Independent Peer Review Report of the 59th SAW/SARC Stock Assessment

Dr Coby Needle

Center for Independent Experts (CIE) Peer Reviewer

August 2014

Executive Summary

This report summarises my notes and conclusions on the SARC 59 assessments of Gulf of Maine haddock and Atlantic sea scallops, presented during the SARC 59 meeting at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA, during 14-18 July 2014. I found the background reading and the meeting itself to be extremely informative and well-structured, and that the ToRs had all been met to the extent possible (the only issues were driven by a lack of some data which the stock assessors could not have addressed). Both assessments were well presented and contained a great deal of relevant information, and I was happy to accept them both as valid representations of stock status. I did not identify any serious weaknesses, but there are a number of issues that I think could improve the assessments still further if addressed: principally, inferred age compositions for haddock, and spatial assessment approaches for sea scallops.

Background

I am an applied mathematician and modeller by training, and I have worked in quantitative fisheries science since 1996. Having served as the Chair of the ICES Working Groups on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK, 2004—2006) and Methods of Stock Assessment (WGMG, 2007—2009), I now lead the Fishery Assessment and Analysis Group at the Marine Laboratory in Aberdeen, Scotland. One of our key roles is the analysis and interpretation of data from the fishing industry. I also still lead on the ICES assessment of Northern Shelf haddock and I am heavily involved in the provision of fisheries advice to the Scottish Government.

Role in the Review Activities

Prior to the SARC 59 meeting in Woods Hole (15-18 July 2014), I thoroughly reviewed the background documents provided for the review panel, along with the extant versions of the stock assessment reports and summary reports for Gulf of Maine haddock and Atlantic sea scallops (see Appendix 1 for a document list). During the SARC 59 meeting I participated in full in the plenary discussions during and after the presentations provided, as well as intersessionally with the other review panel members and the SARC chair. I took copious notes during these discussions which form the basis of my comments below. In addition, I was tasked with fulfilling the role of SARC Expert on Gulf of Maine haddock, so that I was primarily responsible for completing the SARC Summary Report on that stock.

Overall Conclusions and Suggestions for Improvements to the SARC Process

My most recent experience with a benchmark-type stock assessment process was the ICES Northern haddock benchmark meeting (WKHAD), for which I was the principal stock coordinator. In comparison with the ICES approach, I found the SARC/SAW assessment process to be highly structured and formalised, and I felt that it benefitted from this. Throughout the meeting it was very clear what was expected, and what the information and subsequent conclusions would be used for, and I considered the meeting and bracketing work to be very instructive and educational. The SAW participants and those who presented at the SARC meeting had clearly invested a great deal of time and effort in producing work of a very high standard, and everyone involved is to be thoroughly commended.

The principal advantage of the ICES process that I felt was lacking in the SARC system was the opportunity given to the review panel to make very constructive contributions to the assessments themselves. We were able to comment and review the existing work, and should we have seen the need we could have rejected the assessments, but the time available for the review was too short for further data or model developments, and I found this somewhat frustrating. ICES reviewers are more involved in the actual assessment work, at least in my experience, and so have more opportunity to make a constructive difference. It is of course more difficult to get reviewers to set aside the additional time needed for this, and it is precluded by the timetable of the SAW/SARC process, but I think I would have appreciated more opportunity to get more involved, had it been possible.

Regarding the assessments themselves, I found that in both cases a tremendous amount of work had been presented and that the great majority of the ToRs had been fulfilled. The Gulf of Maine haddock assessment did as good a job as could have been expected given the commercial and recreational catch data available, although there was a good deal of data inference and imputation required to generate full age-structured datasets for landings and discards for both of these fisheries. It might be beneficial in future work to explore the stock dynamics indicated by approaches that require less complete datasets (such as length- or survey-based assessment methods). Given the clear stock signals, particularly from the recent strong year-classes, it is unlikely that these approaches would necessarily lead to different perceptions of stock dynamics, but they would at least increase confidence in the baseline catch-at-age stock assessment, which is driven to an extent by inferred age compositions. The Atlantic sea scallop assessment collated a great deal of fascinating biological and fishery data and provided a compelling summary of the past and likely future dynamics of this valuable stock, which (like GoM haddock) looks to be in good shape following strong recent year-classes. My main criticisms for the sea scallop assessment would be a lingering concern that the splitting of the assessment into open and non-contiguous closed areas might be violating basic assessment assumptions of closed populations (although I would not be sufficiently concerned about this to reject the assessment); and that the wealth of information presented meant that the actual time for discussion and evaluation was quite limited.

Overall, I was very impressed by the ability and dedication of the participating scientists, which facilitated a successful and enjoyable meeting.

Comments for each ToR of each assessment

A. Gulf of Maine (GoM) haddock

Note that some of the following text is based in part on the corresponding text in the SARC Summary Report. I was the lead review panel member for GoM haddock, and as there was little or no disagreement in the panel about our conclusions the text I wrote for the Summary Report is very close to the following independent report. However, the text has been modified in a number of places. I also carried out some further analyses using the haddock data, and these are only reported here as they were not discussed during the SARC meeting.

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

This ToR was met, and I accept the Panel findings regarding it.

GoM haddock catch data consist of four components: commercial landings, commercial discards, recreational landings and recreational discards. A great deal of work was presented at SARC 59 on the estimation of commercial discards, although these are a small component of the overall catch. While this was also true historically of recreational landings and discards, the recreational catch is now a very significant part of the overall yield (see Figure CLN.A1).

Recreational discards were included in the assessment for the first time at SARC 59, based on surveys conducted by another US government department. However, recreational discard mortality remains unknown, and there are no studies planned to address this lack. Two sets of results were presented assuming recreational discard mortality of 50% or 100%, with the former being chosen for the final run, but this choice is essentially ad hoc and provides a key source of uncertainty in the assessment. In the future, it will be important to collect empirical data on post-release survival from the recreational fishery.

The only full time-series of catch yield data exists for commercial landings. For the other three components, there are varying amounts of imputation and inference used to generate time-series yield estimates of equal length. In my opinion, it would be more appropriate in future to run the assessment only for those years for which real data exist (from around 1989 onwards). The pattern of historical stock development means that there is no real benefit to extending the time series further back than this, and an assessment based only on observed data is likely to be more justifiable.

Age data are only collected for the commercial landings component of the catch, and for the surveys. The age compositions of the remaining catch components are inferred by the application of the survey age-length keys (ALKs) to length measurements. There is a potential for this to cause bias in age compositions, particularly for the recreational fishery for which the gear used and areas fished are considerably different to the commercial fishery. It may be that the ALKs are similar between components, but it would be inappropriate to assume this on the basis of limited evidence, and the review panel recommends that a trial programme of age-reading from the recreational fishery be carried out to determine if the recreational ALKs are different or not. The extant catch-at-age data suggest very low selectivity on younger ages and highly variable catch curves (Figure CLN.A2), and although estimated year-class strength is reasonably consistent from age to age along cohorts (Figure CLN.A3), there is quite a lot of noise about these bivariate relationships and I hypothesise some of that could be due to inferred age compositions. Length or survey based assessment methods should also be considered (at least comparatively): I give an example of a survey-based assessment for this stock in my comments for ToR A4.

The draft report did not contain sufficient information on the methodology of the recreational landings and discards sampling programme. This was addressed during the SARC 59 meeting to the satisfaction of the review panel.

Weights-at-age in the catch data show considerable reductions over time for older fish, with sharp declines followed by more stability in the most recent 5 years (see Figure CLN.A4). No hypothesis was proposed to attempt to explain this, and it does not seem to be related to commensurate changes in maturity (as is the case for many gadoids in European waters). The review panel queried whether there may be spatial differences in weights-at-age that might cause changes in the whole-stock weights-at-age through spatially heterogeneous exploitation and sampling, and recommends that this possibility be considered in future work.

The spatial distribution of catches was illustrated through maps, time-series plots of the centres of distribution of catches, and Gini indices (which measure the concentration of catch locations). There was less spatial information provided about fishing effort, and it may be that VTR-derived fishing locations are not very reliable due to the inconsistent way in which they are recorded. More use could be made of VMS data for this purpose, given that 100% of the relevant commercial vessels are fitted with these systems. Uncertainty in catch data was expressed through estimated CVs.

Following a review panel request, a table was produced giving the ratio of SARC 59 catch estimated to the previous catch estimates. As expected, there were relatively large differences at ages 1 and 2, but as the catches are small at these ages the changes do not make a large difference to the assessment results.

