Independent Peer Review Report for the 63rd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for ocean quahog.

Martin Cryer

March 2017

Executive Summary

The benchmark stock assessment for ocean quahog implemented in the Stock Synthesis III software package by NEFSC scientists was reviewed by a panel of four independent reviewers between 21 and 23 February 2017. I consider the assessment to be a credible basis for developing fishery management advice on the fishery. The assessment team and their stock assessment workgroup met all their terms of reference (ToR) although areas for future work and improvement were identified. The stock assessment entailed some uncertainty about absolute biomass and almost all information on biomass scale comes from Bayesian priors on survey dredge efficiency that, in turn, came from field depletion estimates. The assessment team therefore focused on relative biomass and trends and these were assessed with much more certainty, and I think this is a good approach. New biological reference points (BRPs) based on relative trends and management strategy evaluation were also developed, and I believe these are superior to the previous pragmatic BRPs. Base model results, a wide range of sensitivity trials, and supplementary analyses outside the model all suggest that the stock is neither overfished nor experiencing overfishing. Projections, using a wide variety of assumed biomass, catch, mortality, and recruitment scenarios, including zero recruitment, suggest that the stock will not become overfished nor experience overfishing in the next 50 years.

I believe that (even) more diagnostics should be presented for a benchmark assessment, that the model development phase, especially the formulation of Bayesian priors, should be more fully described, and that Markov chain Monte Carlo (MCMC) runs should be used routinely for Bayesian models to estimate posteriors (uncertainty estimates), diagnose problems, and provide for stochastic projections. But this is a good, robust stock assessment that has dealt with a number of difficult problems including the lack of contrast in most input data and very low rate of fishing mortality. It provides a scientifically credible basis for developing fishery management advice.

Background

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal meeting of a Panel of stock assessment experts charged with the peer-review of selected stock assessments and models. This report is an independent peer review of the benchmark stock assessment for ocean quahog presented at the 63rd SARC meeting held at Woods Hole, Massachusetts, 21–23 February, 2017. The SARC panel comprised a chairman, Dr Ed Houde, and three reviewers appointed by the Center for Independent Experts (CIE), Drs Mike Bell, Anthony Hart, and me. This report constitutes my own personal review and perspective of the assessment. It is designed to be read as a stand-alone document, but there are strong overlaps with the Summary Report developed collaboratively with the other members of the review Panel. I agree with all statements made in the Summary Report, and some of the text may be very similar, but this report includes further detail on parts of the stock assessment where I have particular interest or knowledge.

Role in the Review

Most of the necessary background papers for the surf clam assessment were made available well before the meeting on 3 February 2017, and the stock assessment report was provided a few days later on 7 February. The important draft stock assessment summary report was provided on 9 February. I read these documents before arriving in Falmouth/Woods Hole but, focusing primarily on the large stock assessment document itself, annotating the electronic documents as I worked through. Because I also conducted previous reviews of the Atlantic surf clam fishery at SARC-56 in 2013 and SARC-61 in 2016, my knowledge of these allied offshore dredge fisheries was already reasonable and this helped with my interpretation of the material.

After meeting briefly with Drs Jim Weinberg and Russ Brown in the early morning of 21 February, the Review Panel worked collaboratively on the Terms of Reference for the Review with the stock assessment team throughout 21 and 22 February. The assessment team agreed to undertake selected additional analyses during the evening of 21 February, and these were reported back on the mornings of 22 and 23 February. Electronic copies of the presentations, rapporteur notes, and additional analyses were made available rapidly, greatly assisting progress. At all times, the review was conducted in a collegial and open manner by all participants, and with good humour. I appreciate this very much and it makes the job of reviewing complex work so much easier.

The Panel met again on 23 February to agree on consensus points and to start drafting the Summary Report. There was strong agreement among the Panel members on almost all aspects of the review and drafting most of the Review Summary Report was straightforward. The Panel Chair allocated the Terms of Reference (ToR) among the Panel members and I tackled ToR5 (stock assessment model) and ToR6 (biological reference points). Toward the end of 23 February, the Chair collated all contributions into a single document. Versions were circulated and comments addressed until an agreed final Summary Report was submitted by the Chair to Dr Jim Weinberg for a check for factual correctness on or about 13 March. I submitted this individual review on 9 March 2017.

Findings as to whether the work provides a scientifically credible basis for developing fishery management advice

I consider the stock assessment for ocean quahog provides a scientifically credible basis for developing fishery management advice on the fishery. I offer detailed comments in relation to each Term of Reference below.

Findings by Term of Reference (ToR)

> **ToR 1.** ***Estimate catch from all sources including landings and discards. Map the spatial and temporal distribution of landings, discards, and fishing effort, as appropriate. Characterize the uncertainty in these sources of data.***

I consider this ToR to have been met.

Landings, estimates of fishing effort, and locations of catches are based on logbook records and are presumed to be accurate because of the procedures in place in the fishery (including cage tagging) and the lack of substantial incentives to misreport. Logbook records are readily available for use in stock assessments. Landings and effort have declined over the past three decades, it is thought because of the limited demand for ocean quahogs. Catch rates have declined in the southern sub-regions (Delmarva and New Jersey) and the fishery is now concentrated off Long Island where 70–80% of landings occurred between 2005 and 2015.

The assessment assumes that discarding is zero. This is apparently true for the directed fishery although limited observer data from the surf clam fishery indicate that, in 2015 and 2016, about 1500 lb of ocean quahog was discarded per 100,000 lbs of surf clam landed (from areas where the two species overlap). This is probably an increase on historical levels of discarding in the surf clam fishery as the depth distribution of Atlantic surf clams has shifted somewhat deeper and into the depth inhabited by ocean quahogs. It was not clear from the stock assessment report whether discards from the surf clam fishery were included in the model but, either way, I agree with the Panel's conclusion that discards are minor and not a concern for the ocean quahog stock assessment. Gear-related incidental mortality is assumed to be 5% and landings are simply scaled by this amount in the model. At the very low level of fishing mortality prevalent in this fishery, incidental mortality is most unlikely to have significant consequences for the ocean quahog assessment unless there are gross trends in the rate that are not accounted for.

