49th NORTHEAST REGIONAL STOCK ASSESSMENT REVIEW COMMITTEE (SARC-49)

Independent reviewer's report on the 2009 Atlantic Surfclam and Butterfish Benchmark Stock Assessments

---

Michael T. Smith

Prepared for

Center for Independent Experts

The Centre for Fisheries and Aquaculture Science
Lowestoft Laboratory
Pakefield Road
Lowestoft
Suffolk NR33 0HT
England, United Kingdom

Phone: +44 1502 524242
Email: mike.smith@cefas.co.uk
www.cefas.co.uk

Table of contents

Executive Summary

Atlantic Surfclam

The SAW carried out a thorough assessment of the Atlantic surfclam fishery in accordance with the specified ToRs which were comprehensively fulfilled. A substantial amount of experimental work has gone in to estimating the efficiency of the survey gear and this has generated additional information on selectivity and influence of gear changes (pumps and cables), all of which add to the credibility of the assessment. The KLAMZ assessment methodology smoothes stock trends and provides a consistent means of carrying out stochastic projections. It provides a useful framework for stock assessment under the existing stock definitions and indicates that the stock is not overfished and that overfishing is not occurring. Nonetheless, moving to an age structured approach could provide many benefits, particularly with regard to estimating and modelling recruitment and, as noted by the SAW, future survey arrangements will make a new assessment framework a pre-requisite.

Evaluations on a regional scale indicate that stock condition and exploitation rates vary across the stock distribution, with Georges Bank unexploited and harbouring around 50% of the biomass. In more southerly regions the stock is experiencing reduced recruitment and stock levels have declined whilst fishing effort has remained stable or increased. This has resulted in increases in fishing mortality to historic maxima in the main areas of the fishery. There is a need to consider the adequacy of the current management regime in ensuring sustainable harvesting of the stock throughout its range. A sound understanding of the (stock and) recruitment dynamics for surfclam (sub-)populations would be beneficial for future management.

Butterfish

The SAW carried out a range of evaluations in order to fulfil the ToRs specified for butterfish as far as was possible given uncertainties in the available data and butterfish biology. New discard estimations were completed, which are likely to have improved the historical estimates of catch. Data from a range of surveys were considered and where possible used for assessment. There are conflicting trends in the survey data and inconsistencies in the biological data relating to natural mortality, so despite the best efforts of the SAW, the assessment results are highly uncertain. The SAW used an innovative approach to providing a prior probability distribution for survey catchability which was required to help scale the assessment. The KLAMZ delay difference model also requires assumptions regarding growth and the timing of recruitment which are difficult to accommodate for butterfish. Different data and methodologies suggest widely differing values for butterfish natural mortality, which further complicate implementation and interpretation of the assessment. Outputs from the stock assessment are therefore highly uncertain, potentially with regards to both scale and trend. An 'envelope' approach was used to indicate that the assessment outputs lie in a plausible parameter space, and this provides some increased confidence in the outputs. Despite the uncertainties in the data and assessment, it appears that fishing mortality for butterfish is very low and especially so when compared with natural mortality, which seems likely to be very high. The assessment biomass trend is driven primarily by the fall survey, which was considered the most representative, and indicates a downward trend. The evidence suggests that the butterfish stock is experiencing reduced recruitment, resulting in declining stock biomass primarily due to factors other than fishing. The SAW proposed new reference points and evaluation against these indicated that overfishing was not occurring and the stock was not overfished. However, reviewers considered that reference points based on long-term equilibrium population dynamics (e.g. MSY) were not appropriate because the stock would not reach a stable equilibrium at constant fishing mortality. Although defining reference points is problematic, the evidence presented indicates that fishing mortality is very low and overfishing is almost certainly not occurring. Although the stock is depleted, it was difficult to evaluate its status in the terms required (i.e. whether or not it is overfished), because fishing did not appear to be the cause of stock decline. The SAW produced short term projections using the KLAMZ model and the assessment results for a range of different states of nature (i.e. assumptions regarding survey catchability and future recruitment).

1. Background

This report provides an independent review of benchmark assessments of Atlantic surfclam and butterfish carried out at the Stock Assessment Workshops (SAW-49) and presented at the 49th Northeast Regional Stock Assessment Review Committee (SARC-49) meeting. The Review Committee was provided with internet access to stock assessment reports and background material prior to the meeting. Prior to participating in the 49th Northeast regional Stock Assessment Review Committee meeting from 30th Nov. – 3rd Dec 2009, I read and carried out a preliminary review of documentation provided. The review panel was chaired by Prof. R. Latour and all reviewers contributed with regards to both assessments.

This report represents my personal review of the two stock assessments considered at SARC 49, but has much in common with the SARC summary report as there was generally a high level of consensus between reviewers. I have therefore tried to summarise most of the issues raised and comments made in the summary report (such that this report can stand alone), whilst attempting to focus on the aspects I personally felt most important and avoid excessive repetition. As required by the CIE statement of work, additional documentation including: a Bibliography of review materials (Appendix 1), a copy of the CIE Statement of Work (Appendix 2) and the Panel membership (Appendix 3) are provided as appendices.

2. Review activities

The Review Committee convened at the Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from November 30 to December 3rd, 2009. The Committee comprised a chair and three panel members. Plenary sessions were open to the public at the meeting and via Webex and conference call.

Monday, 30 November 2009: A presentation of the Stock Assessment Workshop (SAW) results for Atlantic surfclam was given by assessors from the working group (Larry Jacobson, Toni Chute). This was followed a discussion of the material presented. The review panel outlined a number of relatively minor points for clarification and expressed that they were generally content that the assessment met the terms of reference.

Tuesday, 1 December 2009: Assessors from the SAW (Tim Miller & Jason Link) presented the data and results for the butterfish assessment. During discussion of the butterfish assessment the panel noted that the assessment was much more uncertain and more clarification was sought regarding background and potential bounds for the assessment outputs (F and biomass). Paul Rago outlined an 'envelope' approach for exploring plausible bounds for biomass and fishing mortality, which was considered very useful and the reviewers requested that he elaborate on this in the subsequent session.

Wednesday, 2 December 2009: In order to allow additional computational time for butterfish, the agenda was altered, with a short discussion of surfclam, followed by editing of the surfclam assessment summary, before additional information on butterfish was presented. There was lengthy discussion regarding the adequacy of the butterfish assessment reference point and stock status determination, including a presentation on the envelope approach, which indicated that the current stock assessment was within reasonable bounds. There was also discussion on how to proceed in the event that the review panel could not endorse the assessment and/or reference points. This was followed by a private meeting of the review panel to determine where there was consensus on individual terms of reference, identify the contentious issues and determine a way forward. The plenary resumed with review panel members highlighting the issues of concern, further presentation by the assessors as well as procedural guidance (via Webex and teleconference) from members of the science and statistics committee. A way forward was agreed involving not accepting the reference point determination and agreeing to a time table for the final session where the review panel would subsequently work privately on the summary report while the assessors worked to prepare a decision table for presentation at the final plenary, during which the assessment summary would also be edited.

Thursday, 3 December 2009: The review panel completed the summary report and the assessors presented the decision table for short term projections. This included scenarios based on long and short term historic recruitment. It was agreed that the SARC would meet subsequently (15 December 2009) using Webex to finish the review of the butterfish assessment and finalise the assessment summary document. The meeting was then closed.

Review panel members were required to prepare an independent report indicating, for each Term of Reference of the relevant SAW, whether or not: i) the ToR was completed successfully and ii) whether the work presented provides a scientifically credible basis for developing fishery management advice. In making these judgments, reviewers took account of the data adequacy and usage, the appropriateness and implementation of analyses and models applied and the interpretation and conclusions drawn.

The SARC chair and panel members prepared a first draft of a report summarising a consensus of their reviews of the assessments. Individual reviewers subsequently prepared their independent reports following the meeting. There were few disagreements between the panel members regarding most issues, and therefore my independent review should largely reflect the SARC summary review report developed at and following the meeting, but with focus on the issues I personally considered in more depth.

The SARC group also reviewed aspects of the assessment summary report for each stock during plenary and subsequently by email and Webex meeting.

3. Review findings – Stock assessment reviews

3a. Atlantic surfclam in the US EEZ

ToR 1. Characterize the commercial catch including landings, effort, LPUE and discards. Describe the uncertainty in these sources of data.

The commercial catch was generally well characterised in terms of landings, discards and an estimate for incidental mortality. Data presented included time series and spatial plots.

Landings were measured volumetrically in bushels which were very closely monitored under the ITQ system as single species landings and reported. Landings data were considered accurate. A conversion factor for bushels to meat weight is applied. This could introduce some error because clam condition and therefore weight varies seasonally with reproductive status and the same volume of different sized clams is likely to have a slightly different weight. The text (p9) notes EEZ landings as stable between 21-25kt during 1985-2008, but table A2 indicates the lower limit is 18.2kt. Work is required to improve the conversion factor.

Discards had occurred historically during a period when a minimum landing size (MLS) was in force, but SAW and industry members noted that discards no longer occurred because it was illegal to land mixed catches and that grading mixed clams was too time consuming. Surfclams and ocean quahogs were generally distributed at different depths and on different grounds and fishers avoided areas where mixed species occurred. Discard data were available for 1982-93, but were estimated very simply by Invertebrate Subcommittee members with experience of the fishery for the period 1979-81. The discard estimates took into account particular recruitment circumstances in the New Jersey grounds following a die off due to hypoxia, however most landings at this time came from the Delmarva grounds (Table A3). The report text could be clarified in future to better explain the derivation of these discards estimates. The historic discard estimation was simplistic and there seemed to be some spatial inconsistency, but this is likely to be a minor problem and discards are now considered negligible.

