gulf of Maine area. In response, it was considered that these efforts were likely too general to be of utility to the SAW 55 and might detract from the RP discussion. Thus, this avenue will not be pursued.

TOR 7: Evaluation of Stock Status

This TOR will be addressed in the assessment report of the SAW 55 WG.

TOR 8: Projections

Work on this TOR will be undertaken during the Models and Reference Points meeting of 29 October – 2 November 2012.

TOR 9: Research Recommendations

Work on this TOR will be undertaken during the Models and Reference Points meeting of 29 October – 2 November 2012.

Georges Bank Cod

TOR 1: Estimation of Catch

Work for this term of reference was undertaken at the Data meeting of 27 – 31 August 2012. There was no discussion during the Models meeting relevant to this term of reference (TOR).

TOR 2: Survey and Commercial Indices of Abundance

Work for this term of reference was undertaken at the Data meeting of 27 – 31 August 2012. There was no discussion during the Models meeting relevant to this term of reference (TOR).

TOR 3: Stock Structure

As noted under the TOR 3 for GOM cod, most of the discussion on this term of reference was undertaken at the Data meeting of 27 – 31 August 2012.

An issue was brought to the attention of the chair by S. Cadrin on an apparent discrepancy between GB and eGB estimates of cod survey swept area biomass. Resolution of this issue is underway and will be reported at the Models and Reference Points meeting of 29 October – 2 November 2012.

TOR 4: Natural Mortality

Discussion conducted on this TOR was reported under GOM cod. As with GOM cod, M profiling suggested that a long-term change in M is plausible. This profiling suggested an historical and recent M of 0.2 and 0.4 respectively. The WG agreed that an option with an M change should be considered as an alternate to a Base model which would assume no change in M (i.e. M = 0.2). In this option, M during 1978 - 1990 would be set equal to 0.2, during 2003 – 2011 at 0.4, with a linear ramping up of M during 1991 - 2002 between 0.2 and 0.4. The implications of this M change model for stock status and the reference points would be compared and contrasted to the Base model.

The WG noted that the 2011 TRAC assessment of the eastern Georges Bank Cod stock which is transboundary with Canada (Yang et al, 2011) also considered two natural mortality options, one with M equal to 0.2 for all years in the assessment (1978 – 2010) and the other with M on ages 6+ set to 0.5 after 1994. Thus, the WG options on natural mortality are consistent with the TRAC analyses.

TOR 5: Estimation of Fishing Mortality, Recruitment and Stock Biomass

The GARM III assessment used a virtual population analysis (VPA) and thus a bridging analysis to ASAP was considered by the WG (WP 30). Of note was the replacement of the MRFSS data by that of MRIP and the use of the updated discard mortality rates discussed at the Data meeting. Both changes had minimal impact on the assessment.

The GARM III VPA had split the time series of the three surveys (NEFSC spring, NEFSC fall and DFO) to address the retrospective bias. This split no longer addresses this issue with the pattern becoming more pronounced.

The ASAP formulation was largely based on the previous VPA. Age one to 10+ for 1978 to 2011 were modeled. The CV on the catch was initially set equal to 0.2. The fishery was modeled using two selectivity time blocks (1978 – 1993, 1994 – 2011). The three surveys were used as in the VPA except no 1994/95 split was imposed. Natural mortality was set equal to 0.2. While both the survey and fishery selectivities were freely estimated, both indicated flat topped patterns. The model displayed a lack of fit to the 1978 – 88 catch which was considered to be largely due to large mesh discards. Most significantly, the retrospective pattern, while diminished compared to the VPA, was still present.

