Summary Report for CIE
Stock Assessment Review of:
Ocean quahogs, Atlantic surfclams, and the skate species complex

SARC44, November 28-December 4, 2006
Woods Hole, MA

Summary Report to:
The Center for Independent Experts,
University of Miami

Prepared by:
Cynthia M. Jones
December 13, 2006

Executive Summary

A stock assessment review committee (SARC-44) met in Woods Hole, MA from November 28 to December 4, 2006 to review and evaluate three stock assessments: 1) Ocean quahog (*Artica islandica*), 2) Skate-species Complex [Barndoor skate, *Dipturus laevis*; Clearnose skate, *Raja eglanteria*; Little skate, *Leucoraja erinacea*; Rosette skate, *Leucoraja garmani*; Smooth skate, *Malacoraja senta*; Thorny skate, *Amblyraja radiate*; Winter skate, *Leucoraja ocellata*])**,** and 3) Surclam (*Spisula solidissima*). The CIE appointed a panel composed of two reviewers, Patrick Cordue and Dr. Vivian Haist, and a chairperson, Dr. Cynthia M. Jones. The lead reviewer for Ocean quahog and Atlantic surfclam was Dr. Haist and Cordue was lead reviewer for the skate-species complex. Both reviewers and the chair read background material on all species and participated in evaluation, discussion, and feedback throughout the meeting. Subsequent to the meeting, both reviewers sent the chair their reports (by December 8, 2006) and likewise, the chair sent the reviewers a copy of the summary report to verify that it reflected their views (prior to December 13, 2006). Both reviewers agreed that the summary report accurately reflected their views.

One of the charges to the panel in the statement of work (SOW) was to explore and report areas of agreement among the panel members, and when possible to suggest alternative approaches when a term of reference (TOR) was not considered to have been met satisfactorily. Throughout the meeting and subsequently in their reports, the panel members were largely in agreement on the evaluation of all TOR. When they gave conflicting assessments of whether a TOR was met, they often had similar evaluations that differed in the degree of their concerns. When possible and when appropriate, they also suggested alternative approaches when a TOR was not met satisfactorily.

For Ocean quahog, the panel agreed that six of the seven TOR were met satisfactorily. For TOR 2 there was disagreement on whether it was completed satisfactorily. One reviewer stated that the "basecase assessment was done satisfactorily" but that the KLAMZ assessment "could have been better." The other reviewer stated that "This term of reference was not completed in its entirety." The tasks in TOR 2 were to "estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years." In general, the panelists felt that estimates were robust to the issues they identified. The panel's main concern was in how the KLAMZ model (a delay-difference model used as the primary assessment tool) was used to assess the trajectory of absolute biomass, because it failed to link long-term average recruitment to virgin biomass. The biomass trajectory was characterized by low biomass early in the series, a peak in 1978, subsequently followed by low biomass. Moreover, the estimation of absolute biomass relies very heavily on the correctness of dredge efficiency (q) in calculating from the NEFSC area-swept estimates of relative biomass in fishery-independent surveys. As an alternative, the panel suggested that the model be run with an explicit link between recruitment and virgin biomass. When this was done during the meeting, the outcome was a biomass estimate of one-half the base run and a fishing mortality (F) that was double. This result led the panel to conclude that uncertainty had not been adequately captured in this TOR and that there was a clear need for more adequate testing. Nonetheless, current biomass is well above the B_{MSY} proxy and F is below F_{MSY}, so the assessment does provide credible guidance for management.

Both panel members also commented that the biological reference points (BRPs) should be re-examined given research that suggests more conservative targets are required for long-lived species with low productivity. Additionally the working group (WG) did point out that $F_{25\%}$ was not an appropriate proxy for F_{MSY} but did not propose an alternative as requested in this TOR.

Another issue arose during the SARC as to the ad hoc treatment of missing data in the clam survey. Currently when a cell is empty, it is filled in by borrowing the values of its neighbors. Both panel members found this approach unacceptable given the availability of statistically rigorous approaches and they suggest the use of GLM to the original data, thus formalizing the approach and allowing for the estimation of year-area effects. Further suggestions appear in their individual reports.

Although outside the TOR, the panel expressed concern that this species was not managed spatially in light of the fact that individuals of this species are sedentary, long-lived and having low productivity. Such a species will be vulnerable to localized depletions and the panel suggested that more modeling be done at some point to explore the management consequences from this type of life history with and without spatial management.

For the skate-species complex, the panel agreed that important terms of reference had not been met satisfactorily. TOR 1-3 were evaluated as not completed satisfactorily or not completed in their entirety. TOR 4-6 were evaluated as being met satisfactorily with TOR 6 seen as being very well done. Seven species were assessed at one time, which added to the difficulty of meeting all the TOR. Most importantly, there is a lack of species-specific catch data for skates and this makes it very difficult to assess the status of the stocks. Further, the working group attempted to estimate F using newly published methods that had not been vetted for the behavior of the estimator when underlying assumptions are violated as is likely the case for trawl survey length frequency data. Both panel members were in full agreement about this and offered suggestions on alternative approaches including new simulations and other ways to use catch data. They suggest that it would be valuable "to obtain at least qualitative species-specific estimates (aiming for plausible upper and lower bounds on removals)."

The panel recognized the value of the skate WG effort to develop more rational BRP than the current ad hoc reference points. However, the panel members both agreed that the use of the new BRPs was premature and they need further testing before being used. Again both panel members were concerned about using methods that relied on estimates of recruitment obtained by the use of cohort slicing (using length-frequencies as a proxy for age) with data that lack clear modes in size. Hence, they recommended that even though the current BRPs are ad hoc, they are the best available.

The Atlantic surfclam assessment met six of the seven TOR satisfactorily and provides an acceptable basis for management. For TOR 3, one panelist evaluated it as not met satisfactorily while the other stated that it was completed, but qualified this evaluation with a detailed review of their concerns. The panelists agreed that improvement was needed in how surfclams were modeled. Unlike Ocean quahog, the data for Atlantic surfclam includes a wealth of age information. Because of the quality and value of this age data, both panelists were in complete agreement that an age-structured model was far more appropriate for surfclams than the delay-difference model (KLAMZ) that was used for this assessment, along with a complete catch history. Hence, this assessment was less informative than it could have been by providing direct estimation of reference points, more complete characterization of uncertainty, and more consistency ”between the data and model assumptions” that would have been provided by a fully integrated age-structured model.

A very similar approach and methods were used in the modeling of surfclams as Ocean quahog and so the same concerns were raised by the panel as were raised for Ocean quahogs. For example, when cells were empty in the NEFSC clam survey, they were filled by borrowing from adjacent cells. Again this ad hoc approach is unnecessary given the well developed statistical theory for handling such problems. The panel suggests using a GLM approach here too. The same problem exists for the disconnect in the model between the estimated mean recruitment and biomass, which is noticed in the difference between real recruitment events and those frequency of those events from the model.

The panel also noted that the WG noted inadequacies in current reference points but offered no alternatives. Both panelists observed that the use of B_{1999} is a questionable proxy for B_0. Both also agreed that this issue would be resolved with the use of a fully integrated stock assessment model based on all available data

1. Introduction

1.1 Background

The SARC-44 CIE panel was presented with three stock assessments to review and evaluate: 1) Ocean quahog (*Artica islandica*), 2) Skate-species Complex [Barndoor skate, *Dipturus laevis*; Clearnose skate, *Raja eglanteria*; Little skate, *Leucoraja erinacea*; Rosette skate, *Leucoraja garmani*; Smooth skate, *Malacoraja senta*; Thorny skate, *Amblyraja radiate*; Winter skate, *Leucoraja ocellata*])**,** and 3) Surclam (*Spisula solidissima*).

<u>Ocean quahog</u> - Ocean quahog are treated as a single stock in US waters from Virginia to Maine based on results of genetic analyses. Although the species occurs at depths of 10-400 m, in US waters it occurs from 20-80 m and is almost entirely found in EEZ waters. Ocean quahogs are long-lived (50-100 y commonly) and annual natural mortality is assumed to be around 2% (M=0.02/y). However, age is not done as part of the stock assessment. Growth subsequent to recruitment is slow (0.51-0.77% per year in weight). Sexual maturity is variable, with the smallest mature male at 36 mm and 6 years but others still immature at 8-14 years. Recruitment is assumed to be small (<1%/y). Successful reproduction appears to be regional and sporadic on the order of decades.

The fishery occurs throughout this range except on Georges Bank because of potential paralytic poisoning. Almost all of the fishable stock occurs in EEZ waters. The fishery has been governed through quotas since 1979, with landings peaking at 22,000 mt in 1992 and averaging 17,000 mt recently. Based on information from logbooks, fishing effort has shifted from southern areas in the mid-Atlantic Bight to Delmarva and southern New Jersey, with increased effort recently in southern New England which now has the bulk of the landings. Effort has shifted as with localized depletion of parts of the population.

The fishery is managed by ITQ except in Maine waters where it is managed under a small quota. For stock assessments, a delay-difference model is use (KLAMZ) and efficiency-corrected swept-area methods are used to estimate absolute biomass. A cumulative catch model was also used to estimate biomass and F for southern Virginia. Discards are assumed to be 0 and incident mortality from dredging assumed to be 5%.

It is important to note that for biomass reference points, overfishing definitions and management measures apply to the entire stock, not just to the fishable portion. However for fishing mortality reference points, F is calculated over the fishable portion of the stock only. Biological reference points include: targets of ½ virgin biomass (estimated biomass in 1978 as the proxy for virgin biomass) and $F_{0.1}$; thresholds of ¼ virgin biomass (1/2 B_{MSY} proxy) and $F_{25\%}$.

The data available for the stock assessment come from the NEFSC triennial clam-dredge survey data from 1977-2005 (which differed in season over the time course) wherein location and abundance, biomass, and length are recorded usually in their entirety, and swept-area calculated. Surveys use stratified random samples. The efficiency of the NEFSC clam dredge survey (1997, 1999, 2002, 2005) was evaluated in a series of depletion experiments conducted by NMFS in collaboration with industry. This fishery is subject to mandatory reporting with landings and effort (trip-level detail at ten-minute squares) available from logbooks. As noted in logbooks, discards are putatively nil in this fishery.

Skate-species Complex – Seven species of skates are found from Maine to Virginia from the subtidal to over 700 m. This species group includes Barndoor skate (*Dipturus laevis*), Clearnose skate (*Raja eglanteria*), Little skate (*Leucoraja erinacea*), Rosette skate (*Leucoraja garmani*), Smooth skate (*Malacoraja senta*), Thorny skate (*Amblyraja radiate*), and Winter skate (*Leucoraja ocellata*). Distributions differ, depending on the species over this area, but there is considerable overlap for many species. Natural mortality is assumed to be equal to K. These species are generally long-lived, slow growing, and with low fecundity, thereby leaving them more vulnerable to overfishing.

The directed commercial fishery for skates use otter trawls, but skates are also caught as bycatch in groundfish trawling and scallop dredging. There is a long history of fishing for skate beginning in the 1800s, but there have been recent periods of high catch in the 1950s and 1960s from the industrial fleet, in 1993, and again recently. Prior to 2003, the catch of skate species was reported only as “unclassified”, whereas since then they must be reported to species. Similarly, information on discards was only available as aggregated data and discards are believed to be 2-4 times larger than landings. Recreational landings as insignificant when compared to commercial landings and are believed to have an order of magnitude greater number of released skates. The release mortality is unknown but taken as 10-15%, which is found for flounders and other demersal species.

Although there are some regulations governing the harvest of skates in U.S. waters, there are difficulties for stock assessment scientists in obtaining reliable catch histories, such as were available in the Ocean quahog and surfclam assessments.

The data available for this species complex comes from the NEFSC bottom surveys and from observers on commercial vessels that catch skate. Fisheries-independent sampling does provide measures of density, length, sex, and sexual maturity. Unlike Ocean quahogs and surfclams, there is no efficiency correction and, thus the density measures are relative and are not absolute measures of abundance.