An estimate of clams killed by incidental mortality (due to contacting the dredge and sorting gear) was estimated as 12% of landings and this was considered likely to be an over-estimate by members of the SAW. Although this 12% was considered an over-estimate and discards are currently considered zero it would be more consistent to estimate it relative to the total number of clams that contact the dredge (i.e. landings + discards) rather than landings only. Further, selection consists of two processes, selection due to the dredge followed by selection due to the sorting gear. These could introduce differing size structured incidental mortality. A more consistent and if possible data based approach to incidental mortality would be useful, particularly as this is one potential fishery induced effect that could impact (post settlement) recruitment.

Length data were presented by region and sampling levels are provided for numbers of trips and lengths sampled, which indicate a relatively low individual sample size (30 clams), but generally adequate overall sampling levels in Delmarva (DMV) and New Jersey (NJ) and rather low and/or inconsistent sampling levels elsewhere. Age at recruitment is discussed and appears to be based on age sampling carried out during surveys. It appears that there is no age sampling of the commercial catch. Improved sampling levels and coverage for length and an age sampling programme for the commercial catch could be beneficial in improving precision as and when the assessment is implemented in SS3. Consideration could be given to a length stratified sampling programme for age and possibly SLMWT for the commercial catch. Using shells after processing could provide a relatively cheap means to obtain the former only, although this approach may lose information regarding the origin of the catch. Improved sampling (coverage and sample size) of the commercial landings for length and the addition of an age sampling programme for commercial landings could be beneficial.

Nominal and relative price and revenue of landings was also presented.

Fishing effort was also generally well characterised by time series of reported hours fished by region, which are prorated to areas without effort data. Fishing effort data were presented in tabular and graphical form both as time series and spatial plots. Table A4 notes that prior to 1981 effort data were less reliable due to effort restrictions, but the report text indicates that effort data were not reliable for 1985-1990 (for the same reason) and considers them reliable before and after this time period.

LPUE data were also characterised through time and space, and showed trends that were relatively consistent with surveys and assessment output presented later. During discussion of the spatial plots of LPUE it was suggested that some of the high values near the shelf edge might be the result of species mis-classification in the database, but in general the spatial distribution of landings, effort and LPUE characterised the distribution of the fishery and its spatio-temporal trends. Previous assessments had used GLM standardised LPUE, but found little difference between this and nominal aggregate LPUE, the latter was presented in this report.

Trends in landings, effort and LPUE for ‘important’ ten minute squares. These highlight some general trends but also the variability in exploitation history at this scale. They are difficult to interpret generally, because of the relatively large number of squares considered.

Landings and effort are considered accurate, with some exceptions and discards currently negligible. Sampling levels are presented for length data, but the aggregation protocol (for raising to annual by region was not described) and there is little other consideration or description of uncertainty in the commercial data. As noted above improved age and length sampling of the commercial catch could benefit the proposed SS3 assessment and some estimates of dispersion for lengths and age samples could be presented.

The stock definition (i.e. EEZ waters) precludes the consideration and use of state catches in the assessment, although these are likely to be part of the same biological stock structure. This was discussed in the panel report. Inclusion of data regarding surfclam populations in state waters would make sense from a biological perspective.

ToR 2. Characterize the survey data that are being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, age-length data, etc.). Describe the uncertainty in these sources of data.

Survey data were based around the triennial NEFSC stratified clam surveys and were described widely. The strata are based around depth and substrate types and organised regionally. Uncertainty was characterised through description of sampling levels CVs and confidence intervals. Information relating to additional state run surveys was also provided in support of the NEFSC data and SAW conclusions.

The report briefly described the protocols for 'borrowing' data to 'fill holes' and notes that despite research recommendations a model based approach has not yet been developed although it appeared practical for ocean quahogs (please correct text which says Atlantic surfclams). A model based approach would be preferable to the current rather ad hoc approach which must introduce a degree of smoothing through time. Such model based approaches could be based on the use of GLMs in either time or space or possibly more complex spatial modelling techniques based upon the stratum means. Spatial modelling within strata would be complicated because the random design means each sampled year would have different stations, but could be illuminating with regards to the efficiency of the stratified sampling design. Presumably some consideration of spatial heterogeneity has been made previously, that lead to the decision to fill holes by borrowing from neighbours in time rather than space. A model based approach to 'filling' data holes would be preferable.

Randomly selected stations that are too rocky or rough to tow (static gear was also mentioned as a reason for not towing) trigger a search for suitable ground in the near vicinity. Non-towable stations are used to estimate the proportion of ground that is not suitable for surfclams in a stratum and region. Care needs to be taken that near vicinity search does not bias the estimate of unsuitable ground by over-estimating the area that can be fished. Presumably acoustics are used to determine ground suitability for dredge operation and their use could be extended to a wider scale and or some clear and consistent protocol based on the proportion of towable ground in the vicinity could be developed for deciding whether or not a station is towable. Table A24 gives details of the swept area raising in which it appears unfishable ground is considered as zero for all regions other than GBK. A clear protocol for designating untowable stations and rationale/methodology for accounting for these in estimating ground area suitable surfclams is important, especially as a substantial part of the stock is now in an area where ground type may be more variable.

Survey coverage has been good on DMV, NJ and Long Island (LI) grounds, is poorer for Southern New England (SNE) and has frequent gaps on Georges Bank (GBK) where no survey was carried out in 2005. As around half the fishable biomass is currently estimated to be on GBK, it is important that this area is surveyed reliably. South Virginia (SVA) has been relatively poorly sampled in recent years as surfclam abundance and the fishery have declined in this region. Ensure survey coverage is adequate with regards to importance of area to the stock and fishery.

Survey data are aggregated by length to represent pre-recruitment (50-119mm) and fishable (120mm+) abundance and biomass and these data are presented as time series on a regional and combined basis, along with CVs and other information relating to sampling levels (Table A9). CVs were lowest in NJ followed by DMV were often high in LI and SVA high and moderate in GBK and SNE. CVs for pre-recruits tended to be higher than for adults, probably reflecting reduced sampling efficiency for smaller clams, but also potentially relating to higher natural variability of clam abundance at this size. CIs were variable in time and by region, but general trends were consistent with those for commercial catch rate data and suggested declining recruitment and biomass in more southerly regions. Although pump voltage problems were identified during the 1994 survey, generally data from this survey did not appear to be outlying and were consistent with overall and regional trends.

Cooperative depletion studies have been carried out using industrial dredgers to estimate patch densities and efficiencies for the survey dredge. This work also takes advantage of the patch model, a methodological development for accounting for spatial structure in depletion studies. The experiments have also included work to estimate the survey dredge selectivity as well as that for commercial dredger fishing in survey mode and repeat tow experiments to look at the effects of variations in gear setup (new cables and pumps). These data have been built up over a long number of years and represent a considerable effort to improve the efficiency corrected swept area biomass estimates produced from the survey.

The SAW text noted that with one exception (depth) there was no clear correlation between efficiency estimates and most other variables , but the apparent negative relationship between commercial dredge efficiency and density was noteworthy (p.15). However, Figure A.41 indicates correlation between commercial and survey catch rates and a negative relationship between patch density and efficiency is also apparent with the median (as used by the SAW) and is likely to overestimate efficiency at high density and underestimate it at low density. This would lead to the reverse bias in biomass estimates, which would be underestimated for high densities and overestimated for low density sites. Fortunately, there is a cancelling effect in these biases but some investigation of the sensitivity would be useful. Adjustment for this relationship may be difficult due to the circularity between the density estimate derived from the survey and the efficiency used to derive that density. Nonetheless, it might be possible to adjust these using an iterative algorithm or use a simple binning approach to adjust on a very broad scale. The data suggested that efficiency was variable at low density and low, but more consistent at high density. This suggests spatial structuring (clumping) of surfclam populations within a given patch, which combined with the relatively low coverage of the survey tow, mean that at low density the tow may hit (or miss) a higher density clump, thereby giving a high (or low) density estimate. At high density the dredge will have high probability of encountering at least one ‘clump’ of clams thus giving a more consistent estimate of efficiency.

It is also apparent from the plots of the depletion experiments (figs A35-A38) that the area depleted generally only covers part of the setup tow. This therefore assumes that densities are consistent over the area of the setup tow that lies out with the depletion area. Expanding the area depleted or reducing the extent of the setup tow would seem to be options to improve the experimental consistency.

Stepwise AIC was used to analyse repeat tow experiments and suggested that for survey vessel repeat tows differences between pairs of tows were greater than differences due to pump or cable effects. Repeat tow experiments between the RV and fishing vessel were confounded with respect to cable and pump effects and based on results for ocean quahogs the latter were ignored. Graphical analysis indicated substantial variability in the data and a linear model indicated significant station effects, but cable effect was not significant. It appears that between tow variation is greater than systematic cable or pump effects in these experiments.

The selectivity experiments added new value to the surveys by allowing adjustment of catch rates (and subsequent biomass estimates) for gear selectivity. Selectivity was also estimated for a commercial dredger fishing in survey mode, however the chart is truncated (Fig A.47, right-hand column) omitting some length classes with high catches that may fit slightly less well.

<u>The experiments to estimate gear efficiency investigated a range of factors (efficiency, selectivity, gear setup variation) and add greatly to the credibility of the assessment. However, they also highlight variability and uncertainty.</u>

Shell length meat weight relationships (please standardise acronym - SLMW, SHMW, SLMT & SLMWT all appear in the text) were updated retrospectively to use fresh weights obtained during surveys. This represents an improvement but as noted above (under commercial catches) this relationship may vary with latitude, depth and seasonally and some background work to examine these effects would be useful to identify if adjustment for these factors is necessary. These relationships were appropriately combined over regions by weighting by survey densities. Resultant parameters suggest a systematic decline in both the a and b parameters through time. <u>Additional work to investigate SLMWT would be beneficial.</u>

Age samples are taken during the survey and the report presented work that had been carried out that validated the ageing process (Appendix A6). This was an important advance that improves the assessment credibility and will be more important as the assessment is moved to the age and length structured SS3 model. Age sampling levels were presented, showing good sampling coverage and levels in DMV and NJ, but more sparse coverage and levels elsewhere. Attempts should be made to improve sampling levels in the less well sampled areas in the future.