A number of sensitivity analyses were considered by the WG. Assuming upper and lower bounds on the Bigelow / Albatross calibration changed post -2005 population estimates only marginally. Assuming 100% discards rates compared to those used in the Base model resulted in little change. On the other hand, either decreasing 1978-94 catch or increasing 1995 – 2011 catch by significant amounts (50% and 300% respectively) would largely resolve the retrospective pattern. The profile of the 1978 – 2011 M was very flat until values greater than 0.5, above which the value of the objective function increased. It was noted that the profile should be conducted without the recruitment deviation option in the model turned on, which was done and showed a similar pattern. Biomass in a run conducted with commercial LPUE included declined faster than that of the Base model until about 1994 after which it remained relatively flat but

higher than in the Base model. Splitting the survey time series in 1994 did not improve the retrospective pattern. Splitting the NEFSC spring time series to account for the change in net type (Yankee 41 to 36) did result in a modest change in historical biomass, compared to the Base model and the WG agreed to include this split in the final formulation. An examination of retrospective changes in the catchability of the three surveys indicated that these have drifted higher since 1995. It was agreed that this diagnostic be added to the SSB, F and recruitment suite used in describing retrospective patterns. A run which included two fleets (Canada and US) resulted in a stronger retrospective pattern and thus the one fleet model was retained for the final formulation. Runs using a range of CVs on the catch suggested a value of 0.1 which was initially accepted by the WG but later reduced to 0.05 after consideration of the CV used in the GOM cod model, which improved the fit to the 1978-1988 catch data.

The WG discussed the implications of the strong retrospective pattern, which was relatively resilient to many changes made to the model. The WG queried how severe a retrospective pattern needed to be for an assessment model to be rejected. It was generally agreed that it is better to accept an assessment model and apply a retrospective adjustment rather than resort to assessment approaches based on indices alone. The latter make implicit assumptions which are more explicitly stated and examined in age-structured population modeling.

The potential causes for the retrospective pattern were discussed, the underlying process being the disappearance of cod from the population (either through fishing or natural causes), increases in survey catchability or a combination of both. It was suggested that autocorrelation in the residuals could also be a cause although this was examined in SNE/MA winter flounder and determined not to be a factor. This requires follow-up as a research recommendation.

The steady increase in the survey catchabilities, particularly that of the DFO winter survey, was discussed at length. There was speculation on the distributional processes during the cod spawning season that could lead to a change in survey catchability although these could not be corroborated. An ASAP run without the DFO winter survey made the retrospective worse. An envelope analysis which compared the catch and survey swept area biomass over a range of assumed q (0.1 – 1.0) and F (0.1 and 1.0) values indicated that the SSB estimates in the ASAP Base run were generally consistent with the biomass envelopes implied by the surveys. Overall, it was considered that, rather than trying to adjust for survey q changes, these be used as a diagnostic of the influence of changes model changes on the retrospective pattern.

Natural mortality as a source of the retrospective pattern was discussed by the WG. The tagging analysis (TOR 4) provided evidence that M during 2003 – 2006 was higher than the previously assumed 0.2. Further, profiling of M for the 2003 – 2011 period indicated an M of 0.4. An M profile conducted on the 1978 – 2002 dataset indicated that M could be as low as 0.1 supporting the contention that M had increased more recently. Runs exploring the year in which an M increase may have occurred indicated 2003. The WG agreed on an M change model in which M during 1978 - 1990 and 2003 – 2011 was assumed to be 0.2 and 0.4 respectively, with a linear ramp during 1991 – 2002. Preliminary ASAP runs with this assumption produced a much improved retrospective pattern.

The WG agreed that two Base model options be considered – M change and M constant (0.2) models. A retrospective adjustment would be made to the latter but not the former model. The rest of the model elements are indicated in Table 3. Further adjustments to these Base models would include changes in the catch, as was done in the sensitivity runs described above. The final report should also include the sensitivity analyses on discard mortality, commercial LPUE and the Bigelow / Albatross calibration.