Considerable effort was spent on estimation of LPUE for commercial and recreational fisheries. During this work, it became clear that changes in fishing practices, gear and location, along with the possibility of hyperaggregation, mean that LPUE for GoM haddock is not a reliable indicator of stock status or dynamics. The review panel consider that future methodological work in this area should be carried generically, rather than on a species-by-species basis, and that LPUE estimation for GoM haddock in particular need not be reconsidered for a number of years. I would recommend further that simple LPUE is almost never a good indicator of stock dynamics, due to fishery aggregation on areas of high abundance, and that more directed and detailed measures of fishing activity (VMS or other remote electronic monitoring, for example) should be considered for this stock.

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

This ToR was largely met, although there was no information on the use of tagging data to estimate stock abundance or exploitation. I accept the Panel findings regarding the ToR.

The NEMFS spring and fall surveys are both considered to be consistent and independent surveys in the assessment. However, for both there have been significant changes over time in vessel, net and door characteristics, and the use of conversion factors to try and ensure consistency may not be appropriate. Given that there are now five years' data from the Bigelow series, it would be possible to treat that as a separate new index, and this was tried as an alternative assessment run (making no significant difference to the outcome). However, this would not address the changes in the net and trawl doors, and an alternative approach would be to develop priors on net, door and vessel effects for the model estimation.

Following the meeting, I generated standard exploratory plots for the NEMFS surveys as would commonly be found in ICES assessment working group reports. Figure CLN.A5 shows that the survey series are reasonably consistent at age, although there are fluctuations in each series that are not always reflected in the other series. The survey catch curves (Figure CLN.A6) are extremely noisy and it is not at all clear that year-class strength is being tracked well, while my comments about the total catch bivariate scatterplots (Figure CLN.A3) apply equally well to the survey scatterplot in Figure CLN.A7 – there is some consistency along cohorts but also considerable variation, and these surveys both appear quite noisy to be used as the only source of fishery-independent data in the assessment.

The MADMF and MENH series are principally inshore surveys that do not cover the full extent of the GoM haddock stock, and do not track cohorts well through time. However, it might be possible to use the age 1 and 2 data from them as indices of recruitment.

The report hypthesised a westerly concentration of the stock over time, as evidenced by survey distribution maps (Figure A.101). The review panel did not find this conclusion convincing, as the plot for 1970-1979 looked very similar to the plot for 2000-2013, and the Gini index (Figure A.102) did not support the hypothesis either.

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

This ToR was partially met. I accept the Panel findings regarding it.

I was particularly interested in this ToR, as I spent much of the early part of 2014 working towards an ICES benchmark assessment of haddock in the North Sea and West of Scotland. These areas were historically assessed as two stocks, but we managed to demonstrate that there were strong biological links between them and that they should be assessed as a single stock. The situation was very similar to that off the eastern coast of the USA, with two adjoining stocks of very different sizes being fished by the same vessels.

However, in the event it seemed that much of the evidence provided in the assessment report for this ToR was rather circumstantial, and I and the review panel concluded that it would not be possible to determine a mixing rate on the basis of the evidence provided. If concerns about mixing rates and stock structure remain, biological analyses such as directed tagging studies, egg dispersal modelling, genetic differentiation determination or otolith microchemistry analysis would be needed to determine the degree of mixing, and hence stock identity.

In terms of management, the conclusion of low mixing between separate stocks is probably appropriate. In terms of scientific evaluation of mixing or stock structure, the review panel concluded that more work would be required to reach firm conclusions.

The tagging studies presented investigated whether haddock from within closed areas were likely to move outside such areas. Around 20,000 tags were applied, and only around 500 returned. This could indicate a very high tagging mortality, and the way in which haddock tagging is carried out in future will need to be carefully considered.

Potential causal factors such as density dependence or environmental conditions were not considered in detail.

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

This ToR was met. I accept the Panel findings regarding it.

This stock was previously assessed using a VPA-type method that considered catches to be measured exactly. In light of the age-composition problems alluded to above, along with other sampling issues from the earlier time period, the SAW meeting concluded that an alternative statistical catch-at-age method should be used. The reasons for this choice were clear – it allowed for uncertainty in catch data, and there was a great deal of local expertise in its implementation (Chris Legault both wrote the code and was a member of the SAW meeting). However, I didn't think that there had been much consideration of alternative approaches to uncertain catch data, including smoothing through it (time-series models), removing the issue of age compositions (length-based models), and not using it at all (survey-based models). I recommend that these or other possibilities be explored for the next benchmark meeting, at least from an exploratory perspective.

The lack of age- or year-specific natural mortality estimates is problematic. In the absence of multispecies models based on stomach contents data it is difficult to see how M estimates could be generated, but it is clear that a fixed M of 0.2 across all ages and years is highly unlikely to be correct. The assessment is essentially providing advice for total mortality (Z) when M is treated in this way, and this in itself is not a problem as long as the true M does not change. If it does, the fishing fleets (or fish stocks) could find themselves penalised unnecessarily.

I noted two additional points regarding the final ASAP assessment. Firstly, the iterative determination of the "effective sample size" seems to be a way to reduce the influence of poorly-sampled years. cf. the implementation of missing years in TSA or SAM assessments. But doesn't it actually measure how well the data conform to the assumptions of the model? Secondly, the higher effective weight on catch-at-age data than on survey-at-age data is difficult to understand given the use of survey ALKs to produce much of the catch-at-age data.

Finally, after the meeting I ran the survey data provided through the SURBAR method, widely used in ICES for exploratory analyses (and occasionally for the final advice). The method is described in full in ICES (2009) and Needle (2012, in prep). For this implementation, I assumed a smoother of the age and year effects of total mortality of $\lambda = 5.0$ as the survey data are quite noisy. The model converged quickly with this setting. The stock summary plots are given in Figure CLN.A8 (line plot) and CLN.A9 (box-and-whisker plot), and indicate SSB was initially high before falling to very low levels in the mid-1980s to mid-1990s, and then rising to a maximum in the early 2000s. The only notable historical recruitment events were for the 1998 and 2010 year-classes, while the size of the 2011 year-class remains extremely uncertain (although it is likely to have been large). The estimated mean Z is extremely uncertain, with little evidence that it changed significantly during the time period, and this is likely to be due to noisy survey data. The residual plot in Figure CLN.A9 shows that the model fits the survey data reasonably well, without strong patterns or significant outliers. Finally, Figure CLN.A10 compares the best-estimate SURBAR results with the corresponding ASAP estimates, converting to mean Z and mean-standardising recruitment and SSB to facilitate comparison (SURBAR produces relative stock estimates only). We can see that the recruitment and SSB estimates are largely very similar, save for the final year-class strength which SURBAR (and therefore the surveys) indicate to be much higher than allowed for in the ASAP assessment. The mean Z estimates are not close, but the SURBAR mean Z values are very uncertain (Figures CLN.A8 and CLN.A9) due to noisy indices so it is not surprising that mortality is not consistently estimated. The conclusion from this work is that the surveys mostly agree with the full assessment in terms of abundance, and less so in terms of mortality, but in any case that exploratory analyses of this type can point to potential issues with the final assessment that may need to be addressed.

ICES (2009). Report of the Working Group on Methods of Fish Stock Assessment. ICES CM 2009/RMC: 12.

Needle, C. L. (2012). Fleet Dynamics in Fisheries Management Strategy Evaluations, PhD thesis, University of Strathclyde, Glasgow.

Needle, C. L. (in prep). Using self-testing to validate the SURBAR survey-based assessment model. Submitted to ICES Journal of Marine Science.

A.5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This ToR was met, and I accept the Panel findings regarding it.

The selectivity ogive presented implies that there could be changes in *F* reference points. However, the fishing mortality on mature fish is probably not that high. It might be useful to have a measure of exploitation rate that would be more stable and more indicative of the exploitation on the mature part of the population.

A.6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

This ToR was met, and I accept the Panel findings regarding it.

Considering the management approach, my view is that the assessment process is conducted too infrequently for GoM haddock. Any haddock stock is liable to undergo highly episodic recruitment events, with large year-classes appearing unexpectedly at irregular intervals, and fixed multi-annual TACs can quickly become inappropriate following one of these events. We suggested that consideration be given to the approach used for salmon in some ICES advice: here a 3-5 year stock projection is made in which the abundance indices are also projected. The update procedure then consists of comparing the realized (observed) indices with the projected indices for each year, and revision action is only taken if the observations lie outwith a prespecified confidence bound about the projections. This could serve to keep management relevant whilst not overburdening the advisory system with unnecessary work.

A.6.a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

This ToR was met, and I accept the Panel findings regarding it.

A.6.b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

This ToR was met, and I accept the Panel findings regarding it.

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

This ToR was met, and I accept the Panel findings regarding it.

A.7.a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, migration from Georges Bank).

This ToR was met, and I accept the Panel findings regarding it.