Survey age distributions were presented as mean numbers at age per tow and recognisable and strong year classes were identified qualitatively. This is important as it identifies that recruitment events appear to be different in different regions, suggesting the stock identity assumptions used for this stock may be questionable.

Age information from the surveys was also used to estimate growth rates by year and region and indicated what appear to be systematic changes in growth rates in some regions. As always correlation between von Bertalanffy parameters is apparent, but SEs suggest reasonable fits. The method of combining growth curves over years and regions by averaging predicted points, then fitting a new curve seems appropriate, although it is not always clear whether points relating to annual or regional curves were weighted (e.g. according to abundance) during the averaging process. Despite reductions in growth rate occurring in some regions, the Schnute growth parameter (J) shows a slight increase through time, reflecting the shift in the population distribution to the north where higher growth rates are achieved.

<u>Improved coverage for age sampling in the less well sampled areas could be beneficial.</u>

Useful additional survey information was available from New York and New Jersey state surveys and was presented in Appendix A3. Both surveys use a commercial dredger operating in a random stratified survey design. Their results are generally supportive of the NEFSC survey results suggesting reductions in recruitment and stock levels, and an increase in mean length in the population as the population becomes composed of older individuals.

The New Jersey survey also takes grab samples which provide additional information regarding juvenile clams too small to be selected by the state or NEFSC dredge surveys. These grab sample data are used to provide an indication of the success of spat settlement and suggest that the reduced recruitment (of spat) occurring in NJ may not be impaired, and that factors reducing recruitment (to the fishery) may be operating after settlement.

<u>Inclusion of data from state waters and a more rationale stock definition for modelling could improve the understanding and assessment of this stock. A programme to collect data indicating year class strength soon after settlement would be very valuable.</u>

ToR 3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and characterize the uncertainty of those estimates.

Stock biomass was estimated using efficiency corrected survey data and fishing mortality was estimated using the catch/biomass ratio. These metrics were presented probabilistically at a regional level and for the stock as a whole, which permits straightforward interpretation against reference points.

The KLAMZ assessment model fitted to scaled survey recruitment and fishable stock biomass indices and with a prior distribution on the scaling parameter and a random walk model for recruitment was also used to model the population dynamics of the stock. The model fitted as expected and had the effect of producing smoothed outputs for the metrics of interest (R, B, F). Although this modelling approach takes population dynamics into account and was independent from the direct swept area estimates, it was constrained by the random walk model for recruitment and as it used the same data (catches and survey data) and was scaled to approximate the survey biomass estimates, it was unsurprising that results were consistent with the swept area estimates. CVs calculated by bootstrap and delta method provided alternative indications of uncertainty/dispersion in the biomass, recruitment and fishing mortality and using both methods were generally less than 20%. The generally opposing trend in CVs calculated by the different methods is worthy of note. Bootstrap CVs were generally, and in more recent time lower and were considered less reliable due to frequent problems with model convergence. Biomass and fishing mortality estimated from KLAMZ were again presented probabilistically, facilitating straightforward evaluation of status against reference points.

The extensive work to estimate efficiencies and selectivity for the surveys increased the credibility of both the swept area and KLAMZ estimates of B and F, but also introduced new uncertainties as discussed in the panel report and under surveys herein.

Spawning stock biomass was not estimated as surfclams mature at ages below those surveyed or modelled, sexes are separate and there is little or no information available on sex ratio. It was not clear how much background information on life-cycle was available. Issues of sexual dimorphism or changes in sex with age, size or density would all complicate the population dynamics and some background information to assess whether sexes need to be considered explicitly could be valuable.

Retrospective assessments were presented for both swept area (Tables A28-A29) and KLAMZ model estimates of stock size and fishing mortality. Historical swept area estimates are reasonably consistent in scale with no obvious systematic effects and widest variation in 2002 estimates. KLAMZ retrospectives suggest a rescaling of the assessment as data are removed/added with some tendency to over-estimate biomass and underestimate F. Historical retrospectives for the KLAMZ model are reasonably consistent in scale and trend and do not indicate obvious systematic trends.

Spatial variations in pre-recruitment through time were described by a time series of spatial plots indicating survey catches of clams <60mm in length.

The assessments as carried out by swept area and in KLAMZ fulfil this ToR, by providing biomass and fishing mortality estimates in a probabilistic manner that permits straightforward evaluation against reference points. However, moving to a more explicitly age and length structured approach should provide benefits in terms of modelling the stock dynamics in more detail and in particular with regards to improved estimation of recruitment. This may in turn help to provide new insights into current recruitment declines. The proposed move to SS3 would also permit spatially structured analysis, which will be a pre-requisite if the triennial survey is replaced by an annual rolling survey, covering different areas each year.

ToR 4. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}; and estimates of their uncertainty). Comment on the scientific adequacy of existing and redefined BRPs.

The rationale for reference points in this assessment was not changed. The F_{MSY} proxy reference point remained the same (F_{MSY}=M=0.15), and the biomass reference point was updated with the new biomass estimate for 1999 (B_{1999}=1086kt meats). This pragmatic approach seems acceptable given the limited stock and recruitment dynamics implemented in the KLAMZ model.

However SAW comments that the reference points could usefully be re-evaluated in the future with respect to the following points are valid and constructive:

1) M=0.15 may overstate the productivity of surfclams. It was also noted that simulation work has indicated it may not be possible to achieve MSY in a spatially structured stock and this needs to be considered in setting the F reference point. However the management plan results in F targets considerably lower than the FMSY reference.

2) The existing biomass target can be achieved by the biomass on Georges Bank alone and hence does not provide protection for the stock elsewhere.
3) The rationale that B1999 represents B_{virgin} is rather *ad hoc* and could potentially be improved by exploring the population dynamics more extensively once the SS3 age/length structured modelling approach is implemented.

A better biological basis and consistent approach for setting the reference points would be preferable.

The sedentary nature of many shellfish stocks makes definition of stock identity, and assessment and management on a broad scale, difficult as many biological and fishery processes may vary more locally. In this assessment there is evidence for differing growth rates and condition factors regionally as well as marked differences in exploitation rate and history. These do not constitute grounds to change the stock definition, but do complicate assessment and management. However, there is also evidence for variable recruitment by regions, with general recruitment declines more apparent in some regions and historical differences in the timing of strong year classes. These differences in recruitment dynamics do suggest that the single stock identification may be questionable and discussion during the meeting suggested that there may be physical barriers to larval transport (e.g. Hudson Canyon) from one region of the stock to another. There is therefore a need to review available information on hydrodynamics and model the potential for larval dispersion under different scenarios of stock biomass. This seems particularly important given the current situation where around half of the biomass is concentrated in one region that may be isolated, in terms of providing larval supply, from the main area of the fishery. An adequate definition of the stock identity, taking account of recruitment dynamics, is required in order to define biologically meaningful reference points. Explicitly accounting for heterogeneous spatial structure in sedentary stocks and their fisheries would also be beneficial, but it may not be not straightforward to find a suitable scale at which this can be achieved.

ToR 5. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 4).

Results from the assessments indicate that the stock was not overfished and that overfishing was not occurring. However, although overfishing was not occurring at the stock level, fishing mortality varied regionally and there was small probability (circa 20%) that F in Delmarva exceeded F_{MSY}=0.15 (Fig A58).

Comments under ToR 4 relate to the adequacy of the stock identity for setting meaningful reference points. The scale of management is also important in determining whether and how fishing can occur within these limits. Under the present regime there is differential exploitation within the overall management unit, which together with environmental factors is leading to a mismatch between resource and exploitation distributions. A redistribution of fishing effort in response to the opening of GBK may alleviate this problem, but if it does then additional management measures might need to control the distribution of fishing effort.

ToR 6. Identify potential environmental, ecological, and fishing-related factors that could be responsible for low recruitment.

Although no definitive conclusions could be reached the SAW report considered a number of data sets and potential factors that could be responsible for low recruitment including:

1) Reduced growth rates, which delay recruitment to the fishery allowing for lower survival due to natural mortality during the time delay. Natural mortality is usually related to size and as such slower growth rates will mean clams may be subject to higher natural mortality throughout their lifetime.
2) Warm water temperatures were thought responsible for die off of surfclams in DMV in the early 2000s and would likely reduce recruitment. Some members of the SAW felt the distribution of surfclams was moving to deeper (cooler) water in response to warming temperatures (climate warming), and that growth and condition could suffer under these conditions, but no strong evidence was presented for this. Temperature anomalies for the DMV area supported the high temperatures experienced in the early 2000s, but suggested that the temperature regime had now returned to more normal levels. Further scallop populations nearby would be expected to be more sensitive to warming temperatures, but had not shown recent systematic reductions in recruitment. However, a wide scale and gradual temperature related effect related to climate change remains a possibility.
3) Predator data for surfclam are very limited, but indicate elasmobranchs may be important predators. Data on abundance of potential predators from bottom trawl survey data was subsequently presented. In discussion, Industry representatives felt that rays in particular might have increased in recent years and be important, while scientists suggested that ray abundance over the survey (=stock) area was limited and had not increased sufficiently to be responsible for the recruitment decline. It was also suggested that crabs might be responsible, but bottom trawl survey data presented during the meeting did not suggest strong increases in abundance for most of the likely species. However it should be noted that trawls may have poor capture efficiency for many invertebrate species, including crabs and particularly those that burrow frequently and effectively such as cancrids.