A large number of maps showing the distribution of fishing effort and catch were included in the stock assessment report and these show convincingly that the distribution of catch has shifted north in the past three decades. Fishery-independent surveys and LPUE both indicate that this shift is probably a response to declining abundance in the south. In turn, these changes in the distribution of abundance are probably a response to ocean warming and can be expected to continue over time.

> *ToR 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.). Use logbook data to investigate regional changes in LPUE, catch and effort. Characterize the uncertainty and any bias in these sources of data. Evaluate the spatial coverage, precision, and accuracy of the new clam survey.*

I consider this ToR to have been met.

Fishery-independent surveys using hydraulic dredges have been conducted across the surf clam and ocean quahog grounds off the east coast of the US for many years. Information from surveys since 1983 was presented in the stock assessment report and used in the stock assessment model, but previous surveys were not used because they were not considered sufficiently comparable with more recent surveys. It seems unfortunate to have to discard such potentially useful historical data (because they could be very informative) and I believe it might be worth reconsidering that choice for the next assessment. This could be especially pertinent if a more detailed examination of the factors driving catch on a tow were to be undertaken, potentially allowing pre-1982 surveys to be "corrected" and made more comparable with more recent surveys.

Details of the survey methodology were given in the stock assessment report, including the change from the small research dredge to a much larger and more efficient modified commercial dredge in 2012. A series of depletion experiments to estimate dredge efficiency is described, including the sophisticated Patch model used to analyse the experimental data. A series of models to estimate the selectivity characteristics of the dredges is also described. Clearly, an impressive amount of work has gone into conducting and analyzing the surveys.

Survey results are presented as estimated numbers per unit area, corrected for selectivity of smaller clams but not for overall efficiency. Given this choice of presentation, the pre- and post-2012 survey series do not line up nicely on the graphs and are not really comparable. I understand that this is the way the data were presented to the model (with separate q parameters estimated in the model) but it would not have been onerous to have additional plots in the report showing the corrected (and more comparable) estimates.

I noted some of the same issues with the survey design and implementation that I noted for the surf clam assessment in SARC61. The report noted that "a few" nearly-random tows were included in some of the earlier surveys. Discussion during this and previous reviews revealed that these tows were targeted on areas that randomly-allocated tows "missed" but were important in some way. Almost inevitably, these areas included some where quahog were expected to be found, potentially in greater than average abundance, and this procedure will introduce bias. Thus, generating and using "nearly random" tows in the analysis is to be discouraged. It would have been better to re-draw another set of random stations or, if the problem was persistent, to stratify such that all areas are allocated at least some stations. The number of such stations was said to be small, so the impact on the index will probably be small unless very large numbers of quahog were caught in several of them. The treatment of un-fishable stations in the analysis was not clear. There are at least two unbiased methods of dealing with stations where it is considered by the skipper that the ground is not suitable for fishing or too dangerous. Because the proportion of such stations was quite small, the impact will also be small and is likely to be problematic only if the approach changed between years (introducing bias into estimated trends). The ability to estimate actual tow length changed substantially during the survey time series. There is likely to be some bias associated with the approximate method used in earlier years, but it appears (to me) that this was not corrected for in the analysis even though data are available to support such analysis. This could introduce or obscure a trend in the time series. It would have been nice to see additional analyses assessing the impact of these issues but, given the small likely effect of each and the relatively loose fit of the model to the abundance indices, it was not a priority for the review meeting. I think these issues should be dealt with before the next benchmark assessment, at a minimum through sensitivity analyses.

In some surveys, not all strata were sampled and tows were "borrowed" from temporally-adjacent surveys to complete the analyses. I agree with the Panel that this is not an ideal approach because it gives more weight to data from the borrowed tows than to the others. Potentially a model-based approach could be applied and I suggest that one is considered for development. However, because the expected change in abundance between surveys even two or three years apart is small for such a long-lived animal experiencing low fishing mortality, and given the relatively loose fit of the model to the abundance indices, this was not a priority for the meeting either.

It was made evident to the Panel during the review that the sampling design for the ocean quahog (and Atlantic surf clam) surveys was sub-optimal because the combined surveys were not focused particularly on either species. I agree with the WG's suggestion that substantial improvements in precision and statistical efficiency are possible. I suggest that, when analyses and/or simulations are done to redesign the surveys, their end-use in stock assessment models is considered as well as the predicted precision of the indices by year. Options could include a very focused and precise survey estimate every three or four years or more frequent but less precise estimates. These will have different performance in the stock assessment model in their ability to track biomass and provide for accurate projections, and the benefits of infrequent but precise surveys will obviously

accrue more slowly. It would probably be beneficial to consider the surveys and stock assessments for ocean quahog and Atlantic surf clam in a combined analysis wherein the allocation of resources for surveys and benchmark stock assessments between the two species is explicitly examined. The near-unfished status, low fishing mortality, and great longevity of ocean quahog all suggest that fewer surveys and more infrequent stock assessments are needed for this species than for Atlantic surf clam.

A simulation study could also consider different approaches to incorporating factors that influence dredge performance and efficiency such as location, depth, speed, tow duration, and dredge characteristics (e.g., dredge width, selectivity, differential pressure, angle of tilt). CIE reviewer Dr Anthony Hart suggested that these factors could be used to model the survey catches purely as an abundance index, potentially as an index of absolute abundance (or biomass density) across all surveys (including those pre-1982). This is a very interesting idea, and quite different from the current model structure using separate estimates of q for the two survey time series, but it would need careful consideration, especially in regard to the relationship between dredge efficiency estimates and model catchabilities. It might turn out to be necessary to retain an estimated q in the model to account for quahogs not vulnerable to the dredge surveys. I suspect there would be very little information to inform the estimation of such a q given the low contrast of the available biomass indices.

The problem comes from the ability of quahog to bury very deeply in the sediment at times, potentially making them invulnerable to even the heavy hydraulic dredge used for surveys since 2012. If a portion of the population cannot be taken in survey dredges, then experimentally-determined catchability is not precisely equivalent to q in the model, and this disparity becomes even more problematic if the depth of quahog burial increases with size (or age). To be clear, experimentally-determined catchability is the proportion of quahog within the sample-able sediment depth that can be caught with a single pass of the dredge, whereas q in the model is simply the proportion of quahog in an area that can be caught with a single pass of the dredge (whatever their depth of burial). The Panel heard about some details of the depletion experiments during the review that led me to think the issue is even more complicated. In particular, some present at the review thought that each pass of a dredge over a particular area of seabed was able to dig or slurry somewhat deeper into the sediment. If this is true, it is likely there would be significant violations of the assumptions of the Patch model. Worse, it would also lead to a biased correction for dredge efficiency because a survey tow will always be the first pass of the dredge over a particular area and will never be able to dig as deep as subsequent passes in a depletion experiment.