Table 3. Specifications Georges Bank Cod ASAP Base Assessment Model

Model Element	Base Case Formulation
Assessment start year	1978
Assessment end year	2011

Start age	Age 1
End Age (plus group)	Age 10+
Fit of age-aggregated survey index: biomass, numbers	Numbers
NEFSC fall & DFO winter surveys	No time blocks
NEFSC spring survey 41/36 split	1978-81, 1982-2011 blocks
Bigelow/Albatross calibration re-estimated in model	No
Start year number at ages estimated	All ages (1 – 10+)
Number of fleets	US & Canada combined
Assumed CV in catch	CV = 0.05
Treatment of error (multinomial vs. adjusted lognormal) in proportions at age	Multinomial
Survey and fishery selectivity assumptions	Survey flat Fishery flat
Fishery selectivity time blocks	Two blocks: 1978 - 1993, 1994 - 2011
Natural mortality	Option 1) 0.2 for all years Option 2) 0.2 for 1978 - 1990, 0.4 for 2003-2011, ramp during 1991-2002
Stock – recruit relationship	S-R internal & external (BH) Proxy (Fx%) with consideration of percentage (see work on winter flounder)
Recruitment Lambda	0
Additional variance for surveys	Square of sums

TOR 6: Reference Points

The WG discussed analyses related to the RP discussion to be held at the next meeting. As with GOM cod, it agreed that both proxy and analytical RPs (either estimated internally or externally to the ASAP model) should be estimated. Regarding the proxy RPs, an analysis corroborating (or not) the current F40% basis of B_{MSY} is required. This could be similar to that undertaken during SAW 52 on SNE/MA winter flounder. Regarding the analytical RPs, given the stock – recruitment relationship indicated in the preliminary ASAP runs, BH and Ricker models may not be viable. The WG recommended that the RP analysis start with the current approach and explore options from there.

The importance of communicating the uncertainties in the analyses was highlighted by the WG. For both GOM and GB Cod, it will be essential that subsequent steps in the management process are fully aware of the relative support for the model options as well as the associated uncertainties.

TOR 7: Evaluation of Stock Status

This TOR will be addressed in the assessment report of the SAW 55 WG.

TOR 8: Projections

Work on this TOR will be undertaken during the Models and Reference Points meeting of 29 October – 2 November 2012.

TOR 9: Research Recommendations

Work on this TOR will be undertaken during the Models and Reference Points meeting of 29 October – 2 November 2012.

References

Hunt, J.J., W.T. Stobo and F. Almeida. 1998. Movement of Atlantic Cod, Gadus morhua, tagged in the Gulf of Maine area. Fish. Bull. 97: 842 – 860.

Miller, T. J. and Tallack, S. T. M. 2007. Estimating instantaneous rates of regional migration and mortality from conventional tagging data. 2007 GARM Data Methods Working Paper C.3.

Paloheimo, J.E. 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: 645 – 662.

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

Table 4. List of Working Papers and Background Documents considered at the SAW 55 Model Issues Meeting