The review panel requested the spring 2014 survey results, and these were collated and provided: the indications are that both the 2012 and 2013 year-classes are high. Otherwise the ToR seemed to be met to our satisfaction.

A.7.b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

This ToR was met, and I accept the Panel findings regarding it.

A.7.c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

This ToR was met, and I accept the Panel findings regarding it.

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

This ToR was met, and I accept the Panel findings regarding it.

The review panel (particularly myself) considered that the use of a fixed TAC for three years following a benchmark assessment is problematic. Notwithstanding the legal requirements for fixed TACs in the US system, a fixed TAC for an episodic spawning stock like haddock is always likely to increase the risk of over- or under-exploitation. The review panel's suggestion for addressing this is outlined in the comments for ToR A6 above.

Stock identity remains a serious concern with this assessment. The Gulf of Maine haddock stock is currently considered to be entirely distinct from the Georges Bank haddock stock (and indeed the Scotian Shelf stock), although survey distribution maps indicate a relatively continuous extent of haddock across area boundaries. Existing data indicate a low level of mixing between the GoM and GB stocks, but further biological data collection and analyses would be required to reach firm conclusions. Following on from our work on the ICES Northern Shelf haddock stock, the review panel suggested that approaches such as directed tagging studies, egg dispersal modelling, genetic differentiation determination or otolith microchemistry analysis would be needed to determine the degree of mixing, and hence stock identity.

The review panel recommended that a trial programme of age-reading from the recreational fishery be carried out to determine if the recreational ALKs are different or not.

To address concerns about the consistency of survey index time series following significant changes in vessel, net and door configurations, the review panel suggested that it would be appropriate to develop priors on net, door and vessel effects for the model estimation.

Finally, we proposed that it would be beneficial for future assessments to incorporate survey-based assessments of the type illustrated here, to compare and contrast with the full catch-at-age assessment in an exploratory fashion.

B. Atlantic sea scallop

Although the review panel was charged with providing reviews on each stock assessment equally, the spread of abilities and experience in the panel meant that I focussed (to a certain extent) on haddock, Yiota Apostolaki on scallops, and Vivian Haist on modelling. Hence, the following comments on scallops represent my own notes and do not necessarily replicate in full the review panel comments to the same extent as for haddock.

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

This ToR was met, and I accept the Panel findings regarding it.

This proved to be a fascinating presentation delivered by an engaging and enthusiastic speaker: the only drawback was that the wealth of information presented was difficult to digest in one sitting, and the depth of explanation meant that quite a long time was taken which precluded a great deal of directed discussion, and which made some of the work more difficult to review. However, I certainly felt that I learned a great deal about sea scallops in this area, and the ToR was dealt with very fully.

One of my principal concerns was whether we were dealing with one stock, or whether a split between Georges Bank and the Mid Atlantic Bight stocks was justified. The lack of recruitment synchrony implies that there could be doubts on this issue. I also raised a question about whether there could be links with the Canadian scallop stock to the north. Adult movement is limited, but there is certainly dispersal of larvae and I'm not sure the stock structure issue has been addressed conclusively. For example, it might be appropriate to consider a joint assessment with Canada for the more northerly scallops.

There is a clear seasonal pattern in meat weight, which (it was explained) is due to a combination of feeding and spawning – meat weights can decline towards the autumn post-spawning and with poorer feeding (just like a fish getting skinny). Natural mortality estimates (Merrill and Posgay) are based on ratio of "clappers" (empty but still linked shells) to live scallops – this has now been revised slightly to account for different estimates of "separation" (i.e. how long a clapper stays intact). I thought that it would have been useful to see the summary plots for the natural mortality calculation in the report. I was interested to note that natural mortality increases with age – it seems that young scallops are vulnerable to crab predation, but can swim away from starfish, whereas old scallops can't swim very well and so more vulnerable to starfish, but also parasite loading. These two latter factors are more lethal than crab predation.

Following several years of analysis on the determination and potential impact of real-time closures in the North Sea (intended to reduce mortality on cod), I was keen to learn more about the scallop closed areas. I queried how a decision is made to reopen a closed area? The response was that closures persist for two to three years, and a decision is made to reopen on the basis of targeted surveys to determine likely catch abundance. These closures are therefore at a much greater scale, both spatially and temporally, than those in the North Sea, and are therefore much more likely to have an effect (and in this case, a beneficial one).

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

This ToR was met, and I accept the Panel findings regarding it.

This presentation also proved to be extremely informative, and addressed the ToR very well. In Scotland we have been working very closely over the past three or four years on trawler discard estimation via on-board cameras, so I was interested to learn about the work being done with the Habcam system for counting and measuring live scallops along transects. It seemed to me that the images would be an ideal case study for automated image analysis, and while the assessors are in the process of working with computer scientists to develop appropriate systems, an operational framework is unlikely to be available for at least a couple of years. Our work on image analysis is at a similar stage of development, and there is scope for future collaboration.

The Habcam system did look particularly impressive, with excellent image quality, scaling calibration through a combination of altimeters and 3D image analysis and no apparent avoidance behaviour. The tendency for younger scallops to be attracted to the lights of the system could in theory affect density estimates, but the Habcam system moves too fast through the water for this to be a problem. The sonar images provided also would be very useful in determining the potential benthic impact of recent trawl activity.

There was also a very interesting presentation on the use of geostatistics to determine optimum transects for Habcam surveys. But (again) I found this quite difficult to review – there was a lot here to digest, and my knowledge of advanced geostatistics was insufficient to make much sensible comment.

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

This ToR was met, and I accept the Panel findings regarding it.

As before, the main problem with this presentation was the sheer wealth of information which made it very hard to digest and review. There does seem to be a good relationship between phytoplankton blooms and recruitment, and important predation from *Astropecten* (starfish) which seems to exclude small scallops from venturing into deep water (> 80m). However, with studies of this kind I would always recommend that R. A. Myers' (1998) work in the potential unreliability of environmental drivers of recruitment be borne in mind, along with Basson's (1999) study on the problems that can arise if environmental drivers are used incorrectly in the development of management advice. There have been enough examples of long-standing environmental correlates of recruitment being used to forecast year-class strength (and hence quotas) in the very year that the correlation broke down, and great care needs to be taken if this is to be done.

Basson, M. (1999) The importance of environmental factors in the design of management procedures. *ICES Journal of Marine Science*, **56**, 933-942.

Myers, R. A. (1998) When do recruitment-environment correlations work? *Reviews in Fish Biology and Fisheries*, **8**, 285-305.

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

This ToR was met, and I accept the Panel findings regarding it.

The stock assessment for scallops was conducted using a bespoke version of the CASA model of Sullivan et al. (1990) to allow for difficulties in scallop ageing. The model appeared to be well-implemented and the diagnostics for the model fit were generally good, and on this basis I could see no reason to reject the assessment.

My only real concern lay with the decision to conduct separate assessments for the closed and open regions of Georges Bank. For me, this stock is a good candidate for an explicitly spatial assessment model. Although Hart et al. (2013) concluded that the use of separate assessments for closed and open areas in Georges Bank provides more quantitatively robust stock estimates, the approach is still problematic because a) there is larval transport between the closed areas, b) the closed areas are not contiguous, and c) closures are not constant through time (so selectivity differences are blurred). Therefore the usual assessment model assumption of a discrete closed population does not necessarily hold. There are a number of groups (including one of my students in Aberdeen, Tanja Buch) currently working on spatial assessment methods which could be appropriate here, with probably the furthest advances so far that I know of being made in Denmark (Kristensen et al, 2014; Nielsen et al, 2014), and I would recommend that these be considered for possible future exploratory assessments of Atlantic sea scallops.

Kristensen, K., Thygesen, U. H., Andersen, K. H., and Beyer, J. E. (2014). Estimating spatio-temporal dynamics of size-structured populations. *Canadian Journal of Fisheries and Aquatic Science*, **71**: 326–336

Nielsen J.R., Kristensen K., Lewy P., and Bastardie F. (2014). A statistical model for estimation of fish density including correlation in size, space, time and between species from research survey data. *PLoS ONE* **9**(6): e99151. doi:10.1371/journal.pone.0099151.

Sullivan, P. J., Lai, H.-L. and Gallucci, V. F. (1990). A catch-at-length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Science*, **47**, 184-198.

B.5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This ToR was met, and I accept the Panel findings regarding it.

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

This ToR was met, and I accept the Panel findings regarding it.

B.6.a. Update the existing model with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

This ToR was met, and I accept the Panel findings regarding it.

B.6.b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

This ToR was met, and I accept the Panel findings regarding it.

B.7. Evaluate the realism of stock and catch projections and compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level).

This ToR was met, and I accept the Panel findings regarding it.