I was surprised to see that quahogs were measured (or reported /analysed) during surveys and in port surveys of landings with a precision of 1 cm. This seems an odd choice given that such a broad size class can contain scores or hundreds of age classes. Especially now that more precise methods of ageing quahogs have become available, it would surely be better to measure with finer precision, potentially 1 mm. This is easily achievable with electronic measures and, in New Zealand, we routinely measure many thousands of scallops with 1 mm precision in each survey. Greater precision in length measurements from previous surveys would have provided more resolution of the shoulders in the length frequency distributions that drive the recruitment “spike” in the model and allow the model to better specify past recruitment. However, this would lead to a biologically reliable recruitment time series only if the growth model was reasonably specified.

Although the Panel and I have identified a number of issues with the surveys, and I have written a fair amount in this section, many of the issues should be considered areas for improvement rather than serious flaws. The overall approach is acceptable and provides a potentially powerful fishery-independent index for the stock assessment model. This could become extremely important if the biomass were to decline in the future because LPUE is unlikely to provide a reliable index of abundance (it is very likely to be hyper-stable given the patterns of fishing).

ToR 3. *Describe the relationship between habitat characteristics (e.g., benthic, pelagic, and climate), survey data, and ocean quahog distribution, and report on any changes in this relationship.*

I consider this ToR to have been met.

Substantial data on the distribution of ocean quahog has been collected over the many years of fishery-independent surveys and a concerted attempt was made to understand the relationship between ocean quahog abundance and habitat characteristics in support of this stock assessment. Temporal shifts in ocean quahog distribution had already been addressed as part of the recent Atlantic surfclam assessment (SARC61), but additional work to identify environmental factors potentially driving ocean quahog abundance was reported for this assessment. A regression tree approach was used to model the relationship between catches in survey tows between 1997 and 2011 and a range of environmental factors. A comprehensive set of environmental variables was compiled, including depth, temperature, a variety of indices of seabed topography, sediment type and chlorophyll concentration. Cross-validation, where models are developed using subsets of the data and tested against the excluded data, suggested that the models had reasonable predictive power. In addition, the separate models for Georges Bank and southern areas had very similar rankings of importance of environmental variables and gross spatial variables like latitude and longitude were not offered to the model. These attributes all suggest that factors genuinely associated with the abundance of ocean quahog have been identified through the modelling. Correlations among the predictor variables make inferences about causality difficult, however, and it should also be borne in mind that recruitment of small individuals to the population (a key driver of abundance in sessile taxa) is likely driven by environmental conditions at the time of settlement several – many decades ago, not by conditions observed when the survey samples were taken. Some factors may change over these time frames, especially temperature.

With the exception of a modest shift in median depth of ocean quahog in surveys in the New Jersey region, no changes in the distribution were noted between 1982 and 2011. Increases in the co-occurrence of Atlantic surf clam and ocean quahog appear to be due predominantly to shifts in the distribution of the surf clams into deeper water. I agree with the Panel that ocean quahog may be buffered from climatic and ocean acidification changes by living in deeper water and deeper in the sediment, and that any response that does occur is likely to be manifested through recruitment rather than survival processes.

ToR 4. *Evaluate age determination methods and available data for ocean quahog to potentially estimate growth, productivity, and recruitment. Review changes over time in biological parameters such as length, width, and condition.*

I consider this ToR to have been met.

The Panel was shown information that supports the idea that annual growth rings are detectable in this species and can be used to estimate the age of individuals. Estimates of age are valuable for the stock assessment but sufficient catch-at-age information has not yet been collected to fit the model to. Rather, the model is fitted to catch at length and survey length frequency distributions (see ToR 5, stock assessment modeling) and the age information is used to parameterize a growth curve. A number of alternative growth models were investigated during the development of the stock assessment and none was found completely satisfactory for the apparently indeterminate growth of this species. The non-asymptotic Tanaka growth model seemed most appropriate, but the fits to the growth curves of individual animals were still not very good, and the SS3 software does not currently implement this model (so a von-Bertalanffy model with wide variance was used). It is not entirely clear how this would affect the estimates of biomass, status, and yield, but it will certainly have had implications for the recruitment deviations estimated by the model. In

particular, the size-frequency and survey data abundances, coupled with the growth model, will have forced the SS3 model to estimate the large peak in recruitment in the late 1990s, and this peak may have looked quite different under a different growth model. The observed time series of recruitment deviations is therefore very uncertain, and must not be considered a definitive estimate of real recruitment patterns. This is important for the interpretation of any estimates of risk made using the stock projections from the model, especially if the stock were to be reduced to biomass near the threshold reference point. The effects of using different growth models on biological reference points were evaluated using management strategy evaluations (MSE) and this does give me some comfort that the impacts on the stock assessment model and estimates of stock status are unlikely to be gross. Given these uncertainties, I agree with the WG's conclusion that non-asymptotic growth is an area for further investigation. I also think it would be useful to have this and other options made available for model building in SS3, but that is not an issue for the assessment team as such, it is something for them to bring up with those who maintain and develop the SS3 package.

Morphometric data were analysed for year and regional effects. Minor differences in morphology were observed among regions and this is consistent with the validated age information from five individual quahogs that were seen to have different growth rates indicating potential differences in productivity among regions. No temporal changes were detected in the morphological metrics and, from this, it might be concluded that no major changes in the condition of ocean quahogs have occurred. However, the observed growth curves of the five aged animals appeared to have periods (of several years) of relatively rapid growth and periods of slower growth, suggesting that there may be temporal changes in productivity. It is not clear whether those changes simply occur in individuals at different times of their life or whether they are more synchronous across members of populations or areas (and therefore more relevant to stock assessment modeling).

