Final Report

Summary Report of the 58th Northeast Regional Stock Assessment Review Committee (SARC 58)

Members of SARC 58

Robert J. Latour, Chair
Catherine Dichmont
Stewart Frusher
Ian Jonsen

Prepared for the Northeast Regional Stock Assessment Workshop
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
Woods Hole, Massachusetts

Meeting dates: 27 January – 31 January, 2014
Report date: 17 February, 2014

Contents

1. Introduction

1.1 Background

The 58th Stock Assessment Review Committee (SARC) met in Woods Hole, MA from 27 January – 31 January, 2014 to review stock assessments for butterfish (*Peprilus triacanthus*), golden tilefish (*Lopholatilus chamaeleonticeps*), and northern shrimp (*Pandalus borealis*). The review committee was composed of Dr. Robert J. Latour (MAFMC SSC and Virginia Institute of Marine Science, Chair) and three scientists affiliated with the Center for Independent Experts: Dr. Catherine Dichmont (CSIRO), Dr. Stewart Frusher (University of Tasmania), and Dr. Ian Jonsen (Dalhousie University).

The SARC was assisted by the NEFSC SAW Chairman, Dr. James Weinberg, his staff, and Dr. Paul Rago (NEFSC). Supporting documentation for the butterfish assessment was prepared by the NEFSC Coastal/Pelagic Working Group (CPWG), and presentations were made by Drs. Charles Adams, John Manderson, and Timothy Miller. Technical documents for the golden tilefish assessment were prepared by the NEFSC Demersal Working Group (DWG), and presentations were made by Mr. Paul Nitschke. Written material for the northern shrimp assessment was prepared by the ASMFC Northern Shrimp Technical Committee (NSTC), and presentations were made by Ms. Kelly Whitmore (MA) and Drs. Anne Richards (NEFSC) and Katie Drew (ASMFC). The CPWG was chaired by Mr. Gary Shepherd (NEFSC), the DWG was chaired by Dr. Mark Terceiro (NEFSC), and the NSTC was supported by Dr. Larry Jacobson (NEFSC), Dr. Yong Chen (University of Maine), and Jie Cao (University of Maine). The contributions of all of these scientists to the SARC process are gratefully acknowledged.

1.2 Review of Activities

Approximately two weeks before the meeting, assessment documents and supporting materials were made available to the SARC panel via an ftp server on the NEFSC website. On the morning of the meeting, the review panel met with Drs. Weinberg and Rago to discuss the meeting agenda, reporting requirements, and meeting logistics. During the SARC meeting, all documents were made available electronically and in print. The meeting opened on the morning of Monday 27 January with welcoming remarks by Drs. Weinberg and Latour. Following introductions, the remainder of Monday was devoted to presentations of the butterfish assessment. Tuesday morning was spent on presentations of the golden tilefish assessment, while Tuesday afternoon was focused on the northern shrimp assessment. Final follow-up discussion for the butterfish and golden tilefish assessments took place on Wednesday and the first few hours of Thursday morning were devoted to follow-up discussion for northern shrimp. The remainder of Thursday was spent editing the Assessment Summary Reports for all three stocks and Friday the meeting was restricted to only the SARC Panelists for report writing. The tone of the meeting was collegial, and considerable time was devoted to facilitate dialog among review panel members, working group scientists, NEFSC population dynamics biologists, MAFMC and ASMFC staff, and industry representatives.

1.3 SARC Process

The review panel was able to reach consensus on all three assessments. Since the last assessments of each species (2009 for butterfish, 2009 for golden tilefish, and 2007 for northern shrimp), considerable research advancements have been made with regard to each assessment, and the incorporation of new information resulting from these research studies led to improved understandings of the population dynamics of all three species. The assessments conducted by the CPWG, DWG, and NSTC were very thorough, and it was apparent that each group devoted

significant time and effort to data analysis, model fitting, evaluation of uncertainty, and report preparation.

The Panel agreed that most all of the Terms of Reference (ToRs) for butterfish were met and that the assessment results could be used as a basis for management advice. ToR 4 was partially met and the Panel encouraged the CPWG to more fully take advantage of the rich ecosystem and multispecies modeling efforts that have been ongoing in the NW Atlantic over the past decade (e.g., Atlantis, EwE, MSVPA, MS Statistical Catch-At-Age) to gain deeper insights about butterfish stock dynamics. Commercial landings (catch+discards) and survey indices were well characterized, although discard estimation for some years was imprecise. Much of the uncertainty in estimated discards noted in the previous peer-reviewed butterfish assessment (SARC 49) was alleviated by fixing the starting year of the assessment at 1989. The results of an impressive butterfish habitat preference study were presented during the meeting, and although the temporally varying habitat index was not formally incorporated in the accepted base assessment model, the Panel strongly supported continued research in this area. A modified Age Structured Assessment Program (ASAP) model, denoted ASAP3, was used as the assessment tool for stock status determination and management advice. Success of the butterfish ASAP3 model largely rested on an externally estimated value of the NEFSC fall offshore survey catchability coefficient (q). Although this approach is somewhat unusual, there was consensus that the CPWG was thoughtful, diligent, and comprehensive in how it generated the estimated q. The Panel also agreed that the approach taken was conservative since the estimated q hinged on the assumption that the survey cruises conducted from the *FSV Henry B. Bigelow* were 100% efficient at sampling butterfish. The accepted ASAP3 based model configuration was a slight variation from the version put forth by the CPWG. Specifically, the Panel recommended excluding the temporally varying habitat index on the basis of parsimony, and only including the fall survey indices because of notable contrasting patterns in abundance among spring and fall abundance trends. The alternative base ASAP3 model yielded low estimated fishing mortality rates, and relatively healthy estimated total stock abundance. There was consensus among the Panel members that overfishing was not occurring and that the stock was not overfished.