B.7.a. Provide numerical annual projections (through 2016). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions

about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

This ToR was met, and I accept the Panel findings regarding it.

B.7.b. Comment on the realism of the projections. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

This ToR was met, and I accept the Panel findings regarding it.

The approach used here seemed to be very similar to that which is commonly used for the same purpose at ICES meetings – stochastic combinations of yield-per-recruit and stock-recruit curves to estimate the maximum yield achievable over a range of F multipliers. I would have preferred that the multiplier range was extended (out to 10, if need be) so that the yield bar charts didn't have those extraneous high bars at the right end – the conclusions on F(msy) would have been the same, and it would have been better presentationally. But overall I agreed with the outcome as given.

We did have interesting discussion on the management approach for this stock. The reference points are set for the whole stock, but the exploitation and selectivity are potentially very different between Georges Bank (open), Georges Bank (closed) and the Mid Atlantic Bight. This is very similar to the situation for Northern Shelf haddock, which is also considered to be a single biological stock with different fisheries and environmental influences in different areas. Any potential overexploitation of subcomponents is dealt with here by temporary area closures (which can be very large) – hence the difference in the F(msy) estimates between GB and MAB may not be all that problematic.

B.7.c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

This ToR was met, and I accept the Panel findings regarding it.

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

This ToR was met, and I had no further comments here.

Appendix 1: Bibliography of materials provided for review

Gulf of Maine haddock

Palmer, M. (2008) Gulf of Maine haddock. Extract from 2008 GARM III report. NEFSC. [2008_GOMHad_GARM3R.pdf]

NEFSC (2012). Technical Documentation for ASAP Version 3.0, NOAA Fisheries Toolbox, September 2012. [ASAP_v3_Technical_Documentation.pdf].

Richard J. Bell and Jon Hare (2014). Gulf of Maine Haddock Stock Recruitment Model: Time Varying Parameters. NEFSC unpublished manuscript. [BellR_Time_varying_par_SR_GoM_Haddock_20140527.pdf]

Michael C. Palmer, Steven Correia, Paul Nitschke (2014). Estimating the year-class size of terminal year cohorts in stock assessment models: the Gulf of Maine haddock (Melanogrammus aeglefinus) example. NEFSC unpublished manuscript. [Estimating_terminal_cohort_size_GoM_haddock_v5_SARC59_WP.pdf].

NEFSC (2012). Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. Northeast Fisheries Science Center Reference Document 12-06. [Groundfish_Update_2012_crd1206.pdf].

Cape Cod Commercial Hook Fishermen's Association (2009). Haddock Migration in New England Waters: Year 1 and Year 2 Analysis of Closed Area and Stock Boundaries: Final Report to the Northeast Consortium. [Haddock_Migration_Final_2009.pdf].

Jon K.T. Brodziak, Laurel Col, Michael Palmer and Liz Brooks (2008). Northeast Consortium Cooperative Haddock Tagging Project: Summary of Reported Haddock Tag Recaptures Through September, 2008. NEFSC unpublished manuscript. [Haddock_Tagging_Sept_2008_v-1.doc]

Michael C. Palmer, Paul Nitschke, Susan Wigley, and Paul Rago (2014). Estimation of haddock bycatch in the northeast United States midwater trawl fishery. NEFSC unpublished manuscript. [Haddock_bycatch_herring_fishery_v3_20140618.pdf].

Tim Miller and Mike Palmer (2014). Estimates of mortality and migration from Gulf of Maine and Georges Bank haddock tag-recovery data, 2005-2010. SARC 59 Working Paper. [Haddock_tagging_miller_SARC59_WP.pdf]

New England Fishery Management Council (2013a). GB haddock stock spillover to GOM haddock stock. Groundfish PDT Memo, August 8, 2013. [PDT_to_SSC_haddock_spillover.pdf]

New England Fishery Management Council (2013b). Spillover of haddock between the Georges Bank and Gulf of Maine stocks. 12. SSC - September 24-26, 2013 – M. [SSC_Report_HaddockSpillover_Sep2013_Final.pdf]

NEFSC (2010). VPA/ADAPT Version 3.0 Reference Manual. NEFSC. [VPA_v3_Reference_Manual.pdf]

NEFSC (2014.) SAW/SARC 59 Benchmark assessment of the Gulf of Maine haddock (Melanogrammus aeglefinus) stock. Gulf of Maine Haddock WP A-1, SAW 59, June 27, 2014. [GOM_haddock WP A1 DRAFT assessment report AOB_0627.pdf]

NEFSC (2014). Gulf Of Maine Assessment Summary Report For 2014. Gulf of Maine Haddock WP A-2, SAW 59, June 27, 2014. [GOM_haddock_WPA2_SAW59_DRAFTAssessSummReprt_AOB_v2.docx]

Atlantic Sea Scallops

NEFSC (2004). Essential Fish Habitat Source Document: Sea Scallop, Placopecten magellanicus, Life History and Habitat Characteristics (Second Edition). NOAA Technical Memorandum NMFS-NE-189. [Scallop_EFH_2004_HartChute.pdf].

Daniel R. Hennen and Deborah R. Hart (2012). Shell height-to-weight relationships for Atlantic sea scallops (Placopecten magellanicus) in offshore US waters. Journal of Shellfish Research, Vol. 31, No. 4, 1133–1144. [Scallop_height_weight_2012_HennenHart.pdf].

Burton V. Shank, Deborah R. Hart, Kevin D. Friedland. (2012). Post-settlement predation by sea stars and crabs on the sea scallop in the Mid-Atlantic Bight. Mar Ecol Prog Ser 468: 161–177. [Scallop_predators_2012_Shank_et al.pdf].

Deborah R. Hart, Larry D. Jacobson, Jiashen Tang (2013). To split or not to split: Assessment of Georges Bank sea scallops in the presence of marine protected areas. Fisheries Research, 144, 74–83. [Scallop_split_area_2-13_Hart et al.pdf]

Deborah R. Hart (2013). Quantifying the tradeoff between precaution and yield in fishery reference points. ICES Journal of Marine Science; doi:10.1093/icesjms/fss204. [Scallop_tradeoffs_2013_Hart.pdf]

Hart, D. R., and Chute, A. S. (2009). Estimating von Bertalanffy growth parameters from growth increment data using a linear mixed-effects model, with an application to the sea scallop Placopecten magellanicus. ICES Journal of Marine Science, 66: 2165–2175. [Scallop_von bert_2009_HartChute.pdf]

NEFSC (2010). 50th Stock Assessment Workshop (SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 10-17. [Sea Scallop Stock Assessment CRD 10-17 50th SAW 2010.pdf]

NEFSC (2014.) SAW SARC 59 Assessment of the Atlantic Sea Scallop for 2014. Atlantic Sea Scallop, WP B-1, SAW 59, July 2, 2014. [Sea Scallop WP B1 DRAFT Assessment Report_70314.pdf].

NEFSC (2014). Sea Scallop Assessment Summary Report For 2014. Sea Scallop WP B-2, SAW 59, July 11, 2014. [SeaScallopWPB2Assess_Summ_SAW59-drh2_AOB0710_v2.docx]

Appendix 2: Statement of Work for Dr Coby Needle

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

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

BACKGROUND

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

SCOPE

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

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

OBJECTIVES

The SARC review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the SSC of the New England or Mid-Atlantic Fishery Management Council. The SARC panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Duties of reviewers are explained below in the "**Requirements for CIE Reviewers**", in the "**Charge to the SARC Panel**" and in the "**Statement of Tasks**". The draft stock assessment Terms of Reference

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

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

PERIOD OF PERFORMANCE

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

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

PLACE OF PERFORMANCE AND TRAVEL

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

STATEMENT OF TASKS

Charge to SARC panel: During the SARC meeting, the panel is to determine and write down whether each stock assessment Term of Reference (ToR) of the SAW (see **Annex 2**) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. **If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.** Where possible, the SARC chair shall identify or facilitate agreement among the reviewers for each stock assessment Term of Reference of the SAW.

If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.

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

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

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (or by email if necessary) the requested information (e.g., 1.name [first, middle, and last], 2.contact information, 3.gender, 4.country of birth, 5.country of citizenship, 6.country of permanent residence, 7.whether there is dual citizenship, 8.country of current residence, 9.birth date [mo, day, year], 10.passport number, 11.country of passport) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/.

Pre-review Background Documents and Working Papers: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the SARC chair and CIE reviewers the necessary background information and reports (i.e., working papers) for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

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

(SARC chair)
Act as chairperson, where duties include control of the meeting, coordination of presentations and discussions, making sure all stock assessment Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For each assessment, review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

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

(SARC CIE reviewers)
For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a reviewer's point of view, determine whether each stock assessment Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a reviewer considers any existing Biological Reference Point or BRP proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist. Review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

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

Tasks after the panel review meeting:

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

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

During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent CIE Report produced by each reviewer.