I believe that at least some of the pattern evident in Figure 53 (purporting to suggest changes in growth / size-at-age over time) is artifactual. At first glance, the figure suggests that, in some areas, 80 mm quahogs with a recent birth year are, on average, younger than 80 mm quahogs with older birth years. However, a quahog with a birth year of 1990 cannot have an age exceeding 25 if it is collected or observed in 2015, whereas those with a birth year of 1900 can have an age >100 years if observed in 2015. I have illustrated this in my Figure 1 below wherein I identify with red triangles parts of the plots where it would have been impossible to make an observation in 2015 (drawn by eye). Some simulations would easily determine whether the observed patterns are entirely artifactual or whether some of the pattern remains after adjustment for this effect.

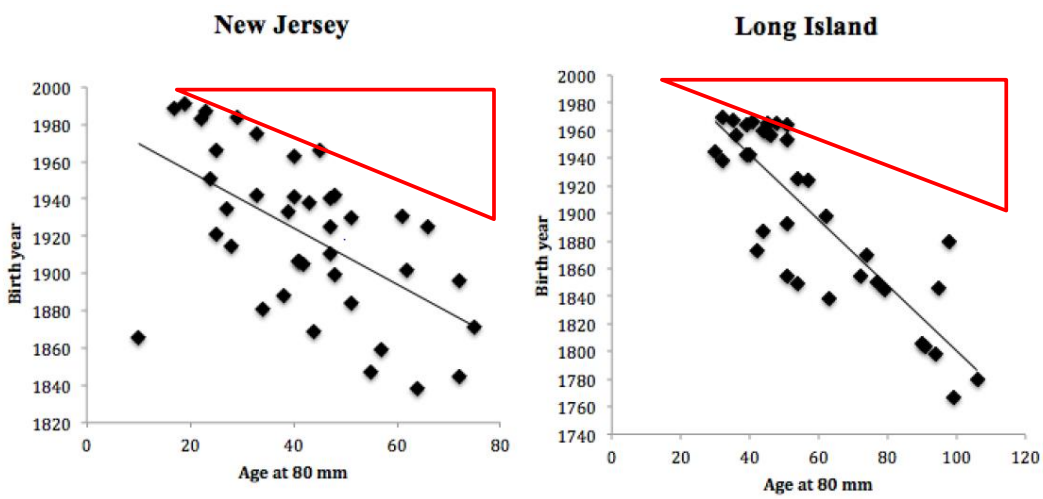

Figure 1 (after Figure 53 of the stock assessment report): Approximate areas of the plots for New Jersey and Long Island (red triangles) where it would not have been possible to make an observation of a given age at size 80 mm in 2015.

I agree with the Panel that all of these issues should be explored further as the ageing method develops.

> *ToR 5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR 4, as appropriate) and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.*

I consider this ToR to have been met.

The primary assessment approach used integrated statistical catch-at-age models implemented in the software package Stock Synthesis III (SS3). I believe the choice of SS3 was appropriate for this assessment, and the model represents a substantial improvement over the previous biomass dynamic models because it allowed the integration of biomass indices, size compositions, dredge efficiency estimates from depletion experiments, and other information. The model incorporated separate areas for Georges Bank and other areas (collectively called the southern area) and the results were combined within the model to infer stock status and provide management advice for the assumed stock. This is a good structure for the model and overcomes some of the difficulties encountered in combining the results from multiple models (a problem that occurred in the surf clam assessment, SARC 61). Modeling the population dynamics of a long-lived animal subject to very low fishing mortality will always be difficult, especially if absolute estimates of population size are required, but I believe the assessment team has done a creditable job of this assessment under difficult circumstances.

Biomass indices from two different fishery independent surveys were fitted in the model but no biomass indices from landings per unit of fishing effort, LPUE, were used formally in the assessment (although many plots and distribution maps were shown). I agree with the Panel that this was appropriate, especially because of the fishing behavior of the fleet. LPUE is almost certain to be hyper-stable and not a good index of abundance. The small proportion of the stock area that is fished each year (and cumulatively) also means that, even if LPUE were to be linearly related to local quahog abundance, it may not be a good index of stock abundance.

The base model accepted by the working group appeared to be appropriate, but the working group had clearly gone through a model-development process that led them to that base model and it is not clear what discussions, data, and diagnostics led them to make some of the important structural and constraining choices. Sometimes it is very helpful to understand that process, but only limited documentation on model development was provided to the Panel. I believe more should have been provided in the documentation to aid a reader's understanding of the rationale for the base model. This would include a much better description of the development of priors for the base model. The stock assessment report alludes to a process of first providing diffuse priors and "potentially" tightening some to bring the model closer to field estimates as model development continued. This does not sound like an ideal process and a better description would have helped me understand the choices and the need to constrain. The stock assessment report refers readers to its Table 20, but I did not find that table very informative on the matter, it simply describes the structure of the base model. Table 21 provides more information on which parameters were estimated with priors, but does not define those priors. I am sure I would be able to find the definitions in the code supplied but, given the influence of these priors in this model, I believe it is very important to define them precisely and describe their derivation in the report.

I also agree that more detail of the methods used should have been provided in the assessment report. Ideally, this would include the actual equations used by SS3, but even links or references to relevant equations or methods in the SS3 manual would help. This level of detail would probably be provided in an appendix, but equations used to estimate some key parameters could appear in

the body of the report. This is a common issue with complex stock assessments (and many other statistical analyses) conducted using integrated packages, but I believe the aim should be to provide a report with sufficient detail to enable a new analyst to replicate the assessment model with little input from the current analyst and potentially using different software.

Model fits to the length compositions were very good, often very good indeed, but the fits to the biomass indices were relatively poor. Generally, I would see this as a bit of a red flag and an indication that model weights should be explored further. However, in this case, I believe the assessment team has found a good balance. There is good reason to suppose that at least one of the survey data points (for the southern area in 1994, and maybe for Georges Bank in 1989) is problematic and probably biased by changes to survey gear performance in that year. In addition, the change between biomass estimates from one survey to the next (positive or negative) is often much larger than should be expected for such a long-lived animal. Because of these particular circumstances, I agree with the Panel that the relatively poor fit to the biomass indices in the base model was acceptable and appropriate; forcing the model to fit the biomass surveys much better would be essentially fitting to irrelevant noise in those indices.