4) The SAW considered a range of fishing effects.
    a) Reduced biomass was considered unlikely to be responsible for recruitment declines to date, although it may contribute in future. Declines in recruitment began to appear while stock biomass was still at a relatively high percentage of the maximum estimated level. NJ state survey grab samples did not suggest that spat settlement was impaired, although data were very noisy in recent years.
    b) Disturbance of sediments by dredges may have caused problems. However, although fishing effort has increased, the proportion of the total stock area disturbed by dredging is relatively small. The report notes that the distribution of clams is patchy and fishery activity is focussed on high density grounds. Simple comparison of area impacted may therefore underestimate the effect on the population. Survey and commercial data indicate that recruitment and fishing occur in the same areas, but that not all areas with good recruitment are fished.
    c) The larval duration of surfclams (19-35 days) is sufficient to allow for dispersal to other areas. However, no information on current regimes was presented and in

discussion members of the group suggested that physical barriers to larval transport may exist.

Mean bottom salinity anomaly presented in the report suggested that low salinity may have persisted in recent years, but this was not discussed.

The possibility that a minimum density is required for successful fertilisation was mentioned in discussion, but lack of evidence for decline in spat settlement does not support this. Discussions also considered the potential for reduced productivity in recent years and subsequently time series of surface chlorophyll concentrations were presented as an indication of productivity trends. Other members of the review panel suggested potential reductions in nutrient levels due to declines in the dumping of sewage sludge, however others felt that that the effect was too widespread to be attributed to this and that the nutrient balance for this area was primarily under the influence of upwelling. Parasitic infections were also considered in discussion, but it was noted that during previous mortality events, parasite loadings had been investigated and surfclams were found to have relatively few parasites.

The SAW identified and considered a range of potential factors that could cause recruitment decline.

ToR 7. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).

a. **Provide numerical short-term projections (1-5 years; through 2015). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment.**
b. **Comment on which projections seem most realistic, taking into consideration uncertainties in the assessment.**
c. **Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.**

The SAW used the KLAMZ model to provide projections under a number of states of nature that were pre-assigned broad probabilities and under four different management scenarios. Projections were bootstrapped (based on 2008 estimated stock conditions) to provide a probabilistic framework for evaluation. These simulations provide a means to address the ToR in a way that is consistent with the stock assessment.

Whole stock projections indicated the stock was likely to continue to decline gradually in the short to medium term under the management plan scenarios, but would decline much more rapidly if fished at the F_{MSY} proxy. The probabilities of falling below the B_{MSY} proxy and the $B_{Threshold}$ in 2015 were presented along with the probability of overfishing (exceeding the F_{MSY} proxy). The simulations suggest that with the exception of fishing at the F_{MSY} proxy (which falls outside the management plan) the probability of overfishing or of the stock falling below the threshold biomass is negligible. They suggest that given the current management criteria and stock definition the stock is relatively robust to overfishing in the short to medium term and address ToR 7c.

However, note that there are regional variations in stock condition which could give rise for concern and the SAW carried out projections for DMV and NJ separately, where the level of data can effectively support the stock assessment model at this scale. Three harvest strategies were considered for these areas; constant catch (2003-2008 level), status quo F and zero F. These scenarios covered a useful range for informing management and indicated that biomass was likely to continue to decline in both areas under most scenarios, the exception being NJ at F=0.

The decision table provides a pragmatic way to present and consider a range of assumptions (ToR 7a) and to indicate which are most realistic (ToR 7b), although the probabilities assigned to states of nature are qualitative and to some extent subjective. The bootstrap provides a means to take account of uncertainties in the assessment and provide probabilistic outputs (ToR 7a). The report does not appear to present the annual probabilities for intermediate years.

The KLAMZ model provides an internally consistent approach to projection with probabilistic output that fulfils this ToR. Under the current stock definition the stock does not appear vulnerable to overfishing, but there are concerns locally.

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

The SAW addressed each of 18 research recommendations (rr) individually and briefly. Some progress was made towards 12 of these with no progress on 6.

The transition to SS3 for assessment was considered in considerable detail and a trial run presented to the SARC and as an Appendix (A5) in the SAW report (rr xvi). Changing to this methodology is considered a pre-requisite given the likely change to a rolling cycle of regional surveys in the future (rr vi). The method also presents considerable opportunity to benefit from a more explicit age structured approach as was illustrated by the persistence of signal from strong year classes in the pilot study and this may provide the opportunity to improve understanding of recruitment dynamics (rr xv). In this regard it may also be advantageous to consider collecting age data from the commercial fishery as well as from the survey (rr iii and ToR 1 herein). SS3 will improve the utility of commercial length data (rrs iii & x), but overall sampling levels are currently quite low in some areas and individual sample size could also be increased (ToR 1 herein). No progress was indicated for conversion factors for the commercial landings (rr iv) and these and SLMWT are also considered under ToR 1 herein.

The transition to SS3 could address or force consideration of many of the research recommendations; 6 are noted above, but once fully implemented the new approach should provide more precise information regarding recruitment trends and this may also help to establish more effective reference points (rr xvii and ToR 4 herein).

Simulations to test the effectiveness of the patch model to heterogeneous stock structures have been carried out and will be taken further. This addresses rrs viii, xii and possibly v.

Fieldwork to estimate efficiency and selectivity for the survey dredge and a commercial dredger addresses rr xiii and possibly in part rr xi.

The report utilised data from state surveys, thereby addressing rr xviii.

Other rrs, where no progress was made include:
i) relating to natural mortality
iv) relating to commercial landings to meat weights conversions. Issues relating to this and SLMWT are briefly discussed under ToR 1 herein.
vii) suggesting new technological survey methods but none specifically. It was felt by the review panel that some means of sampling younger/smaller clams (and potential benthic predators) would be useful. A small meshed beam trawl was suggested by one reviewer. Quantification of catches to provide abundance indices from such a sampling programme would be beneficial.
ix) relating to refining logbook data and improve precision. This may require management action, but is achievable given the precision of modern electronic geolocation devices used on most fishing vessels.
xiv) relating to a model based method for ‘filling holes’ in survey data. This was discussed and preliminary results obtained using GLM models for the ocean quahog survey were encouraging.

Eight new research recommendations were suggested which are all worthy of attention:
Three (ii, v & vi) relate to improved port sampling which should if possible be followed up along with improved commercial age sampling if sufficient resources (for ageing) are available.
Two(i & viii) relate to research into recruitment processes generally and the density of clams required for successful spawning in particular. This was noted under ToR 6, but limited data for spat settlement suggest the recruitment decline may occur post settlement.
The remainder are:
iv) information on maturity and fecundity at length/age, which could be broadened to examine sex ratios in the population and how this relates to adequate spawning densities (see previous).
iii) determination of suitable ground area for surfclam on GBK. This assumes greater importance as the proportion of the stock estimated to be on GBK increases. Errors in the area assumed suitable for the stock on GBK could have a significant influence on the biomass estimated for this component of the stock.
vii) determine whether carrying capacity for surfclams has changed over time. Although relevant, this appears a very wide ranging research aim without proposed methodology. In this meeting the SAW and SARC have briefly considered a number of issues such as primary productivity and climate change (temperature) effects that relate to this issue.

In addition to these, a review of available data on current regimes and simulations to evaluate the potential for larval dispersal from different areas of stock distribution could be very informative with regards to stock identity, potential metapopulation structure and recruitment dynamics.

<u>The SAW has considered a wide range of research recommendations, addressed many and highlighted areas for future work.</u>

3b. Butterfish

ToR 1. Characterize the commercial catch including landings, effort and discards by fishery (i.e., *Loligo* fishery vs other fisheries). Characterize recreational landings. Describe the uncertainty in these sources of data. Evaluate the precision of the bycatch data with respect to achieving temporal management objectives throughout the year.

Largely completed: The SAW presented data available on landings and revised historical estimates of discards. The consistent approach applied to discard estimation should represent an improvement to the historic data used in the assessment. Effort data were not available because the butterfish fishery is not directed, most butterfish being taken as bycatch in trawl fisheries targeting other species (e.g. squid, silver hake). The SAW presented *Loligo* landings as a proxy for butterfish effort, as most of the butterfish bycatch is taken in the *Loligo* fishery. There were major uncertainties in historical data relating to non-reporting of butterfish landings by foreign fleets historically and estimation of the quantities of butterfish bycatch discarded. There was some discussion during the meeting regarding temporal management objectives (for the *Loligo* fishery) and the use of data from the previous year together with 'near real time' estimates of butterfish bycatch. Current sampling levels for butterfish are modest and in year temporal management seems likely to require higher sampling levels and an effective 'near real time' reporting mechanism. Also the population dynamics of butterfish may result in high variability in year class strength from year to year (e.g. Fig B28), which with age 0 recruitment could make the previous year's data (particularly relating to discards) very different from the current year.

The assessment report presented time series data on landings from the fishery including historic declared landings of butterfish by Japanese fishing fleets fishing for *Loligo pealeii*. However, Spain and Italy did not report the butterfish bycatch taken by their squid fishing fleets during 1970-1976 and there was no observer coverage until 1977. Landings are therefore likely to be underestimated during this period. Japanese market demand and landings of butterfish have decreased and since 2001 there has been no directed fishery and relatively low landings (c.500t).

Historic discard data (1976-1986) were considered representative of the directed butterfish fishery and likely to underestimate discards of bycatch taken in other fisheries. The SAW used the standardized bycatch reporting methodology (Wigley et al., 2006) to produce estimates of by-catch from 1989 to 2000 and revised estimates of discards for 1965-1988 based on an average ratio (for 1989-1999) for bottom trawls using mesh <4 inches and landings of all species taken by this gear type and on a regional spatial scale. <u>This standardised approach is welcomed, but the precision and uncertainty in these estimates is high (CVs frequently around 1.5) and the approach requires a number of assumptions (e.g. mesh size, all species, etc). Since 1989 discard data have been available and discards estimates generally have lower CVs (0.2-1.7) although the report notes that their precision is also generally poor.</u>

Discard estimates for butterfish based on *Loligo* catches were also presented (although not used for assessment). These seem to be higher (than the all spp. estimates) in the most recent years but are lower further back in time. Also in general, when the estimates of butterfish discards are high the all spp. estimate is higher and when the estimates are relatively low the squid based estimates are higher. CVs for squid based estimates were extremely variable. Differences between the estimates produced by different raising methods further highlight the uncertainty in these data. Sampling levels for discard estimation were not explicitly presented in the report, but length compositions from the observer programme indicate around 1000 to 4000 butterfish were measured in most years with more than this in a few years.