Surfclam – Atlantic surfclam are treated as a single stock along their range in US waters, although subspecies have been identified. Their range is from North Carolina to Maine at depths from below the sub-tidal in state waters to 60 m in the EEZ, but densities are low at depths greater than 40 m. Over time, biomass has shifted north in response to fishing and climate change. Ages of over 20 years are common and maximum observed age exceeds 35 years. Natural mortality is assumed to be M=0.15. First age of reproduction can be as early as one with full maturity reached by two years. Recruitment is high occasionally thus producing important cohorts, but recruitment has been low recently since 2000 based on survey age composition. Data on length-at-age show an apparent reduction in growth since 1993, with a concomitant delay by 1-2 years in the age of recruitment.

The fishery occurs throughout this range except on Georges Bank because of potential paralytic poisoning. There is substantially variability in biological characteristics and in exploitation rates over different regions. Similarly to Ocean quahogs, the majority of landings and effort have shifted from North Carolina north to New Jersey where landings have been greatest since 1985. Discarding was substantial during the 1980s but has declined since 1990 when minimum size limits were abandoned. Surfclams are fully available to the commercial gears by 110-120 mm length. Between 1982-1992 recruitment to the fishery occurred at 5 ½ years, but since the noted change in growth it is closer to 7 ½ years. Incidental mortality is assigned to be 12% which is an upper bound and landings are inflated by this amount to estimate catch.

Since 1979, the fishery has been managed through quotas and under an ITQ system in the EEZ established in 1990. For stock assessments, a delay-difference model is used (KLAMZ) and efficiency-corrected swept-area methods are used to estimate absolute biomass and F. In addition to model-based F estimates, F is also calculated from the ratio of catch to efficiency-correct swept-area biomass. Additionally, trends in LPUE were used to corroborate results of the KLAMZ model. .

Biological reference points are based on the corrected biomass in 1999 as the proxy for virgin biomass. B_{MSY} is set as ½ B_{1999}. Threshold biomass is ¼ B_{1999}. F_{MSY} is set to equal natural mortality (F=M=0.15/y).

The data available for stock assessment are similar to Ocean quahog. Fishery-independent data come from the NEFSC triennial clam-dredge survey data from 1977-2005 (which differed in season over the time course) wherein location and abundance, biomass, and length are recorded usually in their entirety, and swept-area calculated. Surveys use stratified random samples. The efficiency of the NEFSC clam dredge survey (1997, 1999, 2002, 2005) was evaluated in a series of depletion experiments conducted by NMFS in collaboration with industry. Landings and effort are available from logbooks under mandatory reporting. Note that previous trip regulations make it difficult to interpret effort data prior to 1985 (This point is a bit unclear in our documents). Additionally, landings, age, and lengths are also available from the commercial fishery through port sampling.

1.2 Review of Activities

The CIE participants received the materials to review starting November 13, 2006 (See Appendix 3), in preparation of the SARC/SAW-44 meeting which was held at the Northeast Fisheries Science Center in Woods Hole, MA from November 28- December 4, 2006.

Formal presentations for the three stock assessments (Ocean quahogs, Skate-species Complex, Atlantic surfclams) were made on November 28 and 29 at the Northeast Fisheries Science Center in Woods Hole, MA. Among the CIE panel, Vivian Haist was lead reviewer for Ocean quahogs and Surfclams and Patrick Cordue was lead reviewer for the Skate-species Complex. Notwithstanding designation of lead reviewers, the entire CIE panel participated in the reviews of all assessments. The meeting agenda is given in Appendix 4. During the meeting, assessment scientists were provided with feedback as to the sufficiency of their analyses. Additional presentations were done from November 29-December 1 in response to questions and requests from the CIE panel. An additional meeting was held on December 2 with Dr. Paul Rago to answer a few further questions about the skate complex assessment. During the meeting a question arose in concern to the length-based mortality estimator used in the skate-complex assessment and a half-hour telephone call to its authors, Drs. Gedamke and Hoenig, was made to obtain more information on November 29. Over the course of the open sessions, questions were also entertained from the audience to the assessment scientists and to the panel members.

The CIE panel met in closed session from December 1- 4, 2006 at the Northeast Fisheries Science Center in Woods Hole, MA to produce the individual and summary reports. Participants finalized their reports subsequently according to the CIE timetable (See Appendix 2). During this time the CIE panel explored their level of agreement, as requested on page 4, paragraph 1 of the CIE SOW. All reports were reviewed and accepted by the CIE panel.

1.3 Comments on SARC-44 Process

In contrast to past SARC meetings, more emphasis was placed on examining where there were (or not) areas of agreement among the panel on issues in the terms of reference. Also the timeline was shorter and emphasized more group interaction by the panel during the writing process than had been requested previously in the panel's experience of CIE reviews. While this could be problematic if the CIE panel held strong differing opinions, it was not for our panel because we had all independently arrived at strikingly similar opinions as to the assessments and their sufficiency. Thus, the request that the summary report identify areas of agreement could be easily met.

The SOW for this SARC also required very quick turn-around of written documents, such that the CIE panel had completed their rough drafts by December 4. The framework and background of the Summary document was also completed by then. Although this timetable did force swift completion, it did not permit the panel an additional interval to "mull over" the assessments and perhaps provide an even better in-depth review and recommendations.

2 .Assessments Reviewed

2.1 Ocean quahogs – Terms of Reference

1. Characterize the commercial catch including landings and discards.

This term of reference (TOR) was completed satisfactorily.

The Ocean quahog fishery is almost exclusively commercial and recreational landings are nil. Additionally, discards are negligible and incidental mortality, which is caused by dredge operation, is assumed to be 5%.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

There was some level of disagreement over whether this term of reference was completed satisfactorily, in its entirety. Dr Haist concluded that it was not met satisfactorily because it had not been done properly in its entirety. Cordue concluded that it had been completed because the base case analysis was robust to the problems in the assessment; these problems were the same as noted by Dr. Haist.

At the level of a base-case analysis, the estimation of stock status is robust. Additionally, because estimation of the fishable biomass does not rely on knowledge of maturity or selection by the gear, it is a more robust estimate than either spawning stock biomass or total biomass. Hence, both reviewers thought that this component of the TOR was met. Estimates of absolute biomass and F and their uncertainty rely on estimation with the delay-difference KLAMZ model and fishery-independent surveys, which used efficiency-corrected swept area. Both panel members were in complete agreement about their concerns over how the KLMAZ model was used. Both commented that runs of the model did not link average recruitment to virgin biomass. "This approach is consistent with allowing a regime shift in that the virgin mean recruitment lead to a large accumulated biomass, which has since declined due to fishing and a shift to much lower mean recruitment (the estimated recruitment was very low for most regions)." Such an approach relies very heavily on the efficiency estimates being correct; "absolute biomass estimates from the KLAMZ model are totally dependent on the survey dredge efficiency estimate." Both panel members also suggested an alternative approach wherein a sensitivity analysis was run given an explicit link between recruitment and virgin biomass, and that there is useful information in the relative biomass estimates. At the panel's request a KLAMZ run was done during the SARC under this scenario and it resulted in one half the base-run biomass and twice the F. The panel concluded that the uncertainty had therefore not been adequately characterized previously due to the lack of testing alternative assumptions in more runs. In this vein, one panel member suggested that KLAMZ-generated recruitment be compared with survey data to determine whether the modeled recruitment is plausible, but recognized that the slow growth of Ocean quahogs might make such a comparison problematic.

Both panel members praised the innovative use of the PATCH model (Rago and Weinberg 2006) for estimating dredge efficiency from depletion experiments. Both offered several suggestions to further improve estimation of uncertainty with this model:

1) errors in dredge position, 2) differences in strip width; 3) assumptions about selectivity of the gear. They then suggest that this approach be used with other depletion studies.

Another issue arose during the SARC as to the ad hoc treatment of missing data in the clam survey. Currently when a cell is empty, it is filled in by borrowing the values of its neighbors. Both reviewers found this too ad hoc given the availability of statistically rigorous approaches and suggest the use of GLM to the original data, thus formalizing the approach and allowing for the estimation of year-area effects. Further suggestions appear in the individual reports.

3. *Either update or redefine biological reference points (BRPs; proxies for BMSY and FMSY), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.*

This TOR was completed, with just a few points of concern. The existing reference points were updated and improved based on the KLAMZ model, length-based yield per recruit, and spawning stock biomass per recruit analysis. The assessment report does note some concern over the use of the reference points, B_{msy}, F_{thr}, and F_{msy} as proxies in the current FMP$\left(\text{e.g. } 0.5B_0 \text{ and } F_{25\%}\right)$ which may likely be too high. One panel member further suggested that the tiered structure adopted by the North Pacific Fisheries Management Council and that the "model based approach of Francis (1992) should also be considered as a standard approach for U.S. stocks and could be used in conjunction with a tier structure or not."

Both panel members were in agreement that the entire catch time series from 1967 should be used in modeling to improve the estimate of virgin biomass in contrast to using B_{1978} as a proxy.

Although the stock is managed as one population, both panel members commented on the value of using a spatially-based approach to such a long-lived sessile species that can be highly vulnerable and subject to localized depletions. Although somewhat beyond this TOR, their reports provide further detailed comments.

4. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was completed satisfactorily.

5. *Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.*

This term of reference was completed satisfactorily.

The KLAMZ model was used for this and is an appropriate tool.

6. *If possible,*

a. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and
b. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

This term of reference was completed satisfactorily.

The projections were done using the base model. While this model does capture a great deal of the uncertainty, it does not capture all. It would be improved by including uncertainty due to future recruitment and without constant F.

7. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments*.

This term of reference was completed satisfactorily.

2.2 Skate Species Complex – Terms of Reference

1. *Characterize the commercial and recreational catch including landings and discards.*

This term of reference was not completed entirely.

The assessment provided a good summary of commercial landings and discards. The recreational fishery is of minor importance and a general summary was provided. Because landings are not reported to species, it is more problematic to attempt a reconstruction of the species-specific catch histories. Beyond this, discard data are needed to estimate catches and discard data also are problematic. However, since 1989 better discard reporting has occurred and coincides with observer data, albeit with some evidence that observers confuse some of the species identifications. The panel members appreciated that it will be tedious to attempt to disaggregate these data and uncertainty will still exist because of misidentification. Despite the difficulties, both panel members agreed that an attempt to do so would provide upper and lower bound estimates that would be helpful in understanding stock status. As one panel member states: "To conduct analytical stock assessments for specific skate species, catch time series are needed."

2. *Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.*

This term of reference was not completed satisfactorily.

Both panel members agreed that without a reliable species-specific catch history, this TOR could not be completed satisfactorily despite a great deal of effort on the part of the stock analysts. No estimates of total biomass or spawning stock biomass were made because this cannot be done without a reliable history of landings, accurate discard data, or efficiency-corrected trawl data. Hence, the estimates that are presented are relative abundances. One panel member discussed the value of providing ball park estimates of trawl efficiency (q) by estimating area and vertical availability, and vulnerability to the survey gear. Although not precise, this approach would yield upper and lower bound estimates, which in themselves would be informative. Without reliable data on absolute total biomass, relative biomass has value and can be used to track potential declines within the survey trend of relative biomass. Both panel members also agreed that bootstrapping was a correct procedure to determine uncertainty in the trawl survey data.

The stock analysts employed new techniques in hopes of being able to estimate fishing mortality in this data-poor complex. The panel saw value in using new techniques, but disagreed with the use of the Gedamke-Hoenig (2006) length-based mortality estimator in stock status decisions until this estimator has been more thoroughly vetted. Both panel members were in complete agreement that the mortality estimators were unreliable and should be rejected. Each had somewhat different concerns, but also agreed with each others concern. Their two main concerns follow.

The Gedamke-Hoenig estimator improves on the Hoenig (1987) length-based estimate of mortality. It provides an estimate of total mortality (Z), from which F is estimated for an assumed M (from M=K). However, it still relies on assumptions that are not likely to be met in the skate-complex assessment. These assumptions include "1) constant and know natural mortality rate (for the skate analyses assumed to be equal to the von Bertalanffy growth K parameter); 2) constant growth with know maximum length (von Bertalanffy growth L_∞); 3) constant recruitment; and 4) constant knife-edged selectivity with know length of full selectivity." To use this estimator for the skate complex, it must first be shown that violation of assumptions (such as trends in recruitment or changes in growth) will not invalidate the estimator by producing incorrect Z estimates. This has not been done and will require extensive and complex simulations.