Parameter estimates for this assessment were presented as those at the maximum of the posterior density (MPD) and their uncertainties were estimated from the model's Hessian matrix. This is an approximate method for a Bayesian model and it would have been better to estimate the Bayesian posterior distributions of estimated and derived parameters using Markov chain Monte Carlo (MCMC) methods. As well as providing better estimates of uncertainty (and correlation between parameters), MCMCs are particularly useful for diagnosing some problems in model convergence or stability, and in providing for stochastic projections. I accept that running MCMCs can be time-consuming, and model development is almost always conducted using MPD fits, but I believe running MCMC(s) should be the norm for a Bayesian assessment model. In New Zealand, it would be extremely unusual for a Bayesian stock assessment model to be accepted by a working group if MCMC chains had not been run and presented in some detail. I provided links to two examples of reports with good Bayesian diagnostics and treatment of MCMC chains to the assessment team shortly after the review meeting[1] and I hope they can adopt some of the diagnostics and presentation approaches used in those reports. If there are impediments stemming from software constraints or organizational culture to running MCMCs routinely, I strongly recommend that these be tackled as a priority.

Internal and historical retrospective analyses were conducted and showed no concerning patterns of retrospective behavior by the model. Toward the end of the review, one of the figures was revised and the new figure indicated lesser shifts in scale than the original, but had the same lack of retrospective trends over time. No serious retrospective behavior is indicated and the adjustment will have no consequential effects on the assessment outcome.

The scale of absolute abundance (and, hence, fishing mortality) was somewhat uncertain in the model, but trends in relative biomass and fishing mortality were much more certain. A wide range of sensitivity analyses were conducted by the assessment team and these demonstrate that the model-estimated trends and the working group's conclusions from the model were robust to many of the modelling choices. Likelihood profiling was used to identify conflicts among data sources, priors, and penalties, and to understand the key drivers of the fits and, in particular, biomass scale. Empirical, area swept biomass estimates with q=1 were broadly consistent with the biomass

[1] Francis, R.I.C.C.; Elliott, G.; Walker, K. (2015). Fisheries risks to the population viability of Gibson's wandering albatross *Diomedea gibsoni*. *New Zealand Aquatic Environment and Biodiversity Report No. 152*. 48 p. http://www.mpi.govt.nz/document-vault/7632

Starr, P.J.; Webber, D.N. (2016). The 2015 stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 5 and development of management procedures. *New Zealand Fisheries Assessment Report 2016/41*. 115 p. http://www.mpi.govt.nz/document-vault/13131

estimates from the model, giving me some confidence that the model-based estimates of biomass and fishing mortality were not grossly biased. I like this approach of testing and corroborating complex model outputs with simpler analyses that are much less likely to include mistakes or peculiar model artifacts, and I commend the assessment team and the working group for conducting them and including them in the material presented.

Even though I am reasonably persuaded that the biomass scale of the model is not badly biased, however, I think the working group's focus on trends and ratios, especially for assessing stock status, was appropriate for two main reasons. First, likelihood profiling showed that almost all of the information on biomass scale came from the priors on survey catchability and there seems to be reason to believe that the depletion estimates of catchability are not equivalent to catchability during the survey. In particular (see also ToR2), if some deeply-buried quahogs can never be caught during the first pass of a dredge but can be caught in a subsequent pass, which potentially digs deeper into the sediment modified by a previous pass, then q is not constant throughout the experiment (potentially invalidating the estimate) and any estimate of overall q is not equivalent to a survey q when only single passes are ever used. Second, as is commonly the case, sensitivity and retrospective analyses show that the model's estimates of trends in biomass and fishing mortality were much more stable than the estimates of absolute values.

Overall, I thought the assessment team did a good job under difficult circumstances and with somewhat limited data.

> *ToR 6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. 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, particularly as they relate to stock assumptions.*

I consider this ToR to have been met.

The existing limit BRPs (SARC 48) were described based on a finfish proxy for $F_{THRESHOLD}$ ($F_{45\%}$ for rockfish = 0.022) and expert opinion for $SSB_{THRESHOLD}$ (= $0.4*B_{1978}$). No estimate of F_{MSY} for ocean quahog is possible because the stock-recruit curve cannot be estimated, so a new F_{MSY} proxy was developed using a comprehensive management strategy evaluation (MSE) simulation that included a wide range of assumptions about the life history parameters and growth of quahog and the uncertainty associated with both stock assessments and management decision-making. Model based estimates of the un-fished biomass, B_0, are now possible using the SS3 model so these were applied in the new $SSB_{THRESHOLD}$ BRP (using SSB rather than fishable biomass in the old $B_{THRESHOLD}$ proxy). Despite uncertainties about biomass scale that are likely to continue in the model, the SS3 model will be able to estimate terminal biomass relative to the new $SSB_{THRESHOLD}$ quite reliably.

Although the values of the old and new limit BRPs are similar, I think the new limit BRPs are substantially better than the old ones because each has a more solid theoretical basis, has better estimates of uncertainty, and performance has been explored in MSE simulations across a wide range of possible states of nature.

During the review meeting, the Panel noted the closeness of the biomass target ($0.5*B_0$) and biomass threshold ($0.4*B_0$) and heard that US legislation meant that this could trigger a rebuilding plan quite soon after the biomass declined below its B_{MSY} target. This proximity of the target and threshold could theoretically be problematic and lead to unnecessarily resource-intensive management processes if biomass declines substantially below its current high level. My personal opinion is that this is best resolved by explicitly increasing the value of the biomass target, perhaps

to 0.6*B_0, recognizing the very low productivity and high potential vulnerability of the species to overfishing. On a related note, I agree with other members of the Panel that, although the biomass threshold of 0.4*B_0 might be considered conservative, it is defensible for an animal of such low productivity. I recommend these matters be considered by the relevant fishery management decision-making groups when BRPs are next reviewed for ocean quahog.

> *ToR 7. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to any new model or models developed for this peer review. (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-6).*

I consider this ToR to have been met.