Recreational catch was investigated and found to be insignificant.

Annual length distributions for the commercial catch were presented, but no description of aggregation protocol was given although recent years take account of market categories. Quarterly length sampling levels were generally low (<5 samples per quarter) to moderate (5-15 samples per quarter) and have improved in recent years, while average sample size (circa 100) appears adequate.

No effort data have been presented, because there is no directed fishery for butterfish so effort is difficult to estimate. At the request of the reviewers, *Loligo* landings were presented as a proxy for effort in the *Loligo* fishery which is where the main catch of butterfish occurs. These indicated *Loligo* landings (≈ butterfish effort) had fluctuated around the same level since 1978.

Charts of observed commercial trawl catches (landings and discards) were presented (during the meeting) of survey catches to providing information relating to seasonal and spatial trends in the distribution of the butterfish fishery and observer coverage. As these related only to observed trips, coverage of the fleet was limited, however they were considered a useful additional piece of information.

In summary, the SAW presented available data on landings and produced revised and likely improved estimates of discards. There are uncertainties in some historic landings and major uncertainties in discard estimates. Effort data are not available, but a proxy suggests it may have remained relatively stable while butterfish catches have declined substantially. Sampling levels are low to moderate. Uncertainty in most commercial data is high. This, combined with high recruitment variability and a requirement for near real time reporting, suggests that in year management of squid fisheries based on butterfish bycatch could prove difficult.

ToR 2. Characterize the survey data that are being used in the assessment (e.g., indices of abundance including RV Bigelow data, NEAMAP and state surveys, age-length data, etc.). Describe the uncertainty in these sources of data.

Completed: The SAW completed this ToR as stated, but the high uncertainty in the data and conflicting survey trends mean that their adequacy for determining stock trends is questionable.

The SAW described survey data available including NEFSC bottom trawl surveys and a number of state run surveys. Only data for Massachusetts and Connecticut state surveys were available within the timeframe for the SAW and presented in the report. Although data from Virginia (VIMS) are only available as abundance, rather than biomass indices (biomass is needed for the KLAMZ stock assessment model), these might be able to add support to the other data series and in future should be presented in the report if possible.

NEFSC biomass indices were available for winter, spring and fall. Butterfish recruit during the fall at age zero and this cohort becomes classified as age 1 after January 1st, so no fish less than age 1 are present in the winter and spring surveys. Charts of survey catches were presented during the meeting providing information on spatial coverage of the surveys. The spatial coverage of surveys in relation to the stock distribution, which varies seasonally, may cause problems in relation to survey catchability and improved knowledge in this area could benefit future assessments.

Abundance indices for the fall survey were much higher than those of the winter and spring surveys, due to the presence of high numbers of age 0 in the autumn, although this is less apparent in recent years. Historically the indices differ widely in magnitude and it is difficult to compare them, but they are more comparable in recent years. Given the timing of surveys and assuming the abundance indices are heavily influenced by the recruiting age class the winter and spring indices should correspond to the previous fall index in fig. B10. There is some evidence for this with stronger y/cs in fall surveys in 1999 (2000 winter), 2003 (2004 winter & spring) and 2006 (2007 winter & spring) and weak y/cs in fall surveys 2001 & 2 (2002 & 3 winter & spring) and 2004 (2005 winter & spring). This suggests that the different surveys are able to detect relative year class strengths with some consistency.

However, the longer term time trends in biomass are less consistent between the surveys with the fall biomass index showing a large drop in the late 1990s, possibly another around 2001 followed by a level but variable period. The spring survey suggests a gradual increase in biomass since the early 1990s and the winter survey is noisy without trend. CVs for the NEFSC surveys are reasonable for the fall survey (circa 0.2), in the range of 0.2-0.4 for the winter survey and noisy (0.2-0.9) for the spring survey. The SAW report also notes that different gear configurations have been used in the past for the spring survey. The NEFSC fall survey has the widest coverage and also catches butterfish over a much wider area (both inshore and offshore) than the spring survey and is considered most likely to correspond with the distribution of the butterfish stock. The fall survey (downward biomass trend) was therefore considered more precise and appropriate in determining trends in butterfish biomass than the spring survey (upward biomass trend).

Ageing has not been validated and SAW members commented that ageing was neither straightforward nor especially difficult (i.e. it was intermediate difficulty). Age compositions were presented for the fall and spring surveys. There was some evidence for separate cohorts in length distributions, albeit relatively rare and relatively poor correspondence between year class strengths in age compositions within surveys.

Low occurrence of older fish in the surveys and lack of year class signals lead to concern that older fish may be leaving the survey area or less available to the survey gear. The low rate of increase in size of older fish in the surveys and comparison with commercial length distributions may also give some support to a reduction in catchability of large fish in the surveys, but this evidence is weak. The low headline height of the bottom trawl used historically for the survey could contribute to lower catchability for the larger (pelagic) butterfish. The trawl used by the new RV has a higher headline height and may be more effective.

Information presented for state surveys indicated that the Massachusetts spring survey had low numbers and biomass per tow and very high CVs (0.8-1.0), while the fall survey was noisy but had higher precision (0.2-0.4), but was not age structured so could not be used in the KLAMZ mode. The Massachusetts fall survey suggested increasing abundance between 1982 and 1998 followed by a step reduction and moderate abundance since 2003. The spring survey suggested a gradual increase in abundance through time (although the scale used in figure B18 obscures this). Biomass trends are similar although the fall survey suggests biomass has increased recently following the step down to very low levels in 2000 and 2001. CVs were relatively stable for the fall survey, but appeared to increase through time with the spring survey, which reduce confidence in the increasing trend suggested by this series. Connecticut state surveys had no information relating to precision. The Connecticut fall biomass index increased from 1992 to 1999 then stepped down in 2000 and has been level of increasing very slightly since then. Abundance for this survey was initially stable after the step down but declined again in the two most recent years. The spring biomass index suggests a very gradually increasing trend over the time series. Both state surveys show more positive biomass trends than abundance trends in recent time suggesting an increase in the size of butterfish in the catch. Neither of these state surveys were age structured, so they could not be used in the KLAMZ assessment model. However, they show reasonably consistent trends with both spring surveys suggesting increasing biomass and both surveys suggesting a large drop in abundance around 2000 after which it stabilises or increases. The NEFSC fall survey may include 2 similar features where biomass steps down to a lower level (circa 1995 & 2001), the latter of these does not quite coincide with the state surveys, being later by one year.

The SAW presented a variety of surveys data along with estimates of their precision. The timing and presence of age structuring limits which surveys can be used in the KLAMZ model with the NEFSC fall survey the only one potentially suitable for providing a recruitment index (age 0). Trends in the surveys differ, with the NEFSC fall survey indicating a decline in biomass since the late 1970s, state run fall surveys indicating a sudden drop in biomass around 2000, but possibly increases before and since this, while spring surveys (NEFSC and state) suggest increasing trends in biomass. Fall surveys had higher catches and lower CVs suggesting they had higher precision than spring surveys, which generally had very high CVs. There were also some concerns relating to the suitability of survey gears for sampling this pelagic species and whether coverage is adequate for a schooling species with seasonal migrations. Coverage by the NEFSC fall survey was considered to better represent butterfish stock distribution. The available survey data thus provide a somewhat inconsistent picture of stock trends.

ToR 3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and characterize the uncertainty of those estimates.

Completed: The SAW completed this ToR as stated, but high uncertainty in catch and survey data, conflicting survey trends, wide variation and inconsistency in possible values for natural mortality, and assumptions required for modelling all introduce uncertainty into the assessment outputs in terms of both scale and trends. However, it is apparent that fishing mortality has declined and is low, particularly in comparison with most likely natural mortality rates.

The KLAMZ (delay difference) model was used to estimate stock size, recruitment and fishing mortality. KLAMZ models the population as (pre-)recruits and a plus group which is considered fully selected. The model was set up with no constraint on recruitment (neither SRR nor random walk were implemented). Survey indices provide information on recruitment and stock biomass trends while landings information is used to estimate fishing mortality. The SAW noted that scaling of the population was difficult in the absence of additional information on butterfish survey catchability.

State surveys were not used in the assessment because of limited coverage and lack of age structure. Four NEFSC survey indices were available and were included in the assessment model, three providing information on 1+ biomass and one relating to age 0+ biomass. Indices were weighted by annual CVs and where disaggregated by age, the survey CV was used for both ages which is likely to be an underestimate. The report notes that the final model has each of the series CVs rescaled so that they all inform the model, but the mechanism for this is not explicit. The magnitude of the survey standardised residuals indicates the model is fitting most closely to the fall age zero index, which is not surprising given no constraining recruitment model. Time series of standardised residuals from the final model indicate substantial systematic trends in the survey fits. These reflect the conflicting trends in the survey indices commented on previously, with winter and spring surveys showing positive trends in residuals in recent years and the fall survey (0 and 1+) indices showing negative trends in recent years. However, the decline in butterfish biomass is also supported by declining commercial butterfish catches, primarily taken in fisheries targeting other species (mainly squid), where landings (as a proxy for effort) have remained relatively stable since the late 1970s. There is therefore some uncertainty in the estimated stock trends.