.

The second concern was that, despite its publication in a peer-reviewed journal, the estimator had not been subjected to sufficient analysis to validate its use (especially in determining stock status). The second panel member stated that "The minimum requirements are: (i) the method is shown to be reliable under specific criteria; and (ii) the criteria are shown to be satisfied for the particular assessment to which it is applied. The first step was not achieved in the paper (except that the estimator was derived under specific assumptions which could never be met in practice). The second step was not addressed by the WG." Simulations to test the behavior of the model require a realistic operational model whose conditions can be use to produce Z values with the estimator. Such a model would include:

- Recruitment variability (independent lognormal and correlated lognormal options)
- Growth variability (spatial and temporal)
- A variety of trajectories for fishing mortality (e.g., increasing, decreasing, stepped)
- Realistic fishing selectivity (option for temporal variation)
- Realistic survey selectivity (option for temporal variation)

- Growth and natural mortality parameters which deviate from M = K
- Variability or perhaps shifts in natural mortality

To thoroughly evaluate the behavior of the estimator, the simulation "must explore a wide range of model misspecification" to test whether the estimator is robust to these misspecification. Barring this approach, the second panel member stated that analytic derivation must be accomplished, which is difficult to do and rarely applicable to fisheries.

Taken together, this panel was in agreement that the use of the Gedamke-Hoenig estimator was premature. Similarly, they found that the length-tuned model estimation of F for thorny skate was innovative, but also was not acceptable. Therefore, they rejected the F values produced for this TOR

3. *Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.*

This term of reference was not completed satisfactorily.

The panel agreed that the existing BRPs were ad hoc and that there is a real need to improve upon them. The existing BRPs rely on the biomass index obtained in the NEFSC trawl survey. The use of these reference points assumes that the stocks were not already overfished by the initiation of the time series.

New BRPs were proposed for five species. The methods to obtain these reference points include Stock-recruit fits and length-based yield per recruit analysis. Both are promising approaches but both also rely on cohort slicing to putative age groupings. When there are clear modes in the length frequencies that clearly reflect year classes, then this is a valid approach. However, this is not true for most of the skates, particularly for barndoor skate which have little evidence of modes. Further work will need to be done before these methods can be considered valid.

Even though the panel saw value in the new BRPs they did not see that sufficient work had been done to justify their use at this time and suggested that a panel of experts be convened for this purpose. They felt that it was not appropriate for a SARC to come up with new BRPs.

4. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was completed satisfactorily.

5. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC-reviewed assessments.*

This term of reference was completed satisfactorily. In their detailed reports the reviewers indicated that several of the recommendations are being actively pursued, and suggest that it would help to prioritize the others.

6. *Examine the NEFSC Food Habits Database to estimate diet composition and annual consumptive demand for seven species of skates for as many years as feasible.*

This term of reference was completed satisfactorily.

The entire panel was impressed with the work that was completed for this TOR. Altogether it was an excellent job. The analysis was thorough, sources of uncertainty were identified, and assumptions clearly stated.

2.3 Surfclam – Terms of Reference

1. *Characterize the commercial and recreational catch including landings and discards*.

This term of reference was completed satisfactorily.

There is a long catch history available from 1965, near the beginning of the commercial fishery. Although there was a short period of heavy discarding (up to 37%) that had relatively minor impact on the model runs. These data are also corrected using an upper bound estimate of 12% for incidental mortality rather than using the average and variance in the model run. Although ad hoc, this has little effect on the model run.

2. *Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.*

This term of reference was completed satisfactorily almost in its entirety, but the panel members both were concerned that the full uncertainty was not modeled, or modeled correctly.

Because the same approaches were used for surfclams as Ocean quahogs, many of the same comments are appropriate. Estimates of absolute biomass and F and their uncertainty rely on estimation with the delay-difference KLAMZ model with input from 1985-2005 and fishery-independent surveys, which used efficiency-corrected swept area. Both panel members were in complete agreement about their concerns over how the KLAMZ model was used. The estimates of efficiency drive the estimates of total biomass, and thus the validity and uncertainty in efficiency estimates are of paramount importance to the assessment. Again, estimate mean recruitment was not linked to starting biomass. Fishable biomass is a more reliable estimate than either total biomass or spawning stock size because it does not rely on selectivity assumptions or maturity ogives.

Both panel members were in complete agreement that bootstrapped estimates were inappropriate given that swept-area biomass included uncertainty in *q*, and this is not captured when using bootstrapping. Both panel members agreed that a Bayesian approach would better reflect true uncertainty and that uncertainty is now underestimated. The individual reports that are attached give more detailed advice.

In surfclam as with Ocean quahog, the PATCH model is seen as a significant improvement in estimating dredge efficiency. Notwithstanding the review panel's appreciation, the PATCH model still does not capture the true uncertainty and this is an area that could be improved as suggested for the Ocean quahog, especially by running sensitivity analyses on the selectivity assumptions.

The same issues of imputation in the survey cells apply to surfclam as they did to Ocean quahog. The current procedure for imputation to empty cells is ad hoc by borrowing from adjacent cells. The panel members again agreed that this approach was inappropriate when there are better and statistically valid methods available, such as GLM. Again both agreed that GLM replace the ad-hoc "borrowing" method now used.

The panel agreed that assessment of surfclam could be improved with the used of the considerable data available on surfclam ages. Surfclams are routinely aged and provide reliable ages. Both panel members recommend the use of a fully integrated age structured model as has been suggested previously.

3. *Either update or redefine biological reference points (BRPs; proxies for BMSY and FMSY), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.*

One reviewer stated that this TOR was completed satisfactorily while the other stated that it was not fully completed.

The working group did discuss some of the inadequacies of current reference points, but did not recommend any alternatives. However, one panel member points out that "use of B_{99} as a proxy for virgin/unfished biomass is perplexing. In such a data-rich setting there is no need for proxies for virgin biomass, B_{MSY}, or F_{MSY}."

4. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was completed satisfactorily.

5. *Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.*

Although not explicitly stated by one reviewer, this TOC was met, albeit with some concerns raised. The reviewers differed on how completely they believed this TOR was met and both express concern over the projected recruitments. Both panel members held similar views that insufficient work had been undertaken on sensitivity analyses with regard to recruitment, which was treated as a nuisance parameter in the model.

6. *If possible,*
c. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and
d. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

This term of reference was completed satisfactorily.

Although this TOR was completed, panel members thought that there was a lack of sufficient simulations to provide the better projections. Both suggest that a Bayesian approach captures uncertainty, while bootstrapping is less appropriate and minimizes true uncertainty.

7. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.*

This term of reference was completed satisfactorily.

The panel suggests that new recommendations be prioritize and that an age-structured assessment model has still not been developed, as recommended in the previous assessment. This panel strongly recommends that development of an age-structure model is a high priority.

Appendix 1 -**Terms of Reference for the 44th Northeast Regional Stock Assessment Workshop** (Revised Sept. 6, 2006)

A. Ocean quahogs

1. Characterize the commercial catch including landings and discards.
2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.
3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.
4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).
5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.
6. If possible,

a. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and

b. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

B. Skate species complex

1. Characterize the commercial and recreational catch including landings and discards.
2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.
3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.
4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).
5. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC-reviewed assessments.
6. Examine the NEFSC Food Habits Database to estimate diet composition and annual consumptive demand for seven species of skates for as many years as feasible.

C. Atlantic surfclam

1. Characterize the commercial and recreational catch including landings and discards.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.
3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.
4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).
5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.
6. If possible,
c. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and
d. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.
7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

Appendix 2– Statement of Work to CIE Panel

Consulting Agreement between University of Miami and Participants SARC 44 Statement of Work

General

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication.

The Center for Independent Experts (CIE) shall provide a chair and two panelists for the 44th Stock Assessment Review Committee panel. The panel will convene at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from November 28 – December 4, 2006 to review three assessments (ocean quahogs, *Arctica islandica*; Atlantic surfclams, *Spisula solidissima*, and a skate species complex [Barndoor skate, *Dipturus laevis*; Clearnose skate, *Raja eglanteria*; Little skate, *Leucoraja erinacea*; Rosette skate, *Leucoraja garmani*; Smooth skate, *Malacoraja senta*; Thorny skate, *Amblyraja radiate*; Winter skate, *Leucoraja ocellata*]). In the days following the review of the assessments, the panelists will write individual independent review reports (referred to below as "Independent" reports), and then the SARC Chairman and panel shall use the Independent reports to write the SARC Summary Report.

Specific Activities and Responsibilities

The CIE's deliverables shall be provided according to the schedule of milestones in the table below. The main CIE deliverable will be the SARC Summary Report that will provide key information for a presentation to be made by NOAA Fisheries at meetings of the New England and Mid-Atlantic Fishery Management Councils in early 2007. The SARC Summary Report shall be an accurate and fair representation of the CIE panel viewpoint on how well each of the Terms of Reference of the SAW were completed (please refer to Annex 1 for the SAW Terms of Reference).

The SARC panelists' duties shall occupy a maximum of 14 days per person (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; and several days following the meeting to produce the Independent Reports and the SARC Summary Report).

The SARC chair's duties shall occupy a maximum of 19 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the meeting to start preparation of the SARC Summary Report; and several days after the meeting to finalize the SARC Summary Report).

Charge to panel

The panel is to determine and write down whether each Term of Reference of the SAW (see Annex 1) was or was not completed successfully during the SARC meeting. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. 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. The chair shall identify or facilitate agreement among the panelists for each Term of Reference of the SAW, where possible.

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

Roles and responsibilities

(1) Prior to the meeting

(SARC chair and panelists)

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

(2) During the Open meeting

(SARC chair)

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

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

(SARC panelists)

For the three stock assessments, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a scientist/reviewer's point of view, determine whether each Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a panelist considers any existing Biological Reference Point proxy to be inappropriate, the panelist should try to recommend an alternative, should one exist.

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

(3) After the Open meeting

(SARC panelists)
Each panelist shall prepare an Independent report addressing each Term of Reference of the SAW for each of the stock assessments reviewed. Independent reports will be included as appendices of the SARC Summary Report. Independent reports need to specify and provide an explanation about whether each Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the "Charge to panel" statement.

If any existing Biological Reference Point (BRP) 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 panelist. See Annex 2 for further details on the contents of Independent reports.

(SARC chair)
Prepare a document summarizing the background to the work to be conducted as part of the SARC 44 process and summarizing whether the process was adequate to complete the Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report.

(SARC chair and panelists)

The Chair and panel will prepare the main body of the SARC Summary Report. Each panelist and the chair will read all panelists' independent review reports with the purpose of discussing whether the panelists hold similar views on each Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar or a consensual view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the reviewers to reach an agreement if they cannot reach one. The chair is not required to express the chair's opinion on each Term of Reference of the SAW, specifically because the chair's role is not that of an independent reviewer.

The SARC Summary report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. If any existing Biological Reference Point (BRP) proxies are considered inappropriate, the SARC Summary report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time. See Annex 3 for further details on the SARC Summary report contents.

The contents of the SARC Summary report will be approved by the panelists by the end of the SARC Summary Report development process. The chair will complete all final editorial and formatting changes prior to the final submission of the SARC Summary Report to the CIE, in consultation with the panelists, as the chair deems necessary. The chair will provide the panelists with a final copy of the final SARC Summary Report provided to the CIE.

The milestones and schedule are summarized in the table below. No later than December 13, 2006, the SARC Chair should submit the SARC Summary Report to the CIE for review[1]. The SARC Summary Report shall be addressed to "University of Miami Independent System for Peer Review," and sent to Dr. David Sampson, via e-mail to David.Sampson@oregonstate.edu and to Mr. Manoj Shivlani via e-mail to mshivlani@rsmas.miami.edu

[1] All reports will undergo an internal CIE review before they are considered final.