The Independent CIE Report can also be used to provide greater detail than the SARC Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

SARC chair:
The SARC chair shall prepare a document summarizing the background of the work to be conducted as part of the SARC process and summarizing whether the process was adequate to complete the stock assessment Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report (see **Annex 4**).

SARC chair and CIE reviewers:
The SARC Chair, with the assistance from the CIE reviewers, will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each stock assessment Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this SARC Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion.

The SARC Summary Report (please see **Annex 4** for information on contents) should address whether each stock assessment Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

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

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

DELIVERY

Each reviewer shall complete an independent peer review report in accordance with the SoW. Each reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each reviewer shall complete the independent peer review addressing each stock assessment ToR listed in **Annex 2**.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at the Woods Hole, Massachusetts scheduled during the tentative dates of July 15-18, 2014.
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than August 1, 2014, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment ToR in **Annex 2**.

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

June 10, 2014	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
July 1, 2014	NMFS Project Contact will attempt to provide reviewers the pre-review documents
July 15-18, 2014	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
July 18, 2014	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
August 1, 2014	Reviewers submit draft independent peer review reports to the contractor's technical team for independent review
August 1, 2014	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
August 8, 2014	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
August 15, 2014	Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements
August 22, 2014	The COR distributes the final reports to the NMFS Project Contact and regional Center Director

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

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

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

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

Acceptance of Deliverables: The deliverables shall be the final peer review report from each reviewer that satisfies the requirements and terms of reference of this SoW. The contract shall be successfully completed upon the acceptance of the contract deliverables by the COR based on three performance standards:

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

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

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

Support Personnel:

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

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

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

Key Personnel:

Dr. James Weinberg, NEFSC SAW Chairman, NMFS Project Contact
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543

James.Weinberg@noaa.gov (Phone: 508-495-2352) (FAX: 508-495-2230)

Dr. William Karp, NEFSC Science Director
Northeast Fisheries Science Center
166 Water St., Woods Hole, MA 02543
william.karp@noaa.gov Phone: 508-495-2233

Annex 1: Format and Contents of Independent Peer Review Report

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

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

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

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

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

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

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

3. The reviewer report shall include the following appendices:

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

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

A. Gulf of Maine (GOM) haddock

1. Estimate catch from all sources including landings and discards. Include recreational discards, as appropriate. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data. Investigate the utility of commercial or recreational LPUE as a measure of relative abundance.
2. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, etc.). If available, consider whether tagging information could be used in estimation of stock size or exploitation rate. Characterize the uncertainty and any bias in these sources of data.
3. Evaluate the hypothesis that haddock migration from Georges Bank influences dynamics of GOM stock. Consider role of potential causal factors such as density dependence and environmental conditions.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.
5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).
   a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
   b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) (see Appendix to SAW TORs for definitions).
   a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, migration from Georges Bank).
   b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
   c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

B. Sea scallop

1. Estimate removals from all sources including landings, discards, incidental mortality, and natural mortality. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these assumptions and sources of data. If possible using sensitivity analyses, consider the potential effects that changes in fishing gear, fishing behavior, and management may have on the assumptions.
2. Present the survey data being used in the assessment (e.g., regional indices of relative or absolute abundance, recruitment, size data, etc.). Characterize the uncertainty and any bias in these sources of data.
3. Investigate the role of environmental and ecological factors in determining recruitment success. If possible, integrate the results into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass for the time series, and estimate their uncertainty. Report these elements for both the combined resource and by sub-region. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.
5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model or model formulation developed for this peer review.
   a. Update the existing model with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
   b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

7. Evaluate the realism of stock and catch projections and compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level).
   a. Provide numerical annual projections (through 2016). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
   b. Comment on the realism of the projections. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
   c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

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

Appendix to the SAW Assessment TORs:

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

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

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

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

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

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

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce MSY and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Rules of Engagement among members of a SAW Assessment Working Group:

Anyone participating in SAW assessment working group meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

One model or alternative models:

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

Annex 3: Draft Agenda

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

July 15-18, 2014

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

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

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, July 15

10 – 10:30 AM

Welcome	**James Weinberg**, SAW Chair		
Introduction	**SARC Chair**		**TBD**
Agenda			
Conduct of Meeting			

10:30 – 12:30 PM Assessment Presentation (Stock A.)

TBD TBD **TBD**

12:30 – 1:30 PM Lunch

1:30 – 3:30 PM Assesssment Presentation (Stock A.)

TBD TBD **TBD**

3:30 – 3:45 PM Break

3:45 – 5:45 PM SARC Discussion w/ Presenters (Stock A.)

SARC Chair **TBD**

5:45 – 6 PM Public Comments

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

<u>Wednesday, July 16</u>

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

<u>Thursday, July 17</u>

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

<u>Friday, July 18</u>

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

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

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

Annex 4: Contents of SARC Summary Report

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

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

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

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

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

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

Appendix 3: Panel Membership or other pertinent information from the panel review meeting

CIE panel membership: Yiota Apostolaki, Vivian Haist, Coby Needle.

Appendix 4: Figures from additional analyses carried out for GoM haddock

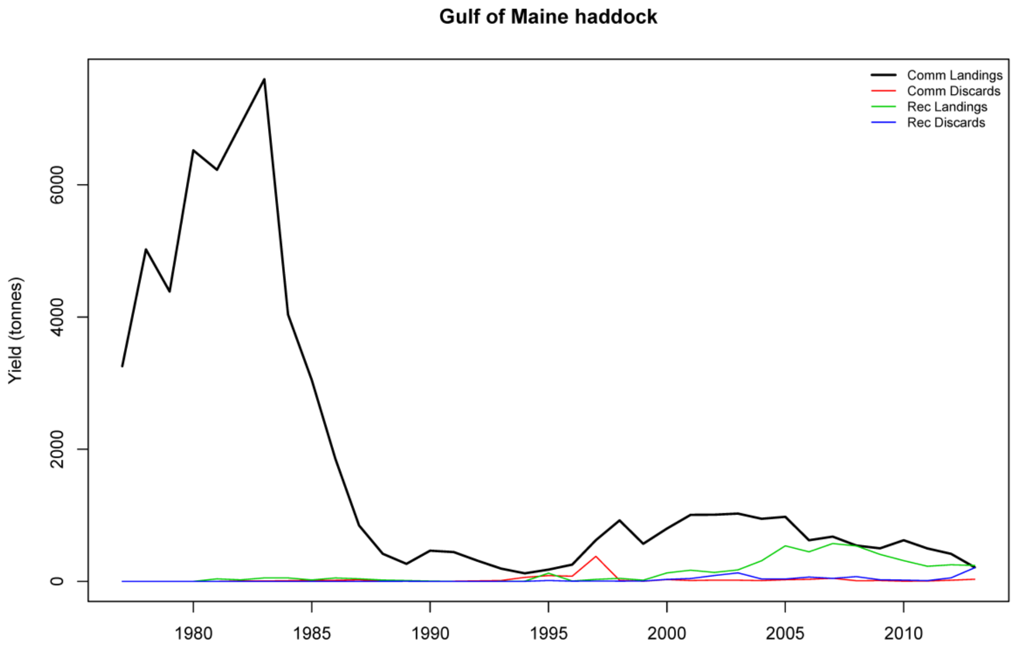


Figure CLN.A1. Gulf of Maine haddock. Fishery yield (tonnes) for the four catch components: commercial landings and discards, and recreational landings and discards.