In order to overcome the problem of scaling the population the SAW introduced a prior probability for survey catchability of 1+ butterfish biomass from the fall survey as this covered the largest proportion of the stock and had lower (combined ages) CVs than other surveys. In order to characterise this, the SAW used information from calibration studies comparing efficiencies for RVs Albatross and Henry B. Bigelow and assumed a beta distribution for the product of the relative efficiency of the RVs, the efficiency of the Henry B. Bigelow and the ratio of survey to stock area. Uniform distributions were used to provide plausible bounds on the efficiency of the Henry B. Bigelow (>0.1 and <0.9) and survey to stock area ratio (>0.5 and <0.9).The beta distribution bounded relative RV efficiency (>0.05 and <1), but mean and variance were estimated from the calibration study and actually had low standard errors. The resulting probability distribution for catchability (q) has most values at the lower end of the range and implied a catchability coefficient of around 0.2 for the Albatross IV. Analyses were carried out to evaluate the sensitivity to small changes in the maxima of the uniform distributions assumed for survey to stock area and Henry B. Bigelow efficiency. As expected these result in changes in the scale of the assessment outputs of a similar scale. The scale of the assessment outputs is dependent on the assumed prior distribution for q and there remains the potential for q to vary between around 0.05 and 0.3 with reasonably high and similar probability (circa 0.1-0.15), which leaves substantial uncertainty in the absolute scaling of biomass and fishing mortality estimates from the model.

Natural mortality (M) was assumed to be 0.8 as in previous assessments and sensitivity analyses on M were carried out for values between 0.6 and 1.0. These indicated changes in scaling that were highest for recruitment and lower for biomass and fishing mortality and made little difference to stock trends. However, this range did not cover the full extent of variation in estimates of natural mortality. ToR 6 explicitly considers natural mortality in the context of predator consumption and this suggests relatively low natural mortalities, but other data and approaches suggest natural mortality may be much higher. For example, catch curves applied to survey data (either by year or year class) suggest total mortality around 1.8-2.0. Estimated fishing mortality would be negligible in comparison, but the suspicion that survey catchability may decline for older fish suggests these may be over estimates. Beverton & Holt (1959) found M/K was usually in the range 1.5-2.5, which given the growth parameters in table B21 implies natural mortality for butterfish in the range 1.0-1.7. Pauly (1980) used a formula including temperature and growth parameters, which, assuming average temperature of 10^{o}C, suggests M = 0.97 for butterfish. Natural mortality rates are poorly defined with different data sets and approaches providing substantially different estimates. There is also the potential for changes in time as predator species compositions have changed. Further work to improve understanding of natural mortality and its level would improve the assessment credibility.

The KLAMZ model incorporates growth using the Schnute model, which is parameterised in terms of the relative difference in weights at the age of recruitment and one year prior to recruitment (J). This causes problems for butterfish for two reasons: they recruit at age zero and size at ages below just under 0.5 becomes negative according to a fitted Schnute growth curve. In order to overcome this, the SAW assumed recruitment at nominal age 1.5 years which permitted estimation of J, commenting that this implies a 7.5 fold increase in weight in the first year of life and that predicted weights at age appeared reasonable, whilst also noting that this may under-estimate growth and productivity and should be reconsidered in future. However, the arbitrary choice of age at recruitment is hugely influential in determining J, since the denominator in its determination is tending towards zero. Moving the timing of recruitment forward or back by very small amounts would greatly change the J parameter. The sensitivity of the stock assessment outputs to this parameter assumption was not tested.

Standard errors associated with the assessment outputs were very high, often similar to the predicted value itself indicating very poor precision. In response to reviewers' concerns regarding the very high levels of uncertainty in the assessment outputs an 'envelope' method, which uses plausible limits for input parameters, was used to provide a pragmatic approach to defining bounds within which biomass and fishing mortality should lie. This was welcomed by reviewers and subsequently presented during the meeting. The 'envelope analysis' provided upper and lower bounds for biomass and F, which enclosed the base model estimates for most of the time series and were in reasonable agreement with the +/- 1 SE CIs for biomass. The KLAMZ model output was outside the 'envelope' in recent time due to the envelope analysis assumption made which used a slightly higher assumption for bounding fishing mortality.

Comparisons with the previous assessment indicated that this assessment resulted in higher and noisier estimates of biomass and recruitment and steeper declines in these. Fishing mortality estimated by this assessment was lower and less variable than in the previous assessment and also showed a downward trend that was not previously apparent. Retrospective analysis applied to this assessment indicated very little evidence of bias in biomass or fishing mortality.

The envelope analysis indicates that the model outputs for this run lie within plausible limits and therefore improves confidence. However, the underlying uncertainties in catch data, survey trends and catchability, natural mortality and growth rate mean that other plausible outcomes may also be possible.

ToR 4. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and F_{MSY}; and estimates of their uncertainty). Comment on the scientific adequacy of existing and redefined BRPs.

Partially completed: The SAW proposed changing the methodology for estimating reference points from the use of the Fox surplus production model to a projection based approach using a bootstrap of all recruitment estimates from the KLAMZ model. The SAW thus proposed using $F_{0.1}$ (=1.04=F20%) as a proxy for F_{MSY} (= threshold) and $F_{30\%}$ (= 0.72) as an F target proxy. They projected the KLAMZ model to derive SSB_{MSY} (=16262t). The methodology proposed by the SAW is internally more consistent than previously and the new reference points were correctly estimated. The SAW also updated reference points using the former rationale and Fox surplus production model. Members of the review panel were generally not happy with adequacy of the reference points due to the very uncertain nature of the data, assumptions and assessment outputs, the use of per recruit analyses (with an annual time step) for estimating reference points for such a short lived and fast growing species and the applicability of equilibrium based population dynamics for a stock that appears to be declining significantly under the influence of factors other than fishing (and is therefore not at equilibrium). Under such conditions it was not clear what the management objective should be (given that stock decline does not appear to be due to fishing – and the stock does not appear to be self-sustaining even in the absence of fishing) and hence what reference points would be required to achieve.

The SAW used deterministic projections to estimate yield and spawner per recruit curves. It would be possible to carry these out stochastically in order to derive distributions for the reference points. Per recruit analyses examine the trade off between growth and fishing and natural components of mortality. As mentioned previously, uncertainty in natural mortality for this species appears to be very high and potential estimates could be in a range from around 0.4 to 1.9. There are problems with the growth model as parameterised in KLAMZ (see comments under ToR 3), although it is not clear if this is the same model applied in the projections. Fishing mortality is estimated to be very small in relation to natural mortality, but its scale is not well determined. In the YPR model it appears that production due to growth is lower than losses due to natural mortality so delaying harvesting by fishing at a lower rate results in yield losses and a much higher fishing rate is required to increase yields by taking fish before they die naturally. The per recruit results used to estimate the reference points seem likely to be very sensitive to the very high uncertainties in growth, M and F.

The assessment results indicate that the stock has declined substantially since the 1970s (from 150kt to 50kt), but apparently primarily due to causes other than fishing as fishing mortality has always been relatively low and has declined from around 0.15 to 0.02. Even despite uncertainties in natural mortality, fishing mortality is low relative to natural mortality and in recent time seems likely to be less than annual variation in M (F <=5% M). Despite this F will have exerted an additional, if relatively small, negative effect. The stock therefore seems to be declining under the influence of factors other than F, which could therefore be considered environmental. If these factors are considered transient then the butterfish stock might return to former levels, but if there is a permanent shift in these environmental factors the butterfish stock might equilibrate to different levels. The current assessment suggests that the stock is not in equilibrium and the use of equilibrium dynamics to set reference points, especially on the basis of fishing mortality rates, does not seem appropriate under these conditions.

If as suggested above the stock is not in equilibrium then the management objectives underpinning the rationale for setting reference points need clarification. Potential aims for the reference points might include:

1) To return the stock to former levels, assuming long term dynamics (i.e. assume the current state is part of a cycle).
2) To halt the decline at the current position (i.e. assume a new equilibrium based around the current stock conditions).
3) To minimise negative impacts on fisheries by continuing at status quo fishing levels (i.e. assuming that the current level of fishing is not the cause of the stock decline and that it is negligible in determining whether and at what level the stock will stabilise).

ToR 5. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 4).

Partially completed: The SAW evaluated stock status against proposed new reference points, updated reference points based on the existing rationale and the existing reference points. However, given uncertainties in the data, assumptions and assessment outputs as well as concerns regarding the rationale and basis for reference point determination it is not clear that the stock status determination is adequate for management. It does appear that fishing mortality is low generally and very low in comparison to likely natural mortality and that the stock has declined, but that this is primarily for reasons other than fishing.

The management framework in the US requires stock status to be framed in terms of whether or not overfishing is occurring and whether or not the stock is overfished. The current assessment suggests the stock biomass is declining, but most likely in response to factors other than fishing. Fishing will exert an additional negative effect but this should be very small given the estimated scale of F relative to M. Fishing is not the primary cause of stock decline and indeed is highly likely to be insignificant at present levels (less than annual variation in M), hence it could be argued that overfishing is not occurring. Stock levels have reduced, possibly to the extent that future recruitment is impaired, but again fishing was/is not the primary cause for this decline and it could also therefore be argued that the stock is depleted but not overfished.

The argument presented above suggests the stock status could be described fulfilling criteria for not being overfished and overfishing not occurring. However, the stock recruitment plot (figure B42) could be interpreted as indicating reduced recruitment at lower SSB levels. The relationship may be spurious in that if SSB is driven by recruitment a few years of reduced recruitment will result in reduced SSB for a short lived species. However one interpretation of the plot would be that no recruitment above around 90kt has occurred at spawning stock biomasses below around 65kt, whilst these have occurred across the range above this SSB level. Simple pragmatic grounds such as these could be used to set a biomass reference point reflecting the stock recruitment relationship. ICES fits two-line (hockey stick) stock recruitment relationships to provide a more objective means of determining potential break points in stock recruitment data as the basis for biomass reference points below which recruitment is impaired. One reviewer suggested the pragmatic approach of using the largest SSB estimate as a proxy for virgin spawning biomass and reducing to a percentage of this e.g. 30%. Given the stock recruit data in figure B42, this approach would provide a very similar estimate in the region of 65kt SSB. Approaches are available for setting biomass reference points, framed in terms other than fishing, and these might be applicable in the current butterfish situation, depending on management objectives under such conditions (see ToR 4).