Milestone	Date
Workshop at Northeast Fisheries Science Center (NEFSC) (begin writing reports, as soon as open Workshop ends)	November 28 - December 1, 2006
SARC Chair and reviewers work at the NEFSC. Reviewers prepare Independent reports. Chair begins to draft SARC Summary Report.	December 1 - 4
Independent SARC reports due to CIE for review and due to CIE Chair	December 8
SARC Summary Report due to CIE for review	December 13
CIE provides reviewed Independent reports to NMFS COTR for approval	December 21
COTR notifies CIE of approval of reviewed Independent reports	December 22 *
CIE provides reviewed SARC Summary Report to NMFS COTR for approval	December 27
COTR notifies CIE of approval of reviewed SARC Summary Report	January 3 *
CIE provides final SARC Summary Report with signed cover letter to COTR	January 4
COTR provides final SARC Summary Report to NEFSC contact	January 4

* Assuming no revisions are required of the reports.

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.

NEFSC Contact person and SAW Chairman:
Dr. James R. Weinberg, NEFSC, Woods Hole, MA. 508-495-2352, James.Weinberg@noaa.gov

Submission and Acceptance of Consultants' Reports

The CIE shall provide via e-mail the final SARC Summary Report in pdf format to Dr. Lisa Desfosse (Lisa.Desfosse@noaa.gov) for review by NOAA Fisheries and approval by the COTR, Dr. Stephen K. Brown, by December 27, 2006. The COTR shall notify the CIE via e-mail regarding acceptance of the report by January 3, 2007. Following the COTR's approval, the CIE will provide the final SARC Summary Report with signed cover letter to the COTR by January 4, 2007. The COTR will then transmit the report to the NEFSC contact, as soon as possible.

ANNEX 1:

Terms of Reference for the 44th Northeast Regional Stock Assessment Workshop

(Revised Sept. 6, 2006)

A. Ocean quahogs

1. Characterize the commercial catch including landings and discards.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.

6. If possible,

    a. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and

    b. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

B. Skate species complex

1. Characterize the commercial and recreational catch including landings and discards.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and

redefined BRPs.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

5. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC-reviewed assessments.

6. Examine the NEFSC Food Habits Database to estimate diet composition and annual consumptive demand for seven species of skates for as many years as feasible.

C. Atlantic surfclam

1. Characterize the commercial and recreational catch including landings and discards.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.

6. If possible,

    c. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and

    d. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

ANNEX 2: Contents of SARC Independent Reports

1. For each assessment reviewed, the report should address whether each Term of Reference of the SAW was completed successfully. For each Term of Reference, state why that Term of Reference 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. 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.

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

3. Any independent analyses conducted by the reviewers as part of their responsibilities under this agreement should be incorporated into their Independent reports. It would also be helpful if the details of those analyses (e.g, computer programs, spreadsheets etc.) were made available to the respective assessment scientists.

4. Additional questions that were not in the Terms of Reference but that are directly related to the assessments. This section should only be included if additional questions were raised during the SARC meeting.

ANNEX 3: Contents of SARC Summary Report

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

   To make this determination, the chairman and panel 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 panel does not reach an agreement on a Term of Reference, the report should explain why.

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

3. The report shall also include as separate appendices the independent review reports prepared by each panelist, the bibliography of all materials provided during SAW 44, and any papers cited in the panelists' reports, along with a copy of the Statement of Work.

   The report shall also include as a separate appendix the Terms of Reference used for SAW 44, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panelist advice.

Appendix 3. Material Made Available For Review by CIE Panelists and Chair

Files to review for SAW/SARC44

Stock	Type of file & Doc. ID	Filename (.pdf)
A. Ocean quahog		
	New files	
	A-1	Quahog - EEZ Assessment Summary - SAW44
	A-2	Quahog - EEZ Assessment - SAW44
	A-3	Quahog - EEZ Assess Appendix - SAW44
	A-4	Quahog - Maine Assessment Report 2005.pdf
	Background	
	A-5	ocean quahog-CRD0403-assessment in 2003.pdf
	A-6	ocean quahog-CRD0404-advisory report in 2003.pdf
	A-7	Ocean quahog Maine Assessment 2003 Background.pdf
	A-8	CJFAS-RagoWeinbergWeidman
	No number	Biological reference points for surfclams and Ocean Quahogs: Rago 6.24.1998
	Rapporteur Reports	
	Rapp OQ SARC 11-28-2006	Meeting notes
	Rapp OQ SARC 11-30-2006	Meeting call back notes
	Powerpoint Presentations	
	OQ-SARC44-Day 1-6	Original SARC 44 presentation
	Bits and pieces for quahog	Presentation of CIE requested material
	Quahog never say done-1	Presentation of CIE requested material
B. Skates		
	New files	

B-1	Skate- Summary Report - SAW44
B-2	Skate - Assessment - Exec Summary
B-3	Skate - Assessment Text & Tables
B-4	Skate - Assessment Figs
B-5	Skate - Assessment Addendum

Background

B-6	Skate - Advisory 1999 - SAW30
B-7	Skate - Assessment 1999 - SAW30
B-8	Skate - 2004 Review – NEFMC
B-9	Skate - 2005 Review – NEFMC
B-10	Skate - FMP excerpt – NEFMC
B-11	Brodziak_Legault_model averaging
B-12	Gedamke_and_Hoenig_length_based_Z
No number	Gedamke 2006 PhD Thesis abstract and excerpts Chapter 4

Rapporteur Reports

Skates_Rapp_notes_final	Meeting notes

Powerpoint Presentations

SARC 44-Bskates-TOR6b

C. Atlantic surfclam

New files

C-1	Surfclam- Summary Report - SAW44
C-2	Surfclam- Assessment- SAW44

Background

C-3	Surfclam - Advisory 2003 - SAW37
C-4	Surfclam - Assessment 2003 - SAW37
C-5	Surfclam - Cooperative Survey 2004

Rapporteur Reports

SurfClamRapport 11-06old	Meeting notes 11-30-2006
SurfClamRapport 11-06b	Meeting call back notes 12-1-2006

Powerpoint Presentations

Surfclam-Take-1-d	Presentation 11-29-2006
Surfclam Remains Of the Day-3	Call back presentation 11-3-2006
SurfclamProjectF=)naturalScale Survey-1	Call back presentation 12-1-2006

Appendix 4. Agenda for SARC 44

44th Northeast Regional Stock Assessment Workshop (SAW 44)
Stock Assessment Review Committee (SARC) Meeting

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

November 28 – December 4, 2006

AGENDA (11-27-06)

TOPIC	PRESENTER	SARC LEADER	RAPPORTEUR
Tuesday, 28 November (1:00 – 5:00 PM)……………………………………………			
Opening			
Welcome	**James Weinberg**, SAW Chairman		
Introduction	**Cynthia Jones**, SARC Chairman		
Agenda			
Conduct of Meeting			
Ocean quahog (A)	**Larry Jacobson**	**Vivian Haist**	**Toni Chute**
SARC Discussion	**Cynthia Jones**		
Wednesday, 29 November (8:30 – Noon)……………………………………………			
Skates (B)	**Kathy Sosebee**	**Patrick Cordue**	**Michelle Traver**
SARC Discussion	**Cynthia Jones**		
Wednesday, 29 November (1:15 – 5:00 PM).……………………………….……			
Atlantic surfclam (C)	**Larry Jacobson**	**Vivian Haist**	**Laurel Col**
SARC Discussion	**Cynthia Jones**		

Thursday, 30 November (8:30 – 5:00 PM) ………………..……………………

Revisit Assessments (A – C) with presenters, as needed.

Friday, 1 December (8:30 AM –) ………………………………………………

Revisit Assessments (A – C) with presenters, if needed.

SARC Report writing. (closed)

Saturday, 2 December - Monday, 4 December ……….……………………………

SARC Report writing. (closed)

Appendix 5. SARC report by Dr. Vivian Haist

Northeast Regional Stock Assessment Review: ocean quahogs, Atlantic surfclams, and skate species complex stock assessments

SARC44, November 28-December 4, 2006
Woods Hole, MA

Prepared for:
The Center for Independent Experts,
University of Miami

Prepared by:
Vivian Haist
Haist Consultancy
1262 Marina Way,
Nanoose Bay, BC
Canada V9P 9C1
haistv@shaw.ca

December 08, 2006

Executive Summary

This report presents results of an independent peer review of three east coast stock assessments (ocean quahogs, Atlantic surfclams, and a skate species complex), conducted for the Center for Independent Experts, University of Miami. The assessments were reviewed during the Northeast Regional Stock Assessment Review Committee meeting (SARC44), held at Woods Hole, MA, November 28 – December 4, 2006.

The ocean quahog assessment provides a scientifically credible basis for fishery management, given the entire EEZ area is managed as a single biological unit. The available data is adequate and reliable, and the KLAMZ delay-difference model is an appropriate tool for conducting the stock assessment.

The summary of ocean quahog status, based on the notion of the EEZ area as a unit stock, ignores stock trends occurring at finer geographic scales and indications of serial depletion of the resource. Finer-scale management of the fishery or rotational harvesting approaches may be required to ensure a sustainable fishery. Biological reference points (BRPs) for ocean quahog should be re-examined in light of studies that suggest more conservative targets are required for long-lived species with low productivity. Given the resource appears to be healthy, redefinition of BRPs is not critical to management advice arising from this assessment.

For the seven skate species under review, assessment of stock status relative to biological reference points is hindered by the lack of catch data for the individual species. This makes determination of status relative to fishing mortality reference points virtually impossible. Substantial effort went into estimating fishing mortality rates from length frequency data, however length frequency data alone cannot provide reliable estimates of current fishing mortality rates for developing management advice.

For skate, there are long time-series of fishery independent trawl survey relative abundance indices, beginning in the early 1960s for most species. The survey data should provide fairly good information on stock trends for most of the skate species. The working group investigated a number of options for redefining skate BRPs. While some of the approaches are promising, additional work is required to investigate assumptions of the analyses. Although ad hoc, current BRPs are the best available at this time.

The Atlantic surfclam stock assessment provides a credible scientific basis for developing fishery management advice. However, the delay-difference KLAMZ model used for the assessment does not use the wealth of age composition data that is available, so the assessment was considerably less informative than it might have been. A fully integrated age-structured model analysis of the surfclam data would allow direct estimation of reference points, better characterization of uncertainty, and greater consistency between the data and model assumptions. The use of B_{1999} as a proxy for B_0 is questionable. The most appropriate approach for redefining surfclam BRPs is through an integrated catch-age analysis – until that is done the existing BRPs are the best available.

Introduction

This document reports on an independent peer review of three east coast stock assessments (ocean quahogs, Atlantic surfclams, and a skate species complex), conducted for the Center for Independent Experts, University of Miami. The assessments were reviewed during the Northeast Regional Stock Assessment Review Committee meeting (SARC 44), held at Woods Hole, MA, November 28 – December 4, 2006. The review followed the guidelines summarized in the Statement of Work (Appendix 2) and review activities are summarized in the main body of this report (Section 1.2).

This report forms one component of the SARC44 Summary Report. It is structured in line with the terms of reference (TORs) for the assessment Working Groups, and addresses the question: *was each TOR completed successfully?* A summary of the opinions of the two CIE reviewers with respect to the TORs is presented in the main body of this document. There are no areas where my views differed from those of the second CIE reviewer.

Summary of findings

Ocean quahog

Ocean quahog is a long-lived (> 100 years old) and sessile species. The ocean quahog resource in the US Exclusive Economic Zone (EEZ) is managed as a unit stock. The stock assessment reported in the 44th SARC concludes that the ocean quahog resource is not overfished and overfishing is not occurring. Estimates of current fishing mortality $(F = 0.0077)$ are substantially lower than the F_{msy} proxy $(F_{0.1} = 0.0278)$. Fishable biomass, while decreasing, is well above the B_{msy} target.

The ocean quahog assessment provides a scientifically credible basis for fisheries management, given the entire EEZ area is currently managed as a biological unit. The available data is adequate and reliable, and the KLAMZ delay-difference model is an appropriate tool for conducting the stock assessment. A time series of fishery independent dredge surveys in conjunction with dredge efficiency estimates from depletion studies provide absolute biomass estimates, which anchors the assessment. Additional sensitivity analyses, applying the KLAMZ model under alternative hypotheses, would significantly improve the assessment.

The summary of stock status, based on the notion of the EEZ area as a unit stock, ignores stock trends occurring at finer geographic scales and the indication of serial depletion of the resource. Finer-scale management of the fishery or rotational harvesting approaches may be required to ensure a sustainable fishery.