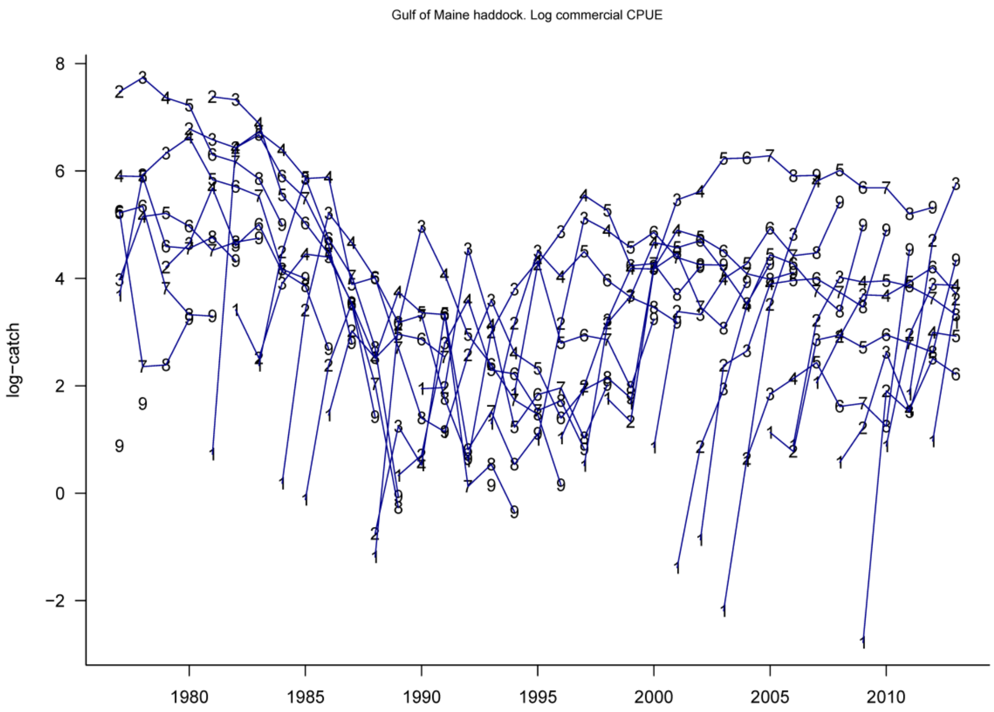


Figure CLN.A2. Gulf of Maine haddock. Log catch curves by cohort for total catch-at-age data.

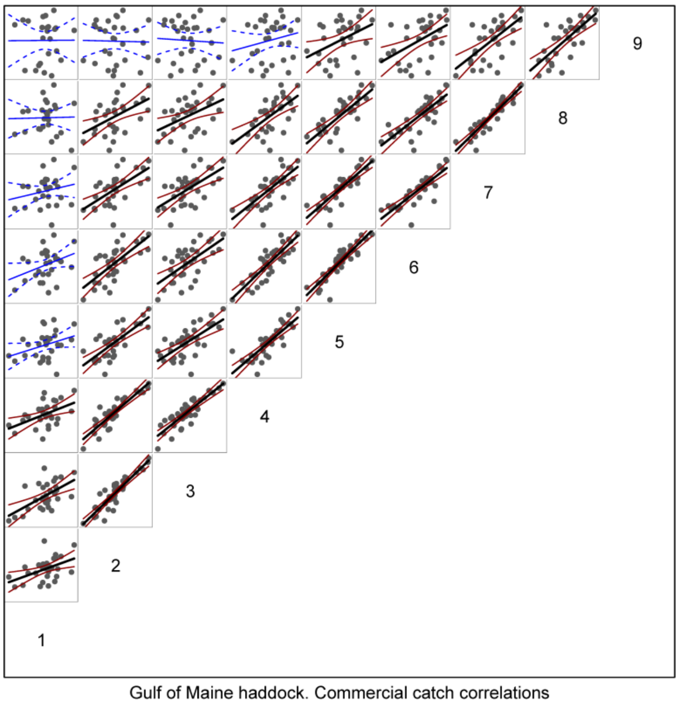


Figure CLN.A3. Gulf of Maine haddock. Correlations in the total catch-at-age matrix, comparing estimates at different ages for the same year-classes (cohorts). In each plot, the straight line is a normal linear model fit: a thick line (and black points) represents a significant ($p < 0.05$) regression, while a thin line (and blue points) is not significant. Approximate 95% confidence intervals for each fit are also shown.

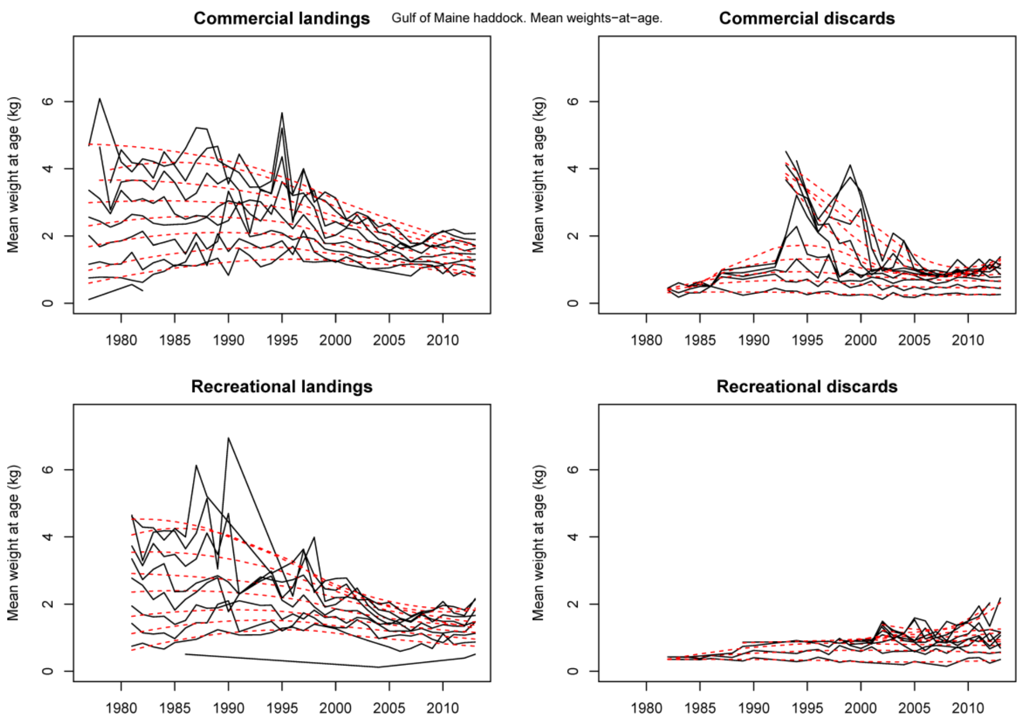


Figure CLN.A4. Gulf of Maine haddock. Mean weights-at-age (kg) by catch component. Red dotted lines give loess smoothers through each time-series of mean weights-at-age.

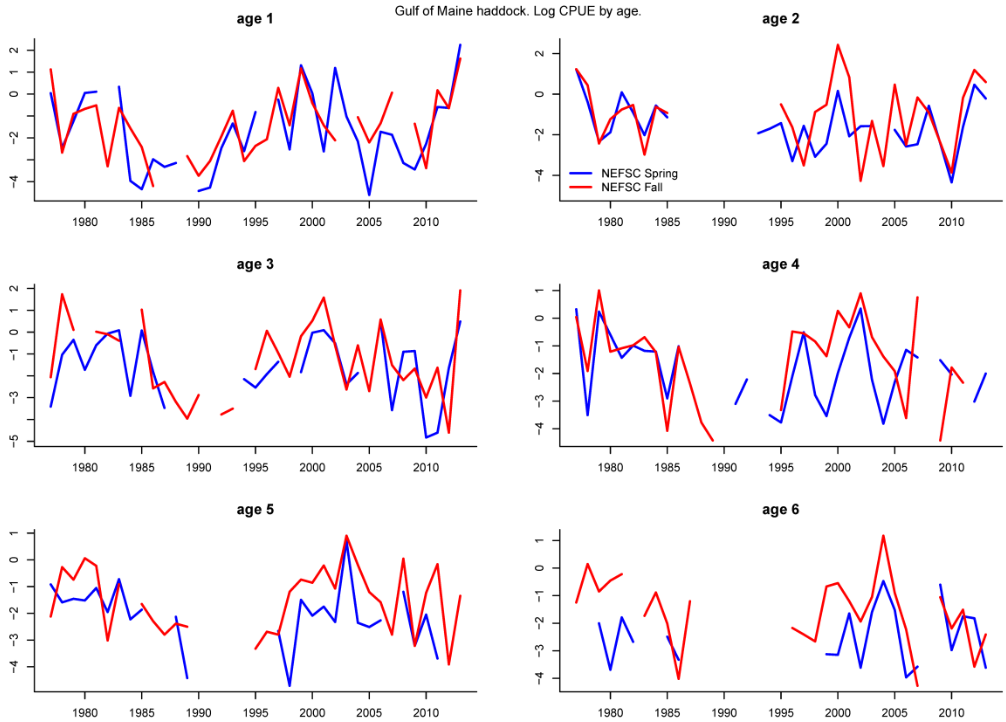


Figure CLN.A5. Gulf of Maine haddock. Survey log CPUE (catch per unit effort) at age.