Data on butterfish maturity presented during the meeting indicated a reduction in the proportion of female butterfish mature at age 1 in 1990 which was not fully reversed, and a further reduction in the last point of the series (2004). There is a slight reduction in size at age over the TS. Over the TS (1984-2004) proportion mature at age 1 drops from circa 0.65 to circa 0.3. Given potentially very high natural mortality rates, which could result the majority of the spawning biomass being aged 1, these reductions could have substantially reduced spawning potential.

ToR 6. Evaluate the magnitude, trends and uncertainty of predator consumptive removals on butterfish and associated predation mortality estimates and, if feasible, incorporate said mortality predation estimates into models of population dynamics.

The SAW completed this ToR, by identifying the main butterfish predators on the basis of stomach contents containing butterfish as >1% diet composition for any five year block. An evacuation rate methodology was then applied to these to estimate daily per capita consumption of butterfish by various predators on a two season basis. These were subsequently scaled to the seasonal total, summed to annual and scaled by estimated predator biomasses and summed over predator species.

Estimated consumption by predators was relatively low (<4kt) prior to 1980, fluctuated between 2 and 12kt during most of the series and recently peaked in 2005 at 16kt, before returning to more normal levels in 2006. Contrasting with landings suggests that landings exceeded consumption by predators before 1980, landings and consumption by predators were similar from 1980 to 1995, after which consumption by predators has been greater than landings. On this basis natural mortality should have been be similar to fishing mortality during the 1980s and early 1990s, at which time estimated F (current assessment) was around 0.1 (0.02-0.14).

The study lists a number of sources of uncertainty, many of which suggest the consumption figures may be underestimates. The review panel had some concerns regarding the effectiveness of the survey in catching a pelagic species such as butterfish and this extended to stomach analysis of predators caught demersally, which might underestimate the occurrence of pelagic prey in stomach contents. Further, the consumption studies did not consider consumption of butterfish by squid, although they did note the high co-occurrence of these two species in time and space and the fishery. Other work noted in discussion (Hunsicker and Essington, 2008) suggested that even if butterfish only formed a relatively small part (5-8%) of their diet, squid could consume quantities of butterfish, similar to estimates of annual recruitment.

Although the report notes that consumption was of the same order of magnitude as landings, this is no longer the case as recent landings are almost an order of magnitude lower. It is also not clear that the landings and consumption have similar trends as suggested. The report appendix provides a number of recommendations for future work many related to dynamically modelling natural mortality using these data. However, at the present time the data on consumption appear rather uncertain and may under-estimate natural mortality by predators. Other indications of mortality such as catch curve estimation of Z (presented during the meeting) and empirical formulae based on growth parameters suggest much higher values for natural mortality than these consumption rates. Whilst it would be beneficial to model M dynamically if sufficient data were available, at the current time the natural mortality rate is one of the major inconsistencies in this assessment and work to rationalise currently differing estimates of its level from different methodologies is required initially. This should extend to considering catchability issues for the surveys including spatial coverage and seasonal variation in butterfish distributions, as well as a range of biological evidence including growth, reproduction (e.g. semalparous lifecycle) and predation, including explicit consideration of squid as a predator.

ToR 7. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).

- **a. Provide numerical short-term projections (1-5years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment.**
- **b. Comment on which projections seem most realistic, taking into consideration uncertainties in the assessment.**
- **c. For a range of candidate ABC scenarios, compute the probabilities of rebuilding the stock by January 1, 2015.**
- **d. Describe this stock's vulnerability to having overfished status (consider mean generation time), and how this could affect the choice of ABC.**

Partially completed: Stochastic projection of the KLAMZ model was used to provide short to medium term projections for the butterfish stock using two different assumptions for recruitment (bootstraps of the full time series or the recent time series), which provide a probabilistic evaluation for exceeding the proposed biomass BRP. However, given the SARC's concerns over accepting either the historic or proposed biomass BRP the ToR as stated cannot be fully completed. During the meeting the SAW brought forward new projections exploring some of the uncertainty outside the assessment framework (e.g. different catchability assumptions) in preparation for compiling a set of decision tables. Differing natural mortality assumptions were not evaluated at this point as these would have increased the number of permutations too much. The group felt this approach was useful in contributing to ToR 7b. ToRs 7c & d are no longer directly relevant given the un-resolved status of the stock.

ToR 8. 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.

Eight research recommendations (rrs) were put forward by SARC 38 and reviewed by the SAW. Several of these (2, 3 & 5) related to discards estimation. The SAW introduced new methods for discard estimation in this assessment that seem likely to have improved some of the catch estimates (possibly particularly those further back in the time series). The SAW investigated the use of ratios based on target species, but ultimately opted for a gear based approach using all species landings. There may be merit in revisiting species directed discards estimation for the fisheries where butterfish are a major by-catch species.

Rr 1 relating to inshore offshore characteristics of butterfish was not fully completed although work was carried out to examine the spatial and seasonal distribution of observed catches and surveys. Given the concerns regarding catchability of butterfish by the surveys this rr could be considered further. A further rr could be to analyse acoustic data relating to surveys to see if (and how often) pelagic fish are able to avoid the (demersal) gear being used by the survey vessel.

Rr 4 relates to natural mortality and rr 6 relates to consumption and inclusion of environmental data relating to distribution and availability to the surveys. Natural mortality was a recurring discussion point throughout the meeting with large inconsistencies between estimates derived from different data and methodologies. Although the group have considered this rr, it needs further work in the future to resolve the current inconsistencies. Uncertainty in survey catchability has also been a major factor in this assessment so additional work to elucidate stock distribution and implications for catchability would be useful.

An rr (7) suggesting exploration of age structured methodology was limited by (lack of) availability of age structured data for the fishery.

The final rr (8) related to reference points for which the basis of estimation was revised in this assessment. High uncertainty in the assessment provides one problem for reference point determination. The unusual circumstance of the stock, which appears to be declining under the influence of factors other than fishing, meant that reference points were considered at some length. The objectives of reference points in these circumstances along with the rationale and methodology for their determination would be useful considerations for future work.

5. Acknowledgements

I would like to thank all the technical working group members contributing to the meeting, for their informative presentations of the SAW results and provision of helpful responses and additional information in response to the SARC's questions. I would also like to thank the staff at the Woods Hole Laboratory and especially Jim Weinberg, Andrea Toran and Paul Rago for their hospitality and guidance throughout the meeting. Thanks are also due to the other members of SARC for productive discussions on the assessments and particularly to Rob Latour (Chair) for keeping the necessary focus under sometimes difficult circumstances.

Appendix 1: Bibliography of documentation provided

Luca M. Cargnelli, L.M., Griesbach, S.J., Packer, D.B. & Weissberger, E., 1999. *Essential Fish Habitat Source Document:* Atlantic Surfclam, *Spisula solidissima*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-142. 22pp.

Mid Atlantic Fishery Management Council, 2007. Amendment 9 to the Atlantic mackerel, squid, and butterfish fishery management plan. Volume I. 415pp.

Mid Atlantic Fishery Management Council, 2007. Amendment 9 to the Atlantic mackerel, squid, and butterfish fishery management plan. Volume II. Appendices. 608pp.

Miller, T.J., Das, C., Politis, P., Long, A., Lucey, S., Legault, C., Brown, R. and Rago, P. 2009. Estimation of *Henry B. Bigelow* calibration factors. Working Paper, Northeast Fisheries Science Center. 286pp.

NEFSC. 1994. Report of the 17th Northeast Regional Stock Assessment Workshop: the plenary. NOAA/NMFS/NEFSC: Woods Hole, MA. *NEFSC* [Northeast Fisheries Science Center] *Ref. Doc.* 94-7. 59pp.

Rago P.J., Weinberg JR, Weidman C. 2006. A spatial model to estimate gear efficiency and animal density from depletion experiments. Can J Fish Aquat Sci. 63:2377-2388.

SAW 38, 2004. Consensus Summary. Atlantic butterfish assessment report. 73pp.

SAW 38, 2004. Consensus Summary Appendix A. Appendix A. (Ocean quahog) KLAMZ Assessment Model – Technical Documentation. 23pp.

SAW 48, 2006. Atlantic surfclam assessment summary for 2006. 15pp.

SAW 48, 2006. Assessment of Atlantic surfclam. 114pp.

SAW/SARC 49, Working Paper A1; Assessment Report, November 13, 2009. Stock assessment for Atlantic Surfclams in the US EEZ. 42pp.

SAW/SARC 49, Working Paper A1a; Appendixes, November 13, 2009. Surfclam appendixes. 91pp.

SAW/SARC 49, Working Paper A1b; Tables, November 13, 2009. Surfclam; Tables. 37pp.

SAW/SARC 49, Working Paper A1c; Figures, November 13, 2009. Surfclam; Figures. 79pp.

SAW/SARC 49, Working Paper A2; Assessment Summary Report, November 13, 2009. 25pp.

SAW/SARC 49, Working Paper B1; Assessment Report, November 18, 2009. Stock Assessment of Butterfish (*Peprilus triacanthus*) in the Northwest Atlantic. 104pp.

SAW/SARC 49, Working Paper B2; Assessment Summary Report, November 18, 2009. Atlantic Butterfish Assessment Summary Report. 5pp.