Biological reference points (BRPs) for ocean quahog should be re-examined in light of studies that suggest more conservative targets are required for long-lived species with low productivity. However, given estimates of fishing mortality that are well below the existing F_{msy} proxy and estimates of current biomass that are above the B_{msy} proxy, redefinition of BRPs is not critical to management advice arising from this stock assessment.

Discussion relative to Terms of Reference

8. *Characterize the commercial catch including landings and discards.*

This term of reference was completed successfully.

Commercial landings in directed fisheries are well known. By-catch in other fisheries is minimal. Recreational fisheries are essentially nil. Discards are also negligible because of automatic sorting equipment releases undersized fish. Fishery induced incidental mortality is assumed to be 5%.

9. *Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.*

This term of reference was not completed in its entirety. Fishable biomass and fishing mortality rates were estimated for 1978 through 2005. The fishable biomass metric is likely more precise than spawning stock or total biomass measures because it does not require expansions for size categories with low survey selectivity. Hence estimating fishable biomass rather than spawning stock or total biomass is appropriate. The method used to estimate uncertainty in the fishing mortality and biomass estimates (based on covariance matrix at maximum of posterior density) is not the best available, given the stock assessment model (KLAMZ) was formulated with a prior on the abundance index scaling parameter. Also, sensitivity analyses employing alternative structural assumptions in the assessment model would have provided a more thorough characterization of the uncertainty in biomass and mortality estimates.

Biomass and fishing mortality are estimated using the delay-difference KLAMZ model. This is a reasonable analytical approach, given the data available. For ocean quahog there are fairly precise estimates of absolute biomass based on survey estimates of relative density in conjunction with estimates of survey dredge efficiency from experimental studies. As demonstrated during the SARC 44 review, absolute biomass estimates from the KLAMZ model are completely dependent on the survey dredge efficiency estimate.

Considerable effort has gone into development of the PATCH model for estimating dredge efficiency from experimental depletion studies. The PATCH model accounts for the “patchy” re-sampling of survey cells when survey tows are used in a depletion study, which is a significant improvement over previous approaches used to analyze the ocean quahog depletion study data. Estimates of survey dredge efficiency are sensitive to the assumed effective size of sampled cells and to the assumption of asymptotic selectivity with known size of full selection. The stock assessment would be improved if the dredge efficiency prior incorporated uncertainty due to those assumptions of the depletion model.

The KLAMZ model formulation used for the ocean quahog assessment decouples the estimated average recruitment level from the estimated virgin biomass level $\left(B_{1978}\right)$. That is, there is no requirement that the average recruitment result in biomass at or near to the B_{1978} level in the absence of fishing. Under this model formulation, the estimated average recruitment level results in unfished biomass levels well below B_{1978}, implying that recruitment over the past 25 years has been well below the longer-term historic levels. While this recruitment trend is plausible, it is informative to conduct a sensitivity run that assumes recent (last 25 years) recruitment is equal to the longer-term recruitment that generated the B_{1978} stock abundance. This sensitivity run, conducted during the SARC 44 review, results in a 50% reduction in biomass levels with a concomitant doubling of fishing mortality estimates and an implied doubling of the survey dredge efficiency parameter. This stock reconstruction is also plausible – given uncertainty in some of the assumed constants in the PATCH model analysis (ie. cell size, asymptotic and known size fully selected), survey dredge efficiency may well be twice the current estimate. The two KLAMZ runs reflect two alternative hypotheses about stock dynamics: high initial biomass generated by high historical recruitment with substantially lower recruitment levels in past 25 years or lower initial abundance and no change in recruitment in past 25 years.

The KLAMZ model analysis could be significantly improved by: 1) beginning the stock reconstruction in 1967 (beginning of catch data series) to obtain a better estimate of B_0; 2) incorporating additional uncertainty in PATCH model estimates of survey dredge efficiency that result from model assumptions (cell sizes, selectivity of gear) when developing the prior for the efficiency parameter; 3) investigating alternative assumptions about recent versus historic average recruitment levels, possibly through a prior in the model; and, 4) using a fully Bayesian implementation of the model to estimate the joint posterior distribution for characterizing uncertainty of model parameters (ie. an MCMC or SIR algorithm).

10. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

This term of reference was completed. Ocean quahog BRPs were updated based on KLAMZ model biomass estimates and length-based yield per recruit and spawning stock biomass per recruit analysis. The assessment document notes that the reference points for B_{msy} and F_{msy} proxies in the current FMP $\left(0.5B_0 \text{ and } F_{0.1}\right)$ are likely too high, citing some recent papers that suggest for long-lived species like ocean quahog maximum sustainable yield is achieved with lower fishing mortality rates at higher average stock biomass. The working group did not recommend redefinition of the BRPs. Further work is required to assess the adequacy of current targets and thresholds.

For ocean quahog, the estimated biomass in 1978, the first year of the stock reconstruction, is taken as a proxy for B_0. The KLAMZ assessment model could be initiated in 1967, closer to the beginning of the fishery, to generate an improved estimate of B_0.

The ocean quahog fishery management plan (FMP) defines the entire EEZ area as a unit stock for management purposes. Although outside the scope of this review, it would be imprudent not to comment on the appropriateness of the "unit stock" concept currently used for management. Ocean quahog are sessile, slow growing, and long-lived. They have a low natural mortality rate and low productivity. This type of resource is highly vulnerable to serial depletion. Data summaries presented in the ocean quahog stock assessment suggest serial depletion, both at the 10-minute square and stock assessment region level. The southern stock assessment regions all appear to be below B_{msy} levels, and fishing effort and landings have shifted north, primarily to Long Island which appears to have had a recruitment event in the 1990s. It is unlikely that the current management approach will lead to a sustainable fishery. Management strategy evaluation (e.g. FAO 2002, Stokes et al. 1999, and papers therein) would be a useful approach to investigate the resilience of the management system to alternative assumptions about reproduction and recruitment dynamics.

11. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was successfully completed. The 2005 fishable biomass was estimated to be 65% of the virgin level (B_{1978}), well above the 50% target. The 2005 fishing mortality rate estimate of 0.0077 is well below the threshold $\left(F_{25\%}\right)$ of 0.0517.

12. *Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.*

This term of reference was completed. Using swept-area biomass estimates from the 2005 survey as the starting point for the projections, uses of the "best" available estimates. However, projecting forward using the KLAMZ model under alternative hypotheses about recruitment would provide a better basis for exploring uncertainty.

13. *If possible,*

e. *provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and*

f. *compare projected stock status to existing rebuilding or recovery schedules, as appropriate.*

This term of reference was completed. The proposed method for characterizing uncertainty assumes error only in the 2005 swept-area biomass estimate. While this is certainly the largest component of the projection uncertainty, uncertainty in future recruitment should also be considered (see TOR 2).

14. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.*

This term of reference was completed satisfactorily.

Skate species complex

For the seven skate species under review, assessment of stock status relative to biological reference points is hindered by the lack of catch data for the individual species. This makes determination of status relative to realistic fishing mortality targets and thresholds virtually impossible. The working group put considerable effort into estimating fishing mortality rates based on survey length frequency data, however those approaches cannot provide reliable estimates of current *Fs* for developing management advice.

For skate, there are long time-series of fishery independent trawl survey relative abundance indices, beginning in the early 1960s for most species. The survey data, which is standardized for vessel and gear changes over the period of the surveys, should provide fairly good information on stock trends. A caveat is that the survey area does not appear to encompass the full geographic range of a few of the species.

The working group investigated a number of options for redefining current BRPs. Both the stock-recruit and length-based yield per recruit approaches are promising, but additional work is required to investigate uncertainty in the recruitment estimates. Although ad hoc, the current skate BRPs are the best available at this time.

Discussion relative to Terms of Reference

7. *Characterize the commercial and recreational catch including landings and discards.*

This term of reference was only partially completed. Commercial landings and discard data were provided for the total skate complex, but no attempt was made to disaggregate these to individual species. Recreational catch of skates is minor, and a general summary of the recreational fishery was provided.

Commercial landings data, available for 1964 through 2005, were generally reported as "unclassified" species prior to 2003 when a change in regulations required skate species to be reported. The assessment reports these data, but does not attempt to disaggregate the landing in the unclassified category to individual species. To conduct analytical stock assessments by species, catch time-series by species are required.

Estimates of skate discards, for both directed and by-catch fisheries, are presented for 1989-2005, the period where observer data is available. Methods for estimating discards are appropriate, and the level of observer coverage in recent years appears to be adequate Estimates of the precision of the discard estimates would be useful to assess their reliability over time. As with the landings data, no attempt was made to disaggregate the total discards of skate to species. Skate discards are significant, as much as 6 times the reported landed catch in some years.

Observer data could be used to disaggregate the unclassified skate landings and to estimate species-specific discards. As pointed out in the assessment document there are considerable problems with species mis-identification in the observer database, which will make the job of disaggregating the data tedious and difficult. It is not likely that catch estimates by species would be very accurate, but a range of plausible estimates based on alternative assumptions could be developed. Additionally, estimates of underreporting, believed to be significant, could be developed and included in the alternative scenarios of catch trajectories.

8. *Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.*

This term of reference was not successfully completed, largely due to limitations in the available data. Biomass estimates were provided as relative indices rather than absolute levels. Fishing mortality rates were estimated using length-based methods, however assumptions of the methods are not met and therefore estimates are unreliable.

The two methods used to estimate fishing mortality rates from the trawl survey length frequency data, Hoenig (1987) and Gedamke and Hoenig (2006), rely on similar assumptions. These include: 1) constant and know natural mortality rate (for the skate analyses assumed to be equal to the von Bertalanffy *K* parameter); 2) constant growth with know maximum length (von Bertalanffy L_∞); 3) constant recruitment; and 4) constant knife-edged selectivity with know length of full selectivity. Hoenig (1987) relies on equilibrium conditions; that is relatively constant fishing mortality rates over a period long enough for the population (and catch sample) length structure to transition to a stable state. Gedamke and Hoenig (2006) attempt to move past the equilibrium assumptions of the Hoenig (1987) method by estimating the mean fishing mortality rates during assumed stanzas of constant mortality and the transitions between them. Both methods provide estimates of total mortality. Assumed values for the natural mortality rate are used to obtain fishing mortality estimates.

Simulation methods have not been applied to either of the methods to determine their robustness to deviations from the assumptions of the methods. Intuitively it is clear that the methods would not provide accurate estimates of fishing mortality under certain conditions (eg. significant changes in growth, trends in recruitment). Realistic simulation experiments are required to assess under what conditions the methods provide reliable estimates of fishing mortality. Reliability needs to be interpreted relative to goals for the analysis – using the terminal fishing mortality estimates in determining overfishing status.

Spawning stock biomass and total stock biomass estimates were presented as relative indices based on long-term trawl survey data. Uncertainty in the biomass indices was estimated using bootstrap methods, which are appropriate for the trawl survey data. No attempt was made to convert the indices to absolute biomass estimates. It may be possible to obtain swept-area biomass estimates using the survey data and priors on the catchability coefficient of the survey vessels. While such an approach may not provide accurate estimates, it could useful in providing a plausible biomass range.

9. *Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.*

A number of alternatives to the existing biological reference points were examined and new BRPs proposed for five of the seven skate species (winter skate, little skate, barndoor skate, thorny skate, and clearnose skate). The proposed BRPs, based on stock-recruit or yield-per-recruit analysis, have a stronger analytical basis than the existing reference points and are a positive step in redefining the skate reference points. However, prior to their acceptance further work is required to assess their stability given uncertainty in recruitment estimates. The fishing mortality reference points from the length-based yield per recruit analyses do not use the recruitment estimates, and they are a scientifically credible improvement over the existing fishing mortality reference points. However, without estimates of current fishing mortality rates, they cannot be used to determine overfishing status.