The assessment team presented a variety of analyses and sensitivity runs that showed quite conclusively that the ocean quahog stock was not overfished nor was it experiencing overfishing at the scale of the stock and at the finer scale of northern vs southern area. Stock status against the existing BRPs was estimated using both existing (KLAMZ, biomass dynamic) and new (SS3 statistical catch-at-age) assessment models. Consistent with the ToR, stock status against the proposed new BRPs was estimated only using the SS3 model, including changing the unit of biomass from "fishable biomass" in the old model to SSB in the SS3 model. I agree with the other members of the Panel that these estimates of status are defensible and robust. In all cases, and across an impressive set of sensitivity runs, the models consistently show that the stock remains at a high proportion of the un-fished biomass and is subject to very low fishing mortality. The stock is neither overfished nor experiencing overfishing.

> *ToR 8. Develop approaches and apply them to conduct stock projections.*
>
> *a) Provide numerical annual projections (5–50 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level), including model estimated and other uncertainties. Consider cases using nominal as well as potential levels of uncertainty in the model. Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).*
>
> *b) Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.*
>
> *c) Describe this stock's vulnerability (see "Appendix to the SAW ToRs") to becoming overfished, and how this could affect the choice of ABC*

I consider this ToR to have been met.

Projections of SSB were made for the years 2017 to 2066 using three different harvest policies based on status quo effort and landings, allocated quota, and fishing at the overfishing threshold. Projections for each year assumed time series average recruitment with uncertainty in starting SSB equal to the model-estimated uncertainty in the final model year. Most catch was assumed to come from the southern region, consistent with the historical time series. Sensitivity runs using arbitrarily low and high values for natural mortality and recruitment were included in the stock assessment report and an additional run with zero recruitment over the projection period was completed during the review meeting (at the request of the Panel). All projections indicated that

biomass will remain above $SSB_{THRESHOLD}$, for the entire resource, including when fishing at $F_{THRESHOLD}$ and under zero recruitment. Thus, projections strongly indicate that there is virtually no risk that the biomass will decline to $SSB_{THRESHOLD}$ within 50 years, assuming natural mortality fluctuates around its assumed value and no unusual high-mortality events occur. I agree with the assessment team, the working group, and the other members of the Panel that the stock's vulnerability to overfishing over the coming decades is therefore very low under all reasonable scenarios and that this conclusion is relatively insensitive to the choice of ABC (Acceptable Biological Catch).

> ToR 9. *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.*

I consider this ToR to have been met.

The assessment team documented progress against 21 research recommendations from the previous assessment (SARC 48) and compiled a list of 14 additional recommendations. Nine of the 21 existing recommendations were reported as complete. These included changes to the survey, estimation of survey selectivity and catchability, development of a length-structured model, and simulation modelling to determine proxy MSY reference points. One recommendation related to the previous survey platform was dropped as being no longer relevant. Significant and ongoing progress was reported for seven recommendations. These related to age and growth studies, maturity, investigation of spatial structure and incorporation into assessment models, survey design and evaluation of the use of underwater photography to survey ocean quahog. No progress was noted for four recommendations, relating to fecundity-at-size, the relationship of dredge efficiency with depth, ocean quahog density and substrate type, and incorporation of size-selectivity in the Patch depletion model.

I agree with the other members of the Panel in endorsing the significant progress that has been made against a very substantial list of recommendations, and the positive impact that this has had on the assessment, including the new statistical catch-at-age model, a much-improved survey, and the better estimation of important biological parameters. Several of the ongoing research recommendations address the mismatch of spatial scale between the assessment and ocean quahog demographic processes, and I strongly support further research on these. Because the current stock assessment model has fixed natural mortality, a mismatch between observed growth patterns and the assumed growth model, and (therefore) an estimated recruitment time series that very likely does not reflect biological reality, I believe the highest priority for future research should be accorded to estimating and modeling biological parameters. Further work on survey dredge efficiency with respect to ocean quahog density and bottom type, and the relationship between the depletion estimates of catchability and the survey q(s) in the model would also be useful to improve the model's performance on estimating biomass scale.

During the meeting, an informally rationalized and prioritized list of research recommendations was compiled by the WG Chair, at the request of the Panel. Those recommendations were grouped into age and growth, survey and fishery topics, roughly prioritized. There was not the opportunity for a full discussion during the meeting, but I agree with the other members of the Panel that the list provides an effective basis for further discussion on priorities by the WG.

Recommendations for Future Work

I have made the following suggestions on future work in this individual report. These might be considered to complement the WG's research recommendations and the two sets are by no means mutually exclusive or comprehensive:

- Consider a more detailed examination of the factors driving catch on a survey tow, including assessing whether pre-1982 surveys can be "corrected" and made more comparable with more recent surveys;
- Tidy up the analysis of surveys addressing "nearly random" tows, treatment of un-fishable stations, correcting historical data for poorly-estimated tow length, and "data borrowing";
- Optimise surveys and re-allocate survey resources for ocean quahog and Atlantic surf clam to drive maximum benefit for the two stock assessments;
- Assess the extent to which changing catchability with multiple passes of a hydraulic dredge affects the accuracy of depletion experiments of dredge efficiency;
- Estimate the proportion of ocean quahog that may be invulnerable to a hydraulic dredge and, if this is not negligible, develop a method of allowing for this inside or outside the stock assessment model;
- Increase the measurement precision for ocean quahogs to 1 mm during surveys and port samples;
- Continue to monitor the distribution of ocean quahogs to assess any effects of climate change and ocean acidification;
- Continue the development of the ageing method and the ageing of ocean quahog from a wide range of geographic regions and depths to better understand their biology, including over historical time-scales;
- Develop methods to overcome the mismatch of scale between the assessment model and the fishery, potentially by developing finer-scale models to understand dynamics in areas that are fished relatively frequently;
- Modify the SS3 package to enable a wider variety of growth curves, including non-asymptotic curves for animals with indeterminate growth;
- Modify the SS3 package to facilitate and automate MCMC estimation of posterior distributions for Bayesian models and underpin analysis and presentation of MCMC diagnostics.

Bibliography

The following papers and documents were provided for the review or cited in this report.

Background Papers

Chute A, Hennen D, Russell R, Jacobson L. 2013. Stock Assessment Update for Ocean Quahogs (*Arctica islandica*) through 2011. NEFSC Ref Doc 13-17; 156 p.