Wallace, D.H. & Hoff, T.B. 2005. Hydraulic clam dredge effects on benthic habitat off the northeastern United States. Am. Fish. Soc. Symp. 41:691-693.

Weinberg, J.R., 2005. Bathymetric shift in the distribution of Atlantic surfclams: response to warmer ocean temperature. ICES Journal of Marine Science, 62: 1444-1453.

Wigley, S. E., Rago, P. J., Sosebee, K. A. and Palka, D. L. 2006. The analytic component to the standardized bycatch reporting methodology omnibus amendment: sampling design and estimation of precision and accuracy. NMFS/NEFSC, Reference Document 06-22.

Appendix 2: CIE Statement of Work

Statement of Work

(T016-07, v 11 September 2009)

External Independent Peer Review by the Center for Independent Experts

49th Stock Assessment Workshop/ Stock Assessment Review Committee (SAW/SARC)

Atlantic Surfclam and Butterfish

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

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

Project Description: The purpose of this SARC49 meeting will be to provide an external peer review of benchmark stock assessments for Atlantic surfclam (*Spisula solidissima)* and butterfish *(Peprilus triacanthus)*. Surfclams are sedentary infaunal bivalves. Butterfish are a schooling pelagic fish. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results form the scientific basis for fishery management in the northeast region. This meeting satisfies Prioritization criteria 1-3. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is attached as **Annex 4.**

The SARC49 review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the Science and Statistics Committee (SSC) of the New England or Mid-Atlantic Fishery Management

Council. The SARC panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Familiarity with statistical models for estimating gear efficiency is desirable, as the surfclam assessment will apply methods for experimentally estimating survey dredge capture efficiency. For butterfish, reviewers should be familiar with schooling pelagic species whose catchability in research trawl surveys is highly variable and influenced by environmental conditions; expertise in discard estimation for pelagic species and in the analysis and interpretation of trawl surveys is desirable.

Reviewer expertise should include statistical catch-at-age, catch-at-length, delay-difference, and traditional VPA approaches. 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 their estimation. Reviewers should have familiarity with the development and interpretation of rebuilding strategies. Experience with the biology and population dynamics of species on the agenda would be useful.

Each CIE reviewer's duties shall not exceed a maximum of 14 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 14 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation).

Location and Date of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts during 30 November through 3 December 2009.

Charge to SARC panel: The panel is to determine and write down whether each Term of Reference of the SAW (see Annex 2) was or was not completed successfully during the SARC meeting. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. Where possible, the chair shall identify or facilitate agreement among the reviewers for each Term of Reference of the SAW.

If the panel rejects any of the current Biological Reference Point (BRP) proxies for B_{MSY} and F_{MSY}, the panel should explain why those particular 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 are the best available at this time.

Statement of Tasks:

1. Prior to the meeting

(SARC chair and CIE reviewers)

Review the reports produced by the Working Groups and read background reports.

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

Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

2. During the Open meeting

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

(SARC chair)
Act as chairperson, where duties include control of the meeting, coordination of presentations and discussion, making sure all Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For the assessment, review both the Assessment Report and the draft Assessment Summary Report.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to discuss the stock assessment and to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(SARC CIE reviewers)
For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a reviewer's point of view, determine whether each 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 proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

3. After the Open meeting

(SARC CIE reviewers)

Each CIE reviewer shall prepare an Independent CIE Report (see Annex 1). This report should explain whether each Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the "Charge to SARC panel" statement.

If any existing Biological Reference Point (BRP) 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 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 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.

(SARC chair and CIE reviewers)
The SARC Chair and CIE reviewers will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each 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 or a consensual 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 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 Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

The contents of the draft SARC Summary Report will be approved by the CIE reviewers by the end of the SARC Summary Report development process. The SARC chair will complete all final editorial and formatting changes prior to approval of the contents of the draft SARC Summary Report by the CIE reviewers. The SARC chair will then submit the approved SARC Summary Report to the NEFSC contact (i.e., SAW Chairman).

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to SARC Summary Report: Each CIE reviewer will assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. CIE reviewers are not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

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 in Woods Hole, Massachusetts during 30 November through 3 December 2009, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 17 December 2009, 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 Dr. David Sampson, CIE Regional Coordinator, via

email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**.

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

26 Oct 2009	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
16 Nov 2009	NMFS Project Contact will attempt to provides CIE Reviewers the pre-review documents by this date
30 Nov – 3 Dec 2009	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
2-3 Dec 2009	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
17 Dec 2009	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
21 Dec 2009	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
29 Dec 2009	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
31 Dec 2009	CIE submits CIE independent peer review reports to the COTR
7 Jan 2010	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

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

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

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

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,
(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Contracting Officer's Technical Representative (COTR)
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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

Key Personnel:

Dr. James Weinberg, NEFSC Stock Assess. Workshop (SAW) Chair, (NMFS Project Contact)
National Marine Fisheries Service, NOAA
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
james.weinberg@noaa.gov Phone: 508-495-2352

Dr. Nancy Thompson, NEFSC Science Director
National Marine Fisheries Service, NOAA
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
nancy.thompson@noaa.gov Phone: 508-495-2233

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each ToR, and Conclusions and Recommendations in accordance with the ToRs. For each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, the Independent Review Report should state why that Term of Reference was or was not completed successfully. To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, regardless of whether or not others read the SARC Summary Report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include as separate appendices as follows:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

ANNEX 2:

Assessment Terms of Reference for SAW/SARC49 (Nov-Dec 2009)

(file vers.: 8/12/09)

A. Atlantic surfclam

1. Characterize the commercial catch including landings, effort, LPUE and discards. Describe the uncertainty in these sources of data.

2. Characterize the survey data that are being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, age-length data, etc.). Describe the uncertainty in these sources of data.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and characterize the uncertainty of those estimates.

4. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and $F_{MSY;}$ and estimates of their uncertainty). Comment on the scientific adequacy of existing and redefined BRPs.

5. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 4).

6. Identify potential environmental, ecological, and fishing-related factors that could be responsible for low recruitment.

7. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).
   d. Provide numerical short-term projections (1-5 years; through 2015). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment.
   e. Comment on which projections seem most realistic, taking into consideration uncertainties in the assessment.
   f. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC.

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

B. Butterfish

1. Characterize the commercial catch including landings, effort and discards by fishery (i.e., *Loligo* fishery vs other fisheries). Characterize recreational landings. Describe the uncertainty in these sources of data. Evaluate the precision of the bycatch data with respect to achieving temporal management objectives throughout the year.

2. Characterize the survey data that are being used in the assessment (e.g., indices of abundance including RV Bigelow data, NEAMAP and state surveys, age-length data, etc.). Describe the uncertainty in these sources of data.

3. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and characterize the uncertainty of those estimates.

4. Update or redefine biological reference points (BRPs; estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, and $F_{MSY;}$ and estimates of their uncertainty). Comment on the scientific adequacy of existing and redefined BRPs.

5. Evaluate stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 4).

6. Evaluate the magnitude, trends and uncertainty of predator consumptive removals on butterfish and associated predation mortality estimates and, if feasible, incorporate said mortality predation estimates into models of population dynamics.

7. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs).
    e. Provide numerical short-term projections (1-5years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment.
    f. Comment on which projections seem most realistic, taking into consideration uncertainties in the assessment.
    g. For a range of candidate ABC scenarios, compute the probabilities of rebuilding the stock by January 1, 2015.
    h. Describe this stock's vulnerability to having overfished status (consider mean generation time), and how this could affect the choice of ABC.

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

Appendix to the SAW Assessment TORs:

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

(The text below is from DOC National Standard Guidelines, Federal Register, vol. 74, no. 11, January 16, 2009)

On "Acceptable Biological Catch":

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability":

"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)

Annex 3: Meeting Agenda

49th Northeast Regional Stock Assessment Workshop (SAW 49) Stock Assessment Review Committee (SARC) Meeting

November 30 – December 3, 2009

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

DRAFT AGENDA* (version: 9-11-09)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Monday, 30 Nov

1:00 – 1:30 PM

Opening
Welcome — **James Weinberg**, SAW Chairman
Introduction — **TBD**, SARC Chairman
Agenda
Conduct of Meeting

1:30 – 3:30 Assessment Presentation Surfclam (Sp. A)
Larry Jacobson **TBD** **TBD**

3:30 – 3:45 Break

3:45 – 5:30 SARC Discussion of Surfclam
Larry Jacobson, SARC Chairman

Tuesday, 1 Dec

9:00 – 10:30 AM Assessment Presentation Butterfish (Sp. B)
Tim Miller **TBD** **TBD**

10:30 – 10:45 Break

10:45 – Noon SARC Discussion of Butterfish
Tim Miller, SARC Chairman

Noon – 1:15 Lunch

1:15 – 2:15 Continue SARC Discussion of Butterfish
TBD, SARC Chairman

2:15 – 3:30 Revisit Surfclam Assessment with Presenters

3:30 – 3:45 Break

3:45 – 5:30 Revisit Surfclam and/or Butterfish Assessments with Presenters

Wednesday, 2 Dec

9:00 – 10:00 AM Revisit Butterfish Assessment with Presenters

10:00 – 10:15 Break

10:15 - Noon Surfclam follow up + review Assessment Summary Report

Noon – 1:15 PM Lunch

1:15 – 3:00 Butterfish follow up + review Assessment Summary Report

3:00 – 3:15 Break

3:15 – 5:15 SARC Report writing. (closed meeting)

Thursday, 3 Dec

9:00 – 2:00 PM SARC Report writing. (closed meeting)

*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 was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

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

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

2. If any existing Biological Reference Point (BRP) proxies are considered inappropriate, include recommendations and justification for alternative proxies. If such alternatives cannot be identified, then indicate that the existing BRPs 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 Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 3: Review Committee members

Robert Latour, chair
John Cotter
Michael Smith
Henrik Sparholt