The primary limitation of the stock-recruitment and length-based yield per recruit analyses for estimating reference points is the potentially large uncertainty in the recruitment estimates. Cohort slicing, an approach that can produce unreliable age estimates, was used to estimate relative year-class strengths. For some of the skate species the cohort slicing method may work. That is, when there are distinct modes for the younger fish in the survey length-frequency data, slicing based on the distribution of the modes can produce reasonably reliable estimate of relative cohort sizes. For some of the skate species there are no clear year-class modes, and alternate methods are required. A potential approach is to apply the iterative algorithm of Kimura and Chikuni (1987) including uncertainty (prior distributions?) in the mean and variance of length-at-age.

Replicate recruitment series could be generated to assess the influence of uncertainty in the growth parameters on the BRP estimates.

The existing BRPs are *ad hoc*. The biomass based reference points, calculated from the distributions of survey indices, assumes the stocks have moved through the B_{msy} level since the beginning of the surveys. Given an unknown exploitation history prior to the beginning of the surveys, the validity of this assumption is unclear. The fishing mortality rate reference points are based on recent trends (3-year running means) in the survey index. Without information on catch trends, the fishing mortality reference points may lead to erroneous conclusions. Updating the existing reference points would only perpetuate the problems with them.

While the existing BRPs are not highly credible, there are no suitable alternatives so they remain the best available BRPs at this time. The stock-recruitment and length-based yield per recruit analyses are promising, and should provide more credible reference points in the near future.

10. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was completed successfully.

11. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC-reviewed assessments.*

This term of reference was completed successfully.

12. *Examine the NEFSC Food Habits Database to estimate diet composition and annual consumptive demand for seven species of skates for as many years as feasible.*

This term of reference was successfully completed.

Food habits were examined for all seven species in the skate complex. Based on food habits data, diet composition, per capita consumption and assumptions about skate abundance, the amount of prey removed by skates was calculated. The analysis was thorough and sources of uncertainty and assumptions of the analysis clearly stated.

Additional comments

An exploratory length-based model was used to analyze data for thorny skate. Following are some ideas for future development of this approach. A strictly length-based model could be used with transition matrices describing the change in size from one year to the next. Nothing is gained by modeling length-at-age, and the effect of length selectivity is lost. Aggregating the length frequency data into larger bins may improve the estimation. The length frequency data has a strong zigzag pattern that suggests there may have been some rounding in the measurement of lengths. This pattern makes analysis of the residuals difficult. The recruitment likelihood is perhaps better formulated as deviations from the mean, rather than deviations from the first recruitment. Parameterizing fishing mortality and recruitment as random-walk processes would reduce the number of effective parameters that are estimated.

Atlantic surfclam

The Atlantic surfclam stock assessment provides a credible scientific basis for developing fishery management advice. However, the delay-difference KLAMZ model used for the assessment does not use the wealth of age composition data that is available from surfclam surveys, so the assessment was considerably less informative than it might have been. A fully integrated age-structured model analysis of the surfclam data would allow direct estimation of reference points, better characterization of uncertainty, and greater consistency between the data and model assumptions.

The use of B_{1999}, the high point in the survey time series, as a proxy for B_0 is questionable given fairly stable and high landings since 1965 (and likely earlier than 1965). The most appropriate approach for redefining surfclam BRPs is through an integrated catch-age analysis – until that is done the existing BRPs are the best available.

Discussion relative to Terms of Reference

8. *Characterize the commercial and recreational catch including landings and discards.*

This term of reference was completed satisfactorily.

Commercial landings data for the stock assessment region (Exclusive Economic Zone or EEZ) were reported for 1965 to 2005. Landings data are relatively accurate because of the tracking system in the ITQ fishery. Recreational catch is low, so was not considered. Significant discarding occurred in the directed fishery in earlier years when there were size limits on landed surfclam (discards rates up to 37% of landings). Inclusion of discard data in the KLAMZ model has a very minor effect on biomass trends because fishing mortality is estimated to be very low. A 12% allowance, considered to be an upper bound, is made in the catch estimates for fishery induced natural mortality. The 12% “upper bound” is based on incidence and mortality of damaged clams in experimental studies of fishery induced mortality. The approach of using the “upper bound" estimate is not risk neutral.

The surfclam fishery has shifted with landing from the southern part of the EEZ (primarily DMV) dominant in the late 1970s and early 1980s and landings from the northern part of the EEZ (primarily NJ) dominant since the mid 1980s.

9. *Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.*

This term of reference was completed, but the estimates of uncertainty were inappropriately calculated given the formulation of the assessment model with a Bayesian prior. Fishable biomass and fishing mortality rates for 1981 through 2005 were estimated using the delay-difference KLAMZ model. Given the KLAMZ model does not estimate year-class strength, fishable biomass estimates will be more reliable than spawning stock or total biomass estimates which would require additional assumptions about selectivity and the maturation ogive. Bootstrap methods were used to estimate bias and uncertainty for KLAMZ model estimates, but Bayesian methods would be more appropriate.

For the surfclam assessment, KLAMZ was formulated with an informed prior distribution for the abundance scaling parameter (Q), implying a Bayesian formulation of the model. For Bayesian estimation, parameter uncertainty is estimated from the joint posterior distribution generally using either an MCMC or SIR algorithm. However, for the surfclam assessment a bootstrap procedure was used to characterize uncertainty and the KLAMZ model run only for MPD estimates (maximum of the posterior distribution) of the bootstrap samples. Thus the purported prior on the Q parameter did not function as a true Bayesian prior. The estimated uncertainty in the Q parameter and resulting biomass estimates obtained from the bootstrap approach will be considerably less than what would have been obtained had a fully Bayesian analysis been conducted. This conclusion follows from the observation that there is little information in the surfclam data and KLAMZ formulation about the value of the abundance scaling parameter Q (as demonstrated through some additional analyses during the meeting). Given little information in the data to modify the prior distribution, the posterior distribution of Q would be very similar to it's prior. For the bootstrap approach, each MPD estimate will tend to be close to the mode of the "prior", resulting in a narrower distribution of Q than would be obtained from a Bayesian analysis. Note that this distinction between the behavior of the bootstrap and the Bayesian estimates of uncertainty is dependent on having an informed prior and their being little information in the data/model about the value of the parameter.

As noted for the Ocean quahog assessment, the formulation of the KLAMZ model de-couples the relative abundance of recruits and the relative abundance of the fishable population. The abundance scaling parameter for recruits is estimated to be double that of the fishable abundance. While this model formulation allows for potential misspecification of the growth parameters, it also eliminates the ability to estimate the virgin stock abundance. An alternative approach would be to develop a prior on the ratio of recruitment and fishable abundance *Q*'s so the B_0 can be estimated.

The PATCH model is major improvement for estimating dredge efficiency from experimental data. While the estimates of survey dredge efficiency from multiple set-up tows/commercial depletion studies are fairly consistent, there are implicit assumptions in the analysis that could generate biased results (asymptotic selectivity with know size of full selection, known position of dredge and size of cells effected by dredge). Sensitivity to these assumptions should be investigated. There is some indication that not all assumptions of the PATCH model are met. The assumption that the dredges are size selective leads to a predicted decline in average size over each tow in a depletion experiment. However, this is not always observed, suggesting that the size selectivity assumptions may be incorrect.

Abundance estimates from KLAMZ model analysis are directly proportional to the efficiency estimate. A prior for the efficiency parameter is developed (CV of median estimates from bootstrapping experiment-specific estimates) and used as the prior for the abundance index scaling prameter (*Q*) in KLAMZ. For future assessments it would be useful to include uncertainty due to potential failure of assumptions in the PATCH model. Also, weighting efficiency estimates from each depletion experiment by a goodness-of-fit measure to obtain the *best* efficiency estimate may be preferable to using the median value across experiments.

Biomass estimates from the 2004 cooperative survey, solely based on commercial dredge efficiency estimates, provides some support that the survey dredge efficiency estimates are reasonably accurate.

10. *Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.*

This term of reference was satisfactorily completed, though significant improvements in redefining BRPs could have been made had an age-structured model been used for the assessment. The B_{msy} proxy, B_{1999}, was updated using results from the KLAMZ model stock reconstructions. The F_{msy} proxy $\left(F_{msy} = M\right)$ was not updated. The working group noted their skepticism with using B_{1999} as a B_{msy} proxy, and suggested that this reference point should be re-evaluated. Given an age-structured model was not used for the stock assessment, the redefined B_{msy} proxy and the current F_{msy} proxy are the best available at this time.

For the surfclam stock, survey age-composition data is available intermittently over the period 1981-2005 and fishery catch-at-length data is available annually over the same period. A substantial amount of ageing has been done, and ageing data appears to be reliable. The survey age-composition data shows periodic strong year-classes consistently through the surveys. Data required for an age-structured assessment are available, and an age-structured model would allow a fully integrated analysis of the stock ensuring consistency and links among all components. Estimates of reference points would be consistent with the assessment.

11. *Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).*

This term of reference was successfully completed.

12. *Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.*

This TOR was completed, though additional consideration should be given to the process of simulating recruitment for the projections. The KLAMZ model was used for the stock reconstruction, so it is an appropriate tool for simulating future stock trajectories under alternative catch scenarios. However, recruitments were treated as nuisance parameters in the KLAMZ stock reconstruction and an appropriate method for simulating future recruitments is not immediately clear.

For the stock assessment, recruitment in the KLAMZ model was simulated as a random walk time process. This is an acceptable approach for the stock reconstruction because there are gaps in the recruitment series that are being fitted (observations approximately every three years), and the time series approach restricts the effective number of parameters in the estimation. Recruitments for the projection phase of the KLAMZ analysis are simulated assuming autocorrelated random variation in the recruitment process with CV and autocorrelation from the empirical estimates (the recruitment series estimated for the stock reconstruction). While this generates patterns in the historic recruitment that match the intermittent observations, the patterns in projected recruitment differ vastly from what is known about surfclam recruitment patterns. The autocorrelated recruitment projections result in recruitment cycles of 50 to 100 years, whereas surfclam appear to have episodic strong year classes at 5 to 10 year intervals. Given relatively short projection periods (2 to 3 years), the lack of congruence between the historic and projection recruitment patterns is not of major concern. In a stock rebuilding or overfishing state, the issue would become important.

13. *If possible,*

    a. *provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and*

    b. *compare projected stock status to existing rebuilding or recovery schedules, as appropriate.*

This term of reference was successfully completed. As noted under TOR 3, estimating uncertainty using a bootstrap procedure is not appropriate given the specification of a Bayesian prior for the survey index proportionality parameter in the KLAMZ model. If marginal posterior distributions of quantities of management interest (eg. projected fishing mortality rates) were based on Bayesian estimation methods (eg. MCMC) modeling assumptions and the estimation process would be consistent.

14. *Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.*

This term of reference was satisfactorily completed. It is noted that a recommendation from a previous surfclam assessment, to develop a stochastic age-structured model for analyzing the surfclam data, was not completed. Development of an age-structured assessment model should be a priority for future surfclams assessments, as it would allow an internally consistent and integrated analysis of all available data.

Additional comments

Significant improvement in PATCH model estimates of survey dredge efficiency might be made if multiple experiments were simultaneously analyzed. Ideas include: estimating a single efficiency parameter for each commercial vessel and a single efficiency for the survey vessel; estimating depth and/or sediment size covariate; model partial selectivity and fit to all size frequency data to estimate survey and commercial selectivity.

References

FAO. 2002. Abstracts of papers presented at the Reykjavik conference on responsible fisheries in the marine ecosystem. Reykjavik, Iceland, 1-4 October 2001. FAO Fisheries Report No. 658, Supplement. FAO, Rome. 90 p.

Gedamke, T. and J.M. Hoenig. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Trans. Amer. Fish. Soc. 135:476-487.

Hoenig, J.M. 1987. Estimation of growth and mortality parameters for use in length-structured stock production models. *In* D. Pauly and G.R. Morgan (eds.) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468 p. International Center for Living Aquatic Resources Management, Manila, Phillipines.

Kimura, D.K. and S. Chikuni. 1987. Mixtures of empirical distributions: An iterative application of the age-length key. Biometrics 43:23-35.

Stokes, T.K., D. S. Butterworth, R.L. Stephenson (Co-Conveners) and A.I.L. Payne (Guest Editor). 1999. Confronting uncertainty in the evaluation and implementation of fisheries-management systems. ICES Journal of Marine Science, 56: 795-796.