WP #	Topic	TOR	Stock	Author	Title
9 (Updated)	Commerical LPUE	2	GOM Cod	Palmer	Commercial catch-per-unit-effort (CPUE) indices for Gulf of Maine Atlantic cod (Gadus morhua).
20	Status estimation	5	GOM Cod	Butterworth, Rademeyer	A Check on the appropriateness of the assumption of multinomial distributions for catch-at-age residuals in the 2011 ASAP assessment of Gulf of Maine Cod
21	Status estimation	5	Both	Legault	Simulation test of age composition error distributions
22	Stock Structure	3	GOM Cod	Chen	Impacts of stock mixing on the GOM cod assessment
23	Status estimation	5	GOM Cod	Butterworth, Rademeyer	Applications of Statistical Catch-at-Age Assessment Methodology to Gulf of Maine cod, October 2012
24	Status estimation	5	GOM Cod	Palmer	Presentation of Statistical Catch at Age assessment of Gulf of Maine Cod
25	Status estimation	5	Both	Butterworth	Suggested Sequential Structure for Simulation Testing of Assessment Methods
26	Status estimation	5	Both	Butterworth	Comments on Legault paper on simulation tests of age composition error distributions
27	Status estimation	5	Both	Brooks	Criteria for evaluating model convergence and stability
28	Stock Structure	3	GB Cod	Shepherd	Alternative approach to analysis of Georges Bank cod tagging analysis
29	Status estimation	5	GOM Cod	Rademeyer	Survey length frequencies of 1964 - 69
30	Status estimation	5	GB Cod	O'Brien	Presentation of Georges Bank cod assessment
31	Natural Mortality	4	Both	Miller	Tag-recover models for GMRI Atlantic cod tagging project
32	Status estimation	5	GOM Cod	Rademeyer	Four runs of Butterworth and Rademeyer model to explore convergence issues
33	Status estimation	5	GOM Cod	Butterworth	The SISAM methods evaluation scheme
34	Status estimation	5	GOM Cod	Butterworth	XSA ADAPT SAM comparison
35	Status estimation	5	GB Cod	Nitschke	Georges Bank cod envelope analysis
36	Status estimation	5	GB Cod	Rago	Georges Bank cod envelope analysis
37	Status estimation	5	GOM Cod	Rademeyer	Comparative run 2
38	Status estimation	5	GB Cod	O'Brien	Reruns & Base formulations for Georges Bank Cod
39	Status estimation	5	GOM Cod	Rademeyer	Comparative run 3
40	Status estimation	5	GOM Cod	Palmer	ASAP analyses on blocks and M
41	Status estimation	5	GOM Cod	Palmer	Survey q sensitivity
Background	Status estimation	5	Both	NMFS, SSC & TRAC	SAW, SSC & TRAC Assessment and related documents during 2009 - 2012
Background	Status estimation	5	GOM Cod	Palmer	Presentation to SSC meeting of 25 Jan 2012: overview of 2011 GOM Cod assessment
Background	Status estimation	5	GOM Cod	Butterworth, Rademeyer	Applications of Statistical Catch-at-Age Assessment Methodology to Gulf of Maine cod
Background	Status estimation	5	GOM Cod	Butterworth, Rademeyer	An Investigation of Differences Amongst SCAA and ASAP Assessment (including Reference Point) Estimates for Gulf of Maine Cod
Background	Status estimation	5	GOM Cod	Butterworth	Response to comments by NMFS on reference point estimation for Gulf of Maine Cod
Background	Status estimation	5	GOM Cod	NMFS	Saw 53 Assessment Panel Report
Background	Status estimation	5	GOM Cod	NMFS	SAW 53 Assessment Report
Background	Status estimation	5	GOM Cod	NMFS	SAW 53 Assessment Summary Report
Background	Status estimation	5	GOM Cod	Pope	Status Quo TACs
Background	Status estimation	5	Both	Butterworth, Punt	Some preliminary examinations of the potential information content of age-structure data from Antarctic Minke whale research catches
Background	Status estimation	5	Both	Restrepo et al	Estimates of selectivity for eastern Atlantic Bluefin Tuna from catch curves
Background	Commercial LPUE	2	Both	NEFSC	Utility of Catch and Landings Per Unit of Fishing Effort (CPUE and LPUE) in Gulf of Maine and Georges Banks cod Stock Assessments
Background	Discards	1	Both	NEFSC	Establishing Discard Mortality Rates for Atlantic Cod Stock Assessments Using a Modified Delphi Technique
Background	Natural Mortality	4	GB Cod	Hunt, Stobo, Almeida	Movement of Atlantic Cod tagged in the Gulf of Maine
Background	Natural Mortality	4	Both	Paloheimo	Tagging analyses off south western nova scotia