Harding JM, King SE, Powell EN, Mann R. 2008. Decadal Trends in Age Structure and Recruitment Patterns of Ocean Quahogs (*Arctica islandica)* from the Mid-Atlantic Bight in Relation to Water Temperature. *Journal of Shellfish Research* ***27***: 667-690.

Hennen DR. 2015. How Should We Harvest an Animal That Can Live for Centuries? *North American Journal of Fisheries Management* ***35***: 512-527.

Hennen D, Jacobson L, and Tang J. 2012. Accuracy of the Patch model used to estimate density and capture efficiency in depletion experiments for sessile invertebrates and fish. *ICES Journal of Marine Science* ***69***: 240–249.

Kilada, R. W., Campana S. E., and Roddick, D. 2007. Validated age, growth, and mortality estimates of the ocean quahog (*Arctica islandica*) in the western Atlantic. *ICES Journal of Marine Science* ***64***: 31–38.

Murawski SA, Ropes JW, Serchuk FM. 1982. Growth of the Ocean Quahog, *Arctica islandica*, in the Middle Atlantic Bight. *Fisheries Bulletin* ***80***:1; 14 p.

Northeast Fisheries Science Center. 2009. 48th Northeast Regional Stock Assessment Workshop (48th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc 09-15; 834 p.

Northeast Fisheries Science Center. 2009. 48th Northeast Regional Stock Assessment Workshop (48th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc 09-10; 50 p.

Rago PJ, Weinberg JR, Weidman C. 2006. A spatial model to estimate gear efficiency and animal density from depletion experiments. *Canadian Journal of Fisheries and Aquatic Sciences* ***63***: 2377-2388.

Ridgway ID, Richardson CA. 2011. *Arctica islandica*: the longest lived non colonial animal known to science. *Reviews in Fish Biology and Fisheries* ***21***: 297-310.

Thorarinsdottir GG, Jacobson LD. 2005. Fishery biology and biological reference points for management of ocean quahogs (*Arctica islandica*) off Iceland. *Fisheries Research* ***75***: 97-106.

Witbaard, R. 1996. Growth variations in *Arctica islandica* L. (Mollusca): a reflection of hydrography-related food supply. *ICES Journal of Marine Science* ***53***: 981–987.

Working Papers for Review

Working Group, Stock Assessment Workshop (SAW 63) 2017. Stock Assessment Report of Ocean Quahog. SAW/SARC 63. February 21-23, 2017. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 404 p.

Working Group, Stock Assessment Workshop (SAW 63) 2017. Stock Assessment Summary Report of Ocean Quahog. SAW/SARC 63. February 21-23, 2017. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 10 p.

Working Group, Stock Assessment Workshop (SAW 63). 2017. Term of Reference 9 – Revised Research Recommendations. 2 p.

Presentations

Working Group, Ocean Quahog. 2017. Ocean Quahog Assessment 2017. PowerPoint presentation. 90 slides.

Papers Cited in this Review

Francis, R.I.C.C.; Elliott, G.; Walker, K. (2015). Fisheries risks to the population viability of Gibson's wandering albatross *Diomedea gibsoni*. *New Zealand Aquatic Environment and Biodiversity Report No. 152*. 48 p. http://www.mpi.govt.nz/document-vault/7632

Starr, P.J.; Webber, D.N. (2016). The 2015 stock assessment of red rock lobsters (*Jasus edwardsii*) in CRA 5 and development of management procedures. *New Zealand Fisheries Assessment Report 2016/41*. 115 p. http://www.mpi.govt.nz/document-vault/13131

Appendix 1. Statement of Work

Statement of Work

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

63rd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC)
Benchmark stock assessment for Ocean quahog

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information may be obtained from www.ciereviews.org.

Scope

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 or not the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fisheries within the jurisdiction of NOAA's Greater Atlantic Regional Fisheries Office (GARFO).

The purpose of this meeting will be to provide an external peer review of a benchmark stock assessment for Ocean quahog. The requirements for the peer review follow. This Statement of Work (SOW) also includes Appendix 1: TORs for the stock assessment, which are the responsibility of the analysts; Appendix 2: a draft meeting agenda; Appendix 3: Individual Independent Review Report Requirements; and Appendix 4: SARC Summary Report Requirements.

Requirements
NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The SARC chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the SARC chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the SOW, OMB Guidelines, and the TORs below. All TORs must be addressed in each reviewer's report. No more than one of the reviewers selected for this review is permitted to have served on a SARC panel that reviewed this same species in the past. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include forward projecting statistical catch-at-age 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 (BRPs) that includes an appreciation for the varying quality and quantity of data available to support estimation of BRPs. For ocean quahogs (a bivalve), knowledge of long-lived, sedentary invertebrates would be useful.

Requirements for Reviewers

- Review the background materials and reports prior to the review meeting
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the SARC Chair with contributions to the SARC Summary Report
- Deliver individual Independent Review Reports to the Government according to the specified milestone dates
- 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 below in the "Requirements for SARC panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent 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 Report produced by each reviewer.
- The Independent 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.

<u>Requirements for SARC panel</u>

- During the SARC meeting, the panel is to determine whether each stock assessment Term of Reference (TOR) of the SAW 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 TOR of the SAW.

- If the panel rejects any of the current BRP or BRP proxies (for BMSY and FMSY 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 tasks in accordance with the SOW and Schedule of Milestones and Deliverables below.

Requirements for SARC chair and reviewers combined:

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.

The SARC Chair, with the assistance from the reviewers, will write the SARC Summary Report. Each 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 will not be submitted, reviewed, or approved by the Contractor.

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.

Foreign National Security Clearance

When 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 reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, country of current residence, dual citizenship (yes, no), passport number, country of passport, travel dates.) to the NEFSC SAW Chair 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/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts.