Report on SARC 44

November 28–December 4, 2006

Woods Hole

Massachusetts

P.L. Cordue

Fisheries consultant

New Zealand

***For* University of Miami Independent System for Peer Review**

4 December 2006

Executive summary

A Stock Assessment Review Committee (SARC) meet from November 28 to December 4, 2006 to review assessments of ocean quahog, Atlantic surfclam, and a skate complex. The CIE appointed Panel comprised of two reviewers and a chair. This is the report of one reviewer and it should be read in conjunction with the other review report and the summary report.

The working group (WG) generally fulfilled their TOR for the quahog assessment. The exception was with regard to redefining reference points. They pointed out the inappropriateness of $F_{25\%}$ as a proxy for F_{MSY} but failed to propose an alternative. The assessment methods were appropriate and the status determination is very robust for the mandated single EEZ stock. The key issue for ocean quahog was outside the TOR of the WG and is outside the TOR of the review. Ignoring evidence of potential serial depletion is imprudent. I recommend that the industry work with scientists and managers to develop a demonstrably sustainable fisheries management plan for ocean quahog.

There are seven species in the skate complex. The task of assessing so many species simultaneously necessarily leads to a formulaic approach to data analysis, presentation, and assessment. The skate WG relied heavily, but not exclusively, upon a recently published method for estimating current fishing mortality from trawl survey length frequency data. The method is unproven and contradicts “accepted wisdom”. It should not be used for status determination in stock assessments without extensive simulation testing to determine when it is able to provide reliable estimates. Other methods used by the skate WG to provide mortality estimates are also unproven or intrinsically unreliable.

The skate WG failed to fulfill several of their TOR, partly due to their reliance on published, but unproven, or unreliable methods. The catch histories for skates will never be well determined, but a much greater effort should be made in the future to obtain at least qualitative species-specific estimates (aiming for plausible upper and lower bounds on removals). The proposed new reference points used promising methods, but both methods relied on estimates of recruitment obtained by cohort slicing. Until the methods are shown to be robust to the errors associated with this approach the existing reference points should be retained. The consequences of doing this are unfortunate for winter skate as it is declared overfished, despite little or no evidence to support this conclusion.

The surfclam assessment is inappropriate given the available data. A fully integrated age-structured model should be used. However, the WG did generally meet their TOR. They again identified problems with reference points but made no alternative proposals. A complete catch history should be made available for future assessments.

Generic approaches should be developed with regard to biological reference points and stock assessment methods in general. The assessment and management of all of the species in this SARC could benefit from the development of rigorous simulation tools which could be used to test estimators and management procedures.

Introduction

A Stock Assessment Review Committee (SARC) met from November 28 to December 4, 2006 to review assessments of ocean quahog, Atlantic surfclam, and a skate complex. The format of the SARC meeting and the schedule for report production required that I essentially complete my report while on site at Woods Hole. This differs from other processes that I have been involved in, where reviewers are able to “pause for thought” before writing their reports. If I had been given more time to write this report, I would have provided more detail in some areas, with more suggestions and recommendations. My choice of words would also have been more carefully considered, but my essential conclusions would be unchanged.

The TOR for the review required that I comment on whether the WGs have meet their TOR and also whether I support the existing and proposed reference points. I have set out my findings below under each of the WG's TOR for each species/complex.

Review findings

Given the time restrictions for producing this report, I have concentrated on the skate assessment for which I was “SARC leader”. The assessment methods used for the quahog and surfclam assessments were very similar to each other. Therefore, the assessments share common strengths and weaknesses. I have made more detailed comments with regard to quahog – simply because it preceded surfclams in the agenda.

I have also concentrated on criticism rather than praise, partly due to time restrictions and partly because that is the nature of review. I would like to counter that by noting that the scientific staff involved in these assessments, and in the collection of data supporting these assessments, show the dedicated and hard-working ethic that, in my experience, typifies fisheries research personnel in most parts of the world.

A. Ocean quahogs

1. Characterize the commercial catch including landings and discards.

This was done satisfactorily. Catches were negligible before 1970 and there is no recreational catch. An allowance of 5% is made for discards and incidental effects. This is so minor and has so little basis that a basecase assumption of 0% may be more appropriate.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty

of those estimates. If possible, also include estimates for earlier years.

At a basecase level this was done satisfactorily and the stock status determination is very robust – so the TOR was met.

However, the KLAMZ assessment could have been better.

The assumptions of the main model runs could have been presented more explicitly. There is no linkage between the estimated mean recruitment and the absolute biomass levels in the model. That is, the model was configured to obtain the best fit to the data while ignoring the relative scales of mean recruitment and biomass. This approach is consistent with allowing a regime shift in that the virgin mean recruitment lead to a large accumulated biomass, which has since declined due to fishing and a shift to much lower mean recruitment (the estimated recruitment was very low for most regions). This approach assumes that all of the information on biomass is contained in the efficiency estimates (and essentially ignores the information in the relative abundance indices).

An alternative assumption is to require an explicit link between virgin biomass and estimated recruitment and to assume that there is information on biomass contained in the relative abundance indices. This approach could be applied to the full data set or just to the relative abundance indices. A run incorporating the link and using only the relative abundance indices was done during the SARC and resulted in biomass estimates half those of the base run and *F*s double those of the base run (for New Jersey). The base run and this alternative run represent two extremes of a continuous spectrum.

The characterization of uncertainty was therefore deficient. The alternative assumptions should have been discussed and a variety of runs presented.

There is also some scope for comparing the estimated recruitment with the survey data and this was not done. For example, estimates of recruitment into the fishable biomass could have been obtained from the survey data for some length class at about the assumed recruitment length (e.g., the length class expected to recruit over the next 5 years). This is problematic given the very slow growth of quahog, but the idea would be to check whether estimated recruitment was *plausible* given the data (and growth assumptions) rather than to provide a definitive diagnostic.

There was much discussion during the SARC on the treatment of the clam survey data and how to deal with missing cells (due to un-surveyed strata in some years). An ad hoc approach was used to fill the missing cells. First, data were “borrowed” from adjacent cells, but then further gaps were filled by application of a GLM to the original data (i.e., not the borrowed data). The calculation of variances for annual surveys with missing cells was inappropriate as no adjustment was made when data were borrowed. Indeed, the more borrowing there was the lower the associated variances were made (i.e., borrowing from the earlier survey and the later survey resulted in double the sample size that there would have been had the stratum actually been surveyed).

A formal approach should be adopted for treatment of the survey data which rigorously deals with empty cells. The best approach may be a GLM which is aimed at extracting the year effects for use as relative abundance indices. If this approach is adopted, care will need to be taken in its application – for instance in checking for year-area interactions.

The use of a patch-model in the analysis of depletion experiments was innovative (Rago and Weinberg 2006). The method appears worthwhile but it still requires some testing to make sure it is robust to deviations from assumptions (e.g., errors in dredge position; choices of strip widths). Also, the method used to choose a best estimate of efficiency is somewhat ad hoc (i.e., use of the median). Consideration should be given to how the whole depletion dataset (i.e., over all experiments) could be analyzed to provide the best estimates of efficiency and selectivity (i.e., there is no need to choose a cut-off length). The consideration should then be extended into the design of future depletion experiments.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

This TOR was almost completed. The existing reference points were updated and improved using new analyses and data. The WG noted problems with the current definition of the reference points, in particular F_{thr}, but did not recommend more appropriate alternatives.

The existing reference points have an internal inconsistency with the biomass referring to the whole stock and the fishing mortality referencing only the exploitable stock. This may have some practical advantages if one is trying to prevent “localized depletion” (albeit on a rather large scale, given that only GBK is closed). However, the inconsistency is unnecessary. A consistent approach should be adopted which is supportive of stated management objectives

The use of B_{1978} as a proxy for B_0 is somewhat perplexing when a full catch history is available with catches beginning in 1967. Clearly, a proxy for B_0 is not needed. It is also clear, as noted by the WG, that $F_{25\%}$ is unsuitable as a proxy for F_{MSY}. Rather than recommend a replacement reference point I suggest that the definition of reference points be considered for quahog in a generic setting. A tier structure where the reference points are defined on the basis of available information has strong precedent in the U.S. (e.g., North Pacific Fisheries Management Council). The model based approach of Francis (1992) should also be considered as a standard approach for U.S. stocks and could be used in conjunction with a tier structure or not.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

This was done.

5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.

This was not discussed, but since KLAMZ was used to do the modeling it was entirely appropriate to use it to do the projections. Adequate information was supplied to allow calculation of TACs and TALs from example projections.

6. If possible,

a. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and

b. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

The TOR was completed. Example projections were done at various constant catch levels and also for a constant F set equal to the estimated $F_{0.1}$. However, the projections were all for the single base run. As discussed above, there were other equally valid runs which could have lead to somewhat different projection results (in particular, runs with different levels of mean recruitment).

7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

This TOR was completed. A good summary of the progress on previous recommendations is provided. Then a list of new recommendations is given. This appears to be outside of the TOR but the terms of reference appear deficient in that new recommendations are not requested. I note that the new recommendations are not prioritized. Perhaps the research planning is done in another process, but it would seem to be imperative to prioritize the recommendations. Further, such recommendations should be much more cohesive and structured.

B. Skate species complex

1. Characterize the commercial and recreational catch including landings and discards.

This TOR was not completed.

A good summary of landings and discards is provided except that some idea of the scale of the "industrial fishery during the 1950s and 1960s" would have been useful. There was an adequate discussion of the issues which make construction of species specific catch histories problematic. However, without species specific catch histories there is little that can be done to fulfill TOR 2.

Although the construction of catch histories cannot be done with great certainty, it is such a fundamental input that every effort must be made to make the determination. The objective would not be to obtain precise estimates. Rather, the aim would be to provide upper and lower bounds on catch histories by species. The first step is to bound the total level of landings and discards. The final step is to split the catches by species, under different hypotheses and assumptions. A very difficult and time consuming task, but it is conceptually feasible.

I fully understand why this was not done. It was probably not feasible in the available time and the WG believed that they had a viable method to use for status determination.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

This TOR was not completed satisfactorily.

No estimates of *absolute* spawning stock biomass or total stock biomass are provided. Survey estimates are presented and discussed but these are *relative* estimates. It would have been possible, and desirable, to provide some "ball park" estimates of absolute biomass by developing species specific priors on the trawl survey proportionality constants (q). Each trawl q has three components: areal availability, vertical availability, and vulnerability (e.g. see Cordue 1996). It is a relatively simple exercise, for a sub-committee of a WG, to bound each of these components, and hence derive, for each species, an upper and lower bound for each trawl survey q. In some cases the bounds may be so wide as to be of little use – but that determination is useful in itself. In the case of some skate species, for some surveys, the bounds may actually be relatively narrow, since vertical availability and areal availability would be close to 1. There may be data available to help with this process (e.g., a monkfish depletion experiment which caught "lots of skates").

The bootstrapping method used to characterize the *precision* of the trawl survey indices was adequate.

Methods which use length frequency data, in conjunction with growth parameters, to estimate total mortality are available in the literature (Beverton Holt 1956, Hoenig 1987, Gedamke and Hoenig 2006). However, the use of these methods in stock assessments which support management decisions is rare. The methods use data which contain very little information on total mortality. What information there is, relates not to *current* total mortality, but to a weighted average of mortality which applied over the life of larger (and generally older) animals in the population. The information is confounded by numerous factors, and any estimates obtained by these methods must be treated with extreme caution – and so are not suitable for use in status determination.

The WG relied, for several species, on the recently published method of Gedamke and Hoenig (2006). It is unfortunate that the authors make claims for the reliability of their estimation method which are not substantiated by their paper and which contradict "accepted wisdom". The method avoids the equilibrium assumptions of earlier methods, but makes many other assumptions which are necessarily violated in practice. Of course, the same is true of any model based estimator: "all models are wrong, some are useful". The usefulness of the Gedamke-Hoenig estimator needs to be demonstrated before it can be accepted in stock assessment for status determination. Below, I summarize what needs to be done before I would be satisfied with the methods use in this context.

The minimum requirements are: (i) the method is shown to be reliable under specific criteria; and (ii) the criteria are shown to be satisfied for the particular assessment to which it is applied. The first step was not achieved in the paper (except that the estimator was derived under specific assumptions which could never be met in practice). The second step was not addressed by the WG.