Table 5. List of Participants of the SAW 55 Models Issues Meeting

Last Name	First Name	Affiliation	E-mail
Blaylock	Jessica	NEFSC	Jessica.blaylock@noaa.gov
Brooks	Liz	NEFSC	Liz.brooks@noaa.gov
Butterworth	Doug	UCT	doug.butterworth@uct.ac.za
Cadrin	Steve	SMAST	scadrin@umassd.edu
Dean	Micah	MA DMF	Micah.Dean@state.ma.us
Deroba	Jon	NEFSC	jonathan.deroba@noaa.gov
Fairbrother	Alison	Public Trust Project	alison.fairbrother@gmail.com
Giacalone	Vito	Northeast Seafood Coalition	summer-breeze@mindspring.com
Hart	Dvora	NEFSC	deborah.hart@noaa.gov
Hogan	Fiona	NEFMC	fiona.hogan@nefmc.org
Miller	Tim	NEFSC	timothy.j.miller@noaa.gov
Nieland	Julie	NEFSC- Woods Hole	julie.nieland@noaa.gov
Nies	Tom	NEFMC	tnies@nefmc.org
Nitschke	Paul	NEFSC	paul.nitschke@noaa.gov
O'Boyle	Robert	Beta Scientific Consulting	betasci@eastlink.ca
O'Brien	Loretta	NEFSC	loretta.o'brien@noaa.gov
Odell	Jackie	Northeast Seafood Coalition	jackie_odell@yahoo.com
Palmer	Mike	NEFSC	Michael.Palmer@noaa.gov
Rago	Paul	NEFSC	Paul.Rago@noaa.gov
Richardson	David	NEFSC-Naragansett	David.richardson@noaa.gov
Serchuk	Fred	NEFSC	fred.serchuk@noaa.gov
Shepherd	Gary	NEFSC	gary.shepherd@noaa.gov
Sosebee	Kathy	NEFSC	katherine.sosebee@noaa.gov
Terceiro	Mark	NEFSC	Mark.terceiro@noaa.gov
Traver	Michelle	NEFSC	michele.traver@noaa.gov
Wang	Yanjun	Fisheries and Oceans Canada	Yanjun.Wang@dfo-mpo.gc.ca
Weinberg	James	NEFSC	James.weinberg@noaa.gov
Wigley	Susan	NEFSC	Susan.wigley@noaa.gov
Wood	Tony	NEFSC	anthony.wood@noaa.gov

Appendix 1.

Draft Agenda*
55th Northeast Regional Stock Assessment Workshop (SAW 55)
Working Group Meeting on Models
15 – 19 October 2012
Stephen H. Clark Conference Room – Northeast Fisheries Science Center
Woods Hole, Massachusetts

TIME/DATE	TOPIC
Monday, 15 October	
09:00 – 09:15	Introductory comments (WG Chair)
09:15 – 12:00	NEFSC model formulation of GOM Cod (Palmer)
12:00 – 13:00	Lunch
13:00 – 17:00	Butterworth & Rademeyer model formulation of GOM Cod (Butterworth)
17:00 – 18:00	Model evaluation criteria (Brooks)
Tuesday, 16 October	
09:00 – 10:00	Analysis of Tagging (Miller)
10:00 – 10:15	Analysis of Tagging (Shepherd)
10:15 – 12:00	Discussion of GOM Cod issues resolution
12:00 – 13:00	Lunch
13:00 – 15:00	NEFSC bridging analysis (VPA to SCAA) of GB Cod (O'Brien)
15:00 – 17:00	NEFSC model formulation of GB Cod (O'Brien)
Wednesday, 17 October	
09:00 – 10:00	Implications of stock structure on cod stock assessment (Chen)
10:00 – 12:00	Discussion of GB Cod issues resolution
12:00 – 13:00	Lunch
13:00 – 17:00	Reruns
Thursday, 18 October	
09:00 – 12:00	Revisit of GOM Cod discussion
12:00 – 13:00	Lunch
13:00 – 17:00	Revisit of GB Cod discussion
Friday, 19 October	
09:00 – 12:00	Revisits of all discussions
12:00 – 13:00	Lunch
13:00 – 15:00	Meeting synopsis
15:00	Adjournment

* Times are approximate, and may be changed at discretion of WG chair; breaks will be held in morning and afternoon at discretion of chair; meeting is open to the public