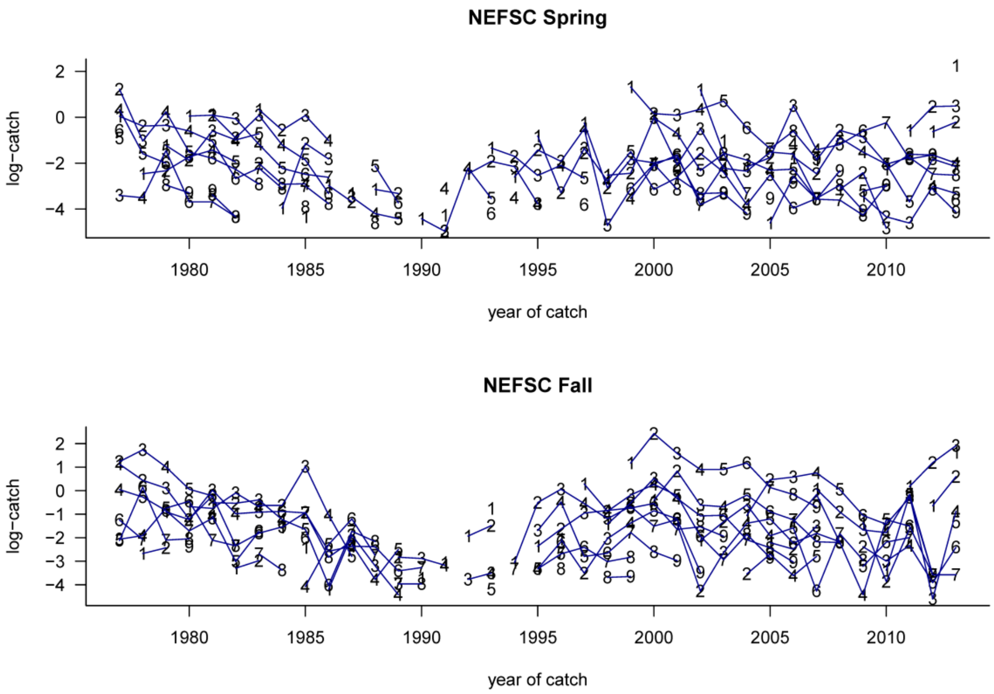
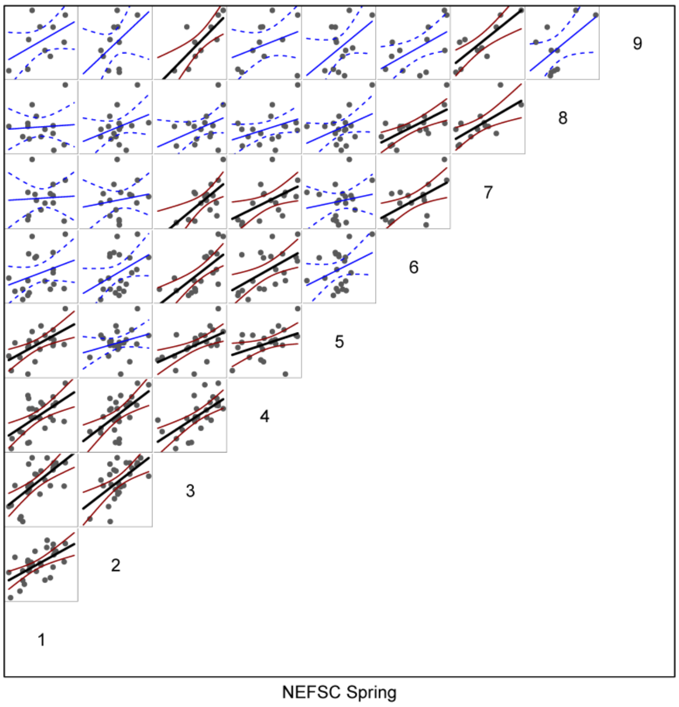


Figure CLN.A6. Gulf of Maine haddock. Survey log catch curves by cohort.



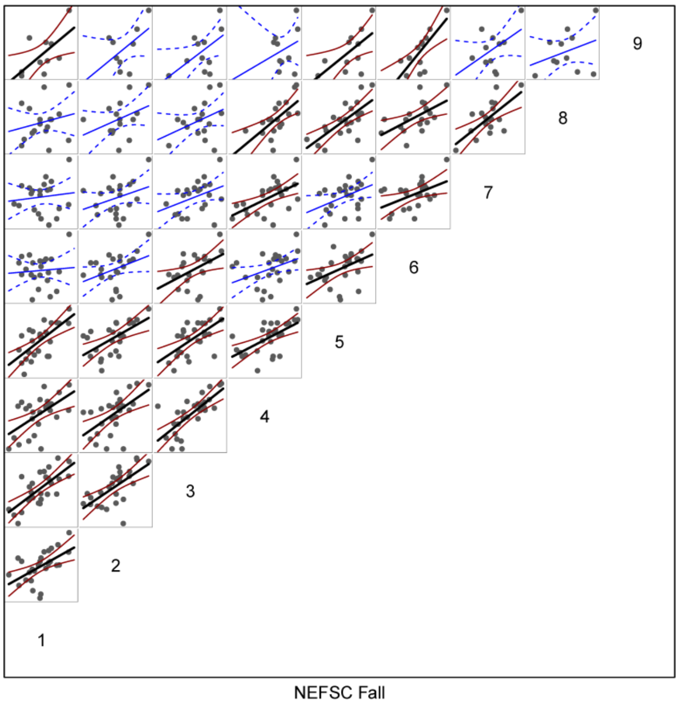


Figure CLN.A7. Gulf of Maine haddock. Correlations in the NEMFS Spring (upper) and Fall (lower) survey catch-at-age matrix, comparing estimates at different ages for the same year-classes (cohorts). In each plot, the straight line is a normal linear model fit: a thick line (and black points) represents a significant ($p < 0.05$) regression, while a thin line (and blue points) is not significant. Approximate 95% confidence intervals for each fit are also shown.

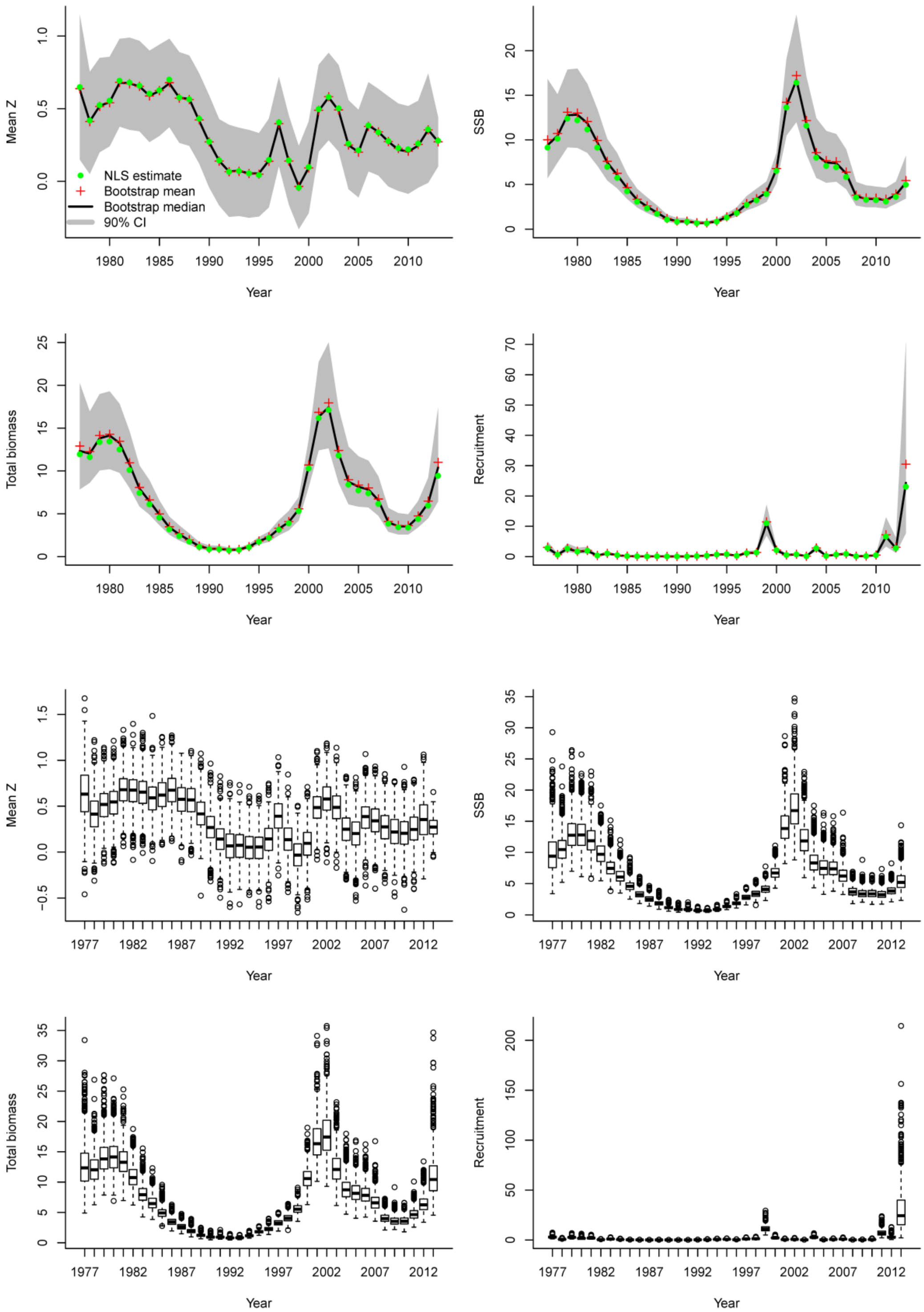


Figure CLN.A8. Gulf of Maine haddock. Summary plots from an exploratory SURBAR assessment, using both available surveys (NEMFS Spring and Fall). Mean mortality Z (ages 6 to 8), relative spawning stock biomass (SSB), relative total biomas (TSB), and relative recruitment. Upper plots: shaded grey areas correspond to the 90% CI. Green points give the model estimates, while red crosses and black lines give (respectively) the mean and median values from the uncertainty estimation bootstrap. Lower plots: box-and-whisker summaries of estimate distribution.