The authors derived a maximum likelihood estimator under certain assumptions. There is a theorem that states that under fairly general conditions a maximum likelihood estimator is *asymptotically* the minimum variance unbiased estimator. However, the authors do not show that their estimator satisfies the conditions of the theorem. It may well do, but that would be incidental because in practice large sample properties of estimators are irrelevant because sample sizes are necessarily limited. In the case of length data, one must note that the effective sample sizes applicable to length frequencies are generally of the order of the number of tows sampled rather than the number of individual animals measured.

The authors claim "reliability" for their estimates on the basis of three tests for which they report results for a goosefish assessment. These are: a sensitivity test with regard to externally estimated growth parameters; a retrospective analysis (which they phrase as a test of "predictive" power); and a comparison of estimates with another method. They fail to recognize that none of these methods address the properties of their *estimator*, but only relate to properties of their *estimates*. There are only two reliable methods for determining estimation properties: analytical derivation, or characterization through simulation.

Analytical derivation is rarely applicable in fisheries stock assessment applications because the estimators are too complex (relying on the minimization of likelihoods with many components). Simulation methods require the use of an “operating model” (a model of “reality”) to generate simulated data; and the application of the estimator to the simulated data to produce simulated estimates. A large number of simulations need to be performed to approximate the distribution of the estimator at many points in the operating model parameter space. In data poor situations, such as skate, it is crucial to explore the operating model space extensively to ensure that the “real” situation is covered by some of the simulations. Of course, there is always the difficulty that the “real” situation is unknown – which is why the assessment is being done in the first place.

To determine when the Gedamke-Hoenig estimator is reliable it is necessary to set up an operating model which incorporates complexities such as those listed below:

- Recruitment variability (independent lognormal and correlated lognormal options)
- Growth variability (spatial and temporal)
- A variety of trajectories for fishing mortality (e.g., increasing, decreasing, stepped)
- Realistic fishing selectivity (option for temporal variation)
- Realistic survey selectivity (option for temporal variation)
- Growth and natural mortality parameters which deviate from M = k
- Variability or perhaps shifts in natural mortality

When using the operating model to test the method for a specific application (e.g., a skate species using a fall survey series of length frequencies) it is necessary to use realistic effective sample sizes and appropriate distributional assumptions. For example, when a scaled length frequency from a stratified trawl survey is used, it is important to try to capture the real distribution of the mean length statistic (which will not, in general, follow a normal distribution – as assumed in the estimation model).

When running the estimation model (during the simulation study) it is crucial to explore a wide range of model-misspecification errors. For example, errors in the assumed growth parameters, error/bias in the length at full selectivity. The scope of an extensive simulation study can be quite daunting. However, it is the only way to determine if there are any *realistic* circumstances under which the Gedamke-Hoenig estimator can provide “reliable” estimates. It is already apparent that for any situations where *F* is changing rapidly that the estimator will fail badly as its estimates of terminal *F* will be badly biased high or low depending on the direction in which *F* is changing (i.e., assuming that the estimator is unbiased for average *F* over the period of the final “stanza”).

The length-tuned model estimation of fishing mortality applied to thorny skate was an imaginative use of a forward projection model. However, the model appears to be vastly over-parameterized given the paucity of data. As with the Gedamke-Hoenig estimator, there may be some *realistic* circumstances under which it could produce reliable estimates – but such circumstances are yet to be demonstrated.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

The WG considered a variety of approaches for revising the existing ad hoc reference points. Two promising approaches were used to derive proposed reference points for five of the seven species. The length-based yield per recruit modeling and the stock-recruit fitting method (Brodziak and Legault 2005) are worth pursuing. However, for derivation of the biomass reference points both methods rely on the estimates of recruitment derived from cohort slicing of length frequency distributions. Cohort slicing may work well when consistent length modes are visible in the length frequency distributions, but for most of the skate species this is simply not the case. In the case of barndoor skate, where so few fish are caught, it is clear that little information, if any, on recruitment can be obtained from the available data.

At this stage, I recommend that the existing reference points be retained. I am sufficiently uncomfortable with the applicability of cohort slicing in the determination of the revised reference points that I cannot support them at this time. I encourage further work to evaluate the robustness of the methods to cohort slicing in the presence of relatively low sample sizes and the absence of modes in the observed length frequencies.

The existing reference points are ad hoc but they are the best available at this time. Better reference points can be developed but it is for an expert group to pursue this in a timely fashion, rather than for this SARC Panel to hastily recommend other ad hoc or poorly researched alternatives. However, I do note that an unfortunate consequence of accepting the existing reference points is that "overfishing" is determined to be occurring for winter skate. An examination of Figure B2.10 shows why this happens – a "high point" in 2002 drops out of the moving average while the "low point" in 2005 comes in. Is this an early warning that the stock size may be moving in the wrong direction? Almost certainly not. The available data are not adequate to provide such a determination.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

This was done.

5. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC-reviewed assessments.

This was completed satisfactorily. There are brief comments in the assessment report on each of the previous research recommendations. Several of them had been actively pursued and considerable progress was made. There is little indication of the priorities for remaining incomplete recommendations and the TOR themselves seems remiss in not requesting new recommendations and a prioritization of the new and incomplete exisiting recommendations.

6. Examine the NEFSC Food Habits Database to estimate diet composition and annual consumptive demand for seven species of skates for as many years as feasible.

This TOR was more than satisfactorily completed. There was a considered and extensive analysis including consideration of a range of trawl survey proportionality constants to estimate absolute annual consumptive demand.

C. Atlantic surfclam

1. Characterize the commercial and recreational catch including landings and discards.

This TOR was completed during the SARC as some mislaid discard data was rediscovered (there had been length restrictions at some stage which had resulted in substantial discarding for a period – but the relatively short duration of the period meant that inclusion of the discards in the model runs had no effect on results).

There is a good summary of landings from 1965, but there is no mention of landings prior to this date. Since, the fishery had total landings of about 20,000 t starting in 1965 there were clearly some earlier catches. The full catch history should be constructed as accurately as possible, to ensure it is available for future stock assessment modeling.

The use of 12% to account for incidental mortality (including discard mortality) is inappropriate given that it is an explicitly derived upper bound. A “best” estimate should be used to maintain “risk neutrality” in the stock assessment.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years.

This TOR was completed.

The methods used were very similar to those for ocean quahog. There are the same problems with the ad hoc method of filling empty cells in the survey strata, including the disconnect between the estimated mean recruitment and the starting biomass, and the consequences for projections.

However, there are some important differences between the surfclam and quahog assessments. There is an extensive and apparently reliable set of age data available for surfclams. The available data is ideal for use in a fully integrated age structured model. Assessment runs could be done assuming non-equilibrium or equilibrium starting conditions. I note the recommendation from the previous SARC to move to an age structured assessment.

I add a technical note on the characterization of uncertainty. This was done using bootstrapping in the presence of a prior on the swept-area biomass q. By using bootstrapping, none of the uncertainty in the q prior is incorporated into the bootstrap distribution. It would be better to simply call the prior a penalty function which is used for stabilizing the point estimates. I would prefer to see either a pure frequentist approach or a pure Bayesian approach.

3. Either update or redefine biological reference points (BRPs; proxies for B_{MSY} and F_{MSY}), as appropriate. Comment on the scientific adequacy of existing and redefined BRPs.

This TOR was not fully completed.

As for quahog, the WG discussed some of the inadequacies of the current reference points but did not recommend any alternatives. The use of B_{99} as a proxy for virgin/unfished biomass is perplexing. In such a data-rich setting there is no need for proxies for virgin biomass, B_{MSY}, or F_{MSY}. It is simply a matter of using all of the available data in conjunction with a fully integrated stock assessment model. I accept that there is a problem with some of the concepts implicit in the legislation given the biological reality of surfclams. However, the assessment is done in the context of the legislation, so the concepts are a given.

4. Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 3).

TOR completed.

5. Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, and for computing TACs or TALs.

The projections were done using KLAMZ which was appropriate, but as with quahog there was the same lack of sensitivity analysis with regard to alternative assumptions which lead to different levels of mean recruitment.

6. If possible,

> **a. provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies and**
>
> **b. compare projected stock status to existing rebuilding or recovery schedules, as appropriate.**

As discussed under TOR 5 and for quahog.

7. Review, evaluate and report on the status of the SARC/Working Group Research Recommendations offered in recent SARC reviewed assessments.

This TOR was completed. As for quahogs I note that new recommendations are not prioritized or organized into a cohesive structure.

Additional comments

All of my additional comments are in the form of recommendations. Some of the issues fall outside of the TOR of the WGs and of the review panel, but they were all raised during the SARC meeting.

Recommendations

For all species I recommend that some generic approaches to determining reference points be adopted. The current ad hoc approach where each species has their own reference points, defined and calculated by all manner of methods, is far from ideal. A tier based structure could be considered (e.g., North Pacific Management Council). The model based method of Francis (1992) could be used as part of a tier structure or not.

Also, for all of the species there is a need for the development of operating models.

Quahogs

The most serious issue for quahog fisheries management is serial depletion. This is outside of the TOR for the WG and for the review. However, ignoring the issue does not make it go away. The WG are well aware of the issue and despite their TOR do present relevant assessment results which are suggestive of potential serial depletion (e.g., regional assessments, greatest depletion in the south, trends in LPUE by ten minute squares, see Figure A15). It is imprudent to manage quahog as a single EEZ stock (excluding Maine). There is a substantial literature on sustainable management of sessile organisms through rotational harvesting. Such management plans should be considered for quahog.

Construct quahog-type operating models for use in management strategy evaluation. Scientists and managers should work with industry to develop a sustainable fisheries management plan.

Skates

A serious effort should be put into reconstructing the full catch histories by species. The objective is to obtain plausible bounds on total removals, rather than precise estimates. This will enable simple forward projection statistical models to be fitted to the relative abundance indices under alternative catch history assumptions. Put into a Bayesian setting, it would also allow prior information on the trawl survey proportionality constants to be utilized. Finally, such models implicitly provide MSY based reference points in terms of $\%B_0$ for any specified harvest strategy and/or definitions of risk (e.g., see Francis 1992). Extensive work is required to produce catch history ranges but the potential reward is an integrated stock assessment complete with reference points and status determination.

Construct skate-type operating models for use in a range of simulation studies – as discussed below.

In the absence of catch histories it may be possible to construct some robust management strategies based on the trawl survey abundance indices and length frequencies. A range of survey based reference points, including the existing ones, should be tested in an operating model framework. That is, set up skate-like operating models, and perform a management strategy evaluation, for the existing reference points and over-fishing definition (or more generally the FMP).

The work on determination of alternative reference points using cohort slicing in conjunction with stock recruitment model fitting or length based yield per recruit methods is also worthwhile pursuing. The key to using these methods is demonstrating that they are robust to inherent errors in cohort slicing (e.g., through simulation studies).

The utility of the Gedamke-Hoenig estimator is yet to be proven in a status determination setting. Simulation studies are required to determine the scope of its utility. I believe it will be very limited – but would be happy to be proven wrong (to be able to successfully manage fisheries just on the basis of trawl survey length frequency data would be remarkable).

Determination of stock structure is an essential pre-requisite to sensible stock assessment and should not be neglected.

Surfclam

The continued use of the KLAMZ model for basecase surfclam stock assessment is strongly discouraged. This is a data-rich situation and it is desirable to move to a fully integrated age-structured forward-projection model. KLAMZ should be used in the next assessment, but only for the purpose of “building the bridge” between the old and new assessments.

Serial depletion appears to be an important issue for surfclams as well as ocean quahogs (but not as pressing).

Construct surfclam-type operating models for use in management strategy evaluation. Scientists and managers should work with industry to develop a sustainable fisheries management plan.

References

Cordue, P.L. 1996: A model based method for bounding virgin biomass using a catch history, relative abundance indices, and ancillary information. New Zealand Fisheries Assessment Research Document 96/8. 48 p.

Francis R.I.C.C.1992: Recommendations concerning the calculations of maximum constant yield (MCY) and current annual yield (CAY). New Zealand Fisheries Assessment Research Document 92/8.