Period of Performance
The period of performance shall be from the time of award through April 7, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

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

No later than January 17, 2017	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
No later than February 7, 2017	NMFS Project Contact will provide reviewers the pre-review documents
Feb. 21 - 23, 2017	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
February 23, 2017	SARC Chair and reviewers work at drafting reports during meeting at Woods Hole, MA, USA
March 9, 2017	Reviewers submit draft independent peer review reports to the contractor's technical team for review
March 9, 2017	Draft of SARC Summary Report, reviewed by all reviewers, due to the SARC Chair *
March 16, 2017	SARC Chair sends Final SARC Summary Report, approved by reviewers, to NMFS Project contact (i.e., SAW Chairman)
March 23, 2017	Contractor submits independent peer review reports to the COR and technical point of contact (POC)
March 30, 2017	The COR and/or technical POC distributes the final reports to the NMFS Project Contact and regional Center Director

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

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each TOR as specified; (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $20,000.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
Dr. James Weinberg, NEFSC SAW Chair
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov Phone: 508-495-2352

Appendix 2. Stock Assessment Terms of Reference for SAW/SARC-63

The SARC Review Panel shall assess whether or not the SAW Working Group has reasonably and satisfactorily completed the following actions.

A. Ocean quahog

1. Estimate catch from all sources including landings and discards. Map the spatial and temporal distribution of landings, discards, and fishing effort, as appropriate. Characterize the uncertainty in these sources of data.
2. Present the survey data being used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, length data, etc.). Use logbook data to investigate regional changes in LPUE, catch and effort. Characterize the uncertainty and any bias in these sources of data. Evaluate the spatial coverage, precision, and accuracy of the new clam survey.
3. Describe the relationship between habitat characteristics (e.g., benthic, pelagic, and climate), survey data, and ocean quahog distribution, and report on any changes in this relationship.
4. Evaluate age determination methods and available data for ocean quahogs to potentially estimate growth, productivity, and recruitment. Review changes over time in biological parameters such as length, width, and condition.
5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR 4, as appropriate) and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.
6. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. 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.
7. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to any new model or models developed for this peer review.
   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-6).
8. Develop approaches and apply them to conduct stock projections.
   a. Provide numerical annual projections (5 – 50 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level), including model estimated and other uncertainties. Consider cases using nominal as well as potential levels of uncertainty in the model. Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
   b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

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

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

Clarification of Terms used in the Stock Assessment Terms of Reference

Guidance to SAW WG about "Number of Models to include in the Assessment Report":
In general, for any TOR in which one or more models are explored by the WG, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the WG and explain their strengths, weaknesses and results in relation to the "best" model. 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. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. 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 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" (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 Maximum Sustainable Yield (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)

Participation among members of a Stock Assessment Working Group:
Anyone participating in SAW 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.

Appendix 3. Draft Review Meeting Agenda

63rd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC)
Benchmark stock assessment for A. Ocean quahog

February 21-23, 2017

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

AGENDA* (version: Feb. 15, 2017)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, Feb. 21

10 – 10:30 AM			
Welcome	**James Weinberg**, SAW Chair		
Introduction	**Edward Houde,** SARC Chair		
Agenda			
Conduct of Meeting			
10:30 – 12:30 PM	Assessment Presentation (A. Ocean quahog)		
	Dan Hennen		**Toni Chute**
12:30 – 1:30 PM	Lunch		
1:30 – 3:30 PM	Assessment Presentation (A. Ocean quahog)		
	Dan Hennen		**Toni Chute**
3:30 – 3:45 PM	Break		
3:45 – 5:45 PM	SARC Discussion w/ Presenters (A. Ocean quahog)		
	Ed Houde , SARC Chair		**Toni Chute**
5:45 – 6 PM	Public Comments		
7 PM	**(**Social Gathering**)**		

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Wednesday, Feb. 22

9:00 – 10:45	Revisit with Presenters (A. Ocean quahog)		
	Ed Houde, SARC Chair		**Alicia Miller**
10:45 - 11	Break		

11 – 11:45 Revisit with Presenters (A. Ocean quahog)
Ed Houde , SARC Chair **Alicia Miller**

11:45 – Noon Public Comments

12 – 1:15 PM Lunch

1:15 – 4 Review/Edit Assessment Summary Report (A. Ocean quahog)
Ed Houde , SARC Chair **Alicia Miller**

4 – 4:15 PM Break

4:15 – 5:00 PM SARC Report writing

Thursday, Feb. 23

9:00 AM – 5:00 PM SARC Report writing

*All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public; however, during the Report Writing sessions on Feb. 22-23, we ask that the public refrain from engaging in discussion with the SARC.

Appendix 4. Individual Independent Peer Review Report Requirements

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 report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the SARC Summary Report.

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, but especially where there were divergent views.

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

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

3. The 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.

Appendix 5. SARC Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background and 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 or not 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 reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the 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 (BRPs) 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 6. SAW 63 Working Group Members and Attendees at the SARC 63 Panel Meeting

NAME	AFFILIATION	EMAIL
Ed Houde	U Maryland Center for Environmental Science	ehoude@umces.edu
Anthony Hart	Western Australian Fisheries	Anthony.Hart@fish.wa.gov.au
Mike Bell	Herriot-Watt University – Intl Centre for Island	M.C.Bell@hw.ac.uk
Martin Cryer	Ministry for Primary Industries, Wellington	Martin.Cryer@mpi.govt.nz
Russ Brown	NEFSC	Russell.brown@noaa.gov
Jim Weinberg	NEFSC	james.weinberg@noaa.gov
Larry Jacobson	NEFSC	larry.jacobson@noaa.gov
Dan Hennen	NEFSC	Daniel.hennen@noaa.gov
Jessica Coakley	MAFMC	jcoakley@mafmc.org
Chris Legault	NEFSC	chris.legault@noaa.gov
Sheena Steiner	NEFSC	sheena.steiner@noaa.gov
Alicia Miller	NEFSC	alicia.miller@noaa.gov
Toni Chute	NEFSC	toni.chute@noaa.gov
Mark Terceiro	NEFSC	mark.terceiro@noaa.gov
José Montañez	MAFMC	jmontanez@mafmc.org
Joe Myers	Bumble Bee/Snow's Foods	
Tom Hoff	Wallace & Associates	tbhoff@verizon.net
Daphne	Rutgers University	dmunroe@hsrl.rutgers.edu
Tom Alspach	Sea Watch International	talspach@goeaston.net
Eric Powell	University of Southern Mississippi	eric.n.powell@usm.edu
Roger Mann	VIMS	rmann@vims.edu
D.H. Wallace	Wallace & Associates	DHWallace@aol.com
Doug Potts	NMFS/GARFO	douglas.potts@noaa.gov
Gary Shepherd	NEFSC	gary.shepherd@noaa.gov