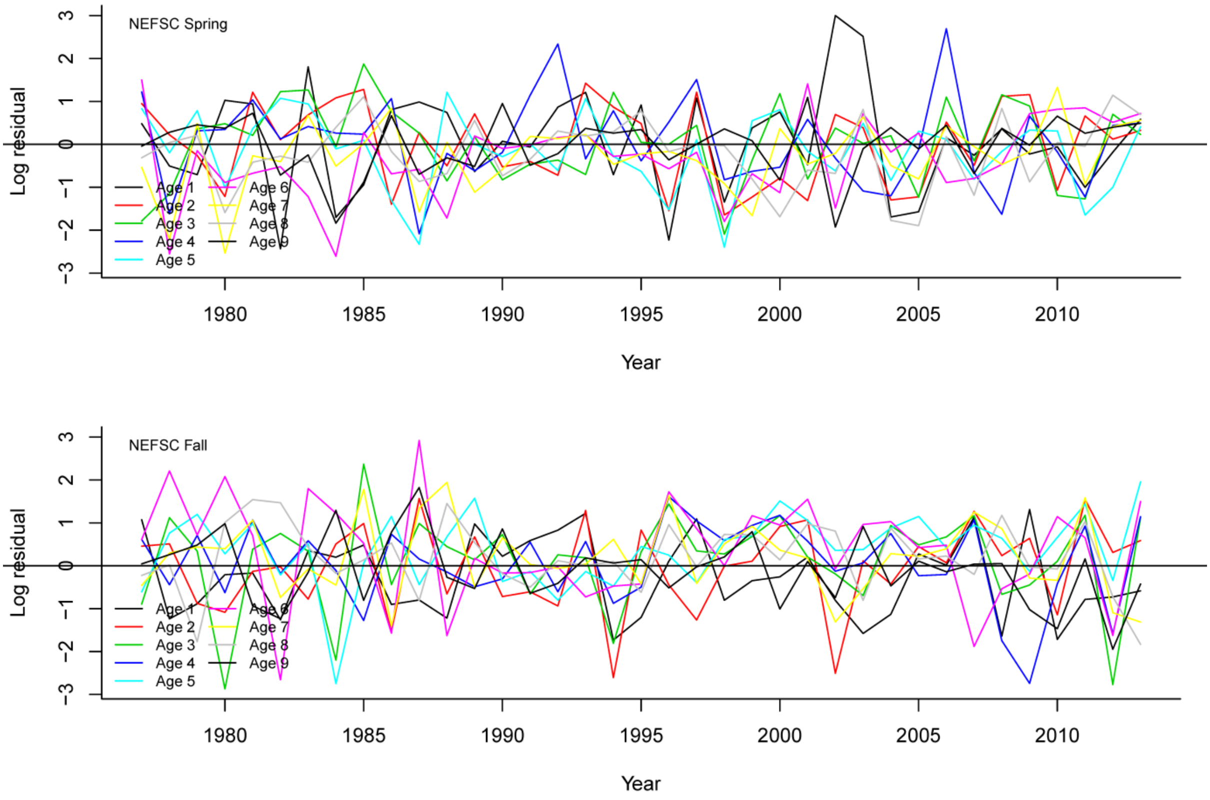


Figure CLN.A9. Gulf of Maine haddock. Log residuals by age from an exploratory SURBAR assessment, using both available surveys (NEMFS Spring and Fall).

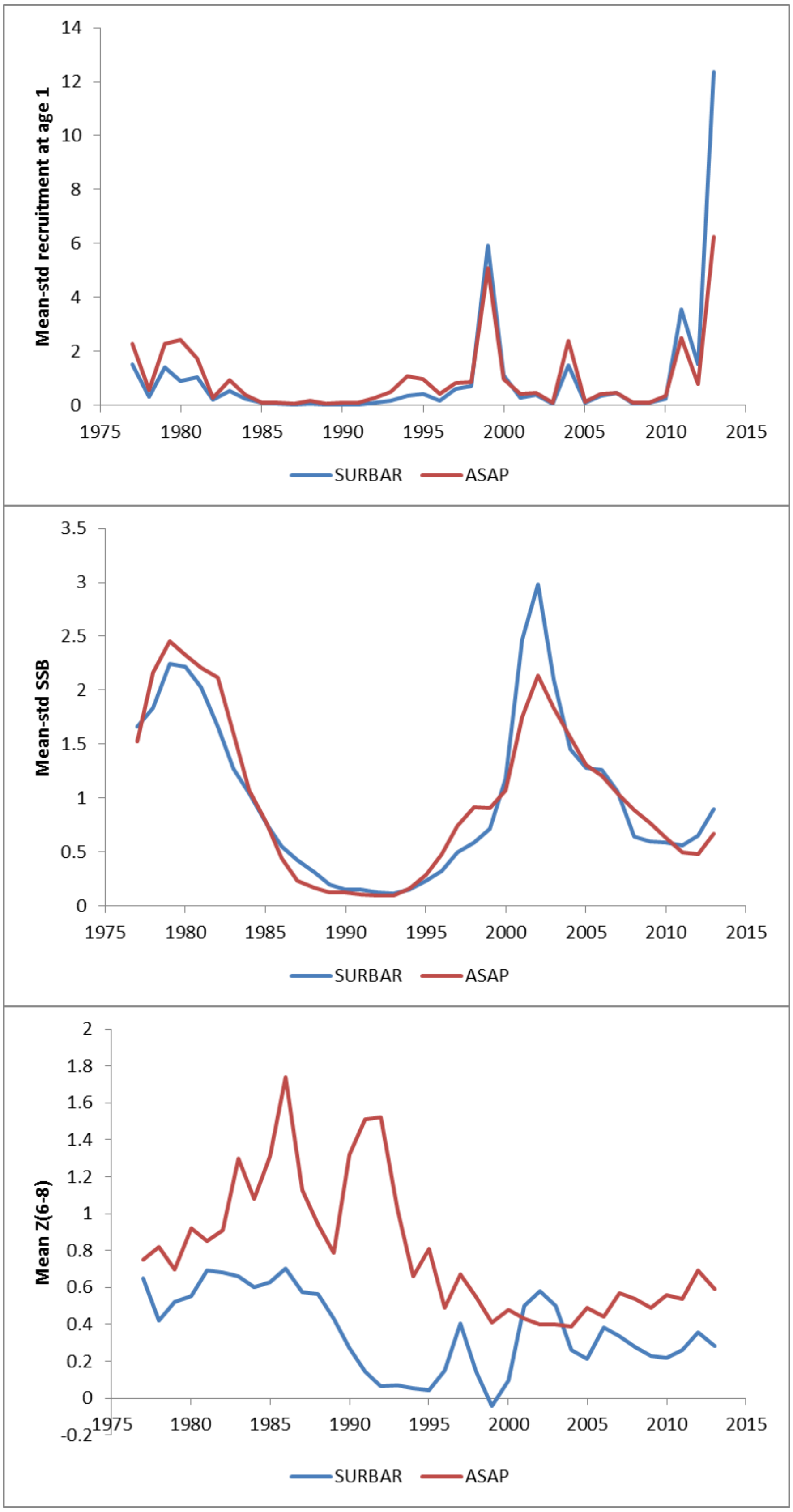


Figure CLN.A10. Gulf of Maine haddock. Comparisons of mean Z (= mean F – 0.2), relative SSB and relative recruitment from a SURBAR analysis and the final ASAP assessment.