The Panel agreed that most of the ToRs for golden tilefish were fully met and that the assessment results could be used as a basis for management advice. ToR 3 was partially met and the Panel encouraged the DWG to more fully explore environmental impacts on golden tilefish stock dynamics. Three independent assessment models were presented, and the Panel was very appreciative of the significant effort put forth to develop each assessment model. Commercial golden tilefish landings (directed catch) were thoughtfully characterized. The DWG relied on commercial landings-per-effort (LPUE) as an index of population abundance, and while this is not ideal, there were very few other options. As in the previous peer-reviewed golden tilefish assessment (SARC 48), the Panel reiterated concerns about the coarse definition of fishing effort and the spatial limitation of the fishery relative to the distribution of golden tilefish. Obtaining more refined trip-level information from commercial fishers, the standardization of market category designations among dealers, and the development of a statistically defensible industry-based survey of golden tilefish remain important areas of future research. Data on the age composition of recent landings were vital to the assessment because they:

- Confirmed the propagation of strong year classes through the population evident in the LPUE indices via modal analyses.
- Facilitated configuration of an ASAP model.
- Provided evidence of dome-shape fishery selectivity.

The Panel accepted the base ASAP model configuration put forth by the DWG with one minor adjustment to how projections were formulated. The assessment model results indicated that full fishing mortality has been decreasing over the last decade and the spawning stock biomass (SSB) has been increasing. There was consensus among the Panel members that overfishing was not occurring and that the stock was rebuilt.

The Panel agreed that not all of the ToRs for northern shrimp were met and that the results of the analytical assessment models should not be used as a basis for management. In lieu of an accepted assessment model, the Panel recommended basing northern shrimp management on observed patterns in the northern shrimp catches, survey indices, and potentially commercial CPUE. Three independent assessment models were presented by the NSTC, and the Panel was very appreciative of the significant effort put forth to develop each assessment model. Northern shrimp commercial landings were well characterized and the survey indices appeared to be informative. The decision to reject each assessment model put forth was not taken lightly, however, consensus of the Panel was that each model had serious problems. The University of Maine (UME) model is a statistical stage- and size-structured approach which conceptually is a good match for northern shrimp. However, the fit of the base model to the survey indices and survey size composition data was very poor, and the Panel concluded that this lack of fit compromised the model's ability to adequately reconcile northern shrimp stock and fishery dynamics. The Catch Survey Analysis (CSA) model is a two-stage approach that is also a conceptually good match for northern shrimp. Examination of the base model results revealed strong patterns in the residuals of fits to the survey indices. During the meeting, the Panel encouraged several variations of the base model that reflected adjusted CVs on the catch and survey time-series and different likelihood weighting schemes. Ultimately, only marginal diagnostic improvements were achieved, but of major concern was that substantially different results could be realized depending on the likelihood weights (overfishing versus not overfishing). The chosen likelihood weights essentially predetermined the resulting stock status determinations and management advice. The Panel therefore concluded that the CSA northern shrimp model was not robust and could not be used as a basis for management. The ASPIC model was applied primarily for heuristic and confirmatory purposes. However, recent northern shrimp recruitment has been highly variable, and the inability of ASPIC to adequately capture strong year classes as they propagate through the population led the Panel to recommend discontinued use of this assessment approach for the foreseeable future.

Special Comment, Northern Shrimp: There was consensus that NSTC's efforts were rigorous and that the analysts did everything possible to develop a reliable northern shrimp assessment. Model performance can sometimes break down with the addition of new data, and unfortunately, the post 2006 data created unforeseen technical problems within the CSA and UME modeling frameworks.

Special Comment, SARC Process: The volume of material presented for each species was quite substantial, and each assessment team should be commended for their efforts. However, the Panel felt that insufficient time was allotted to thoroughly review all aspects of the material presented, which unfortunately hinders the feedback component of a collegial peer-review. The Panel therefore recommended that more consideration be given to the breadth and depth of assessment material likely to come forward when developing the SARC schedule and associated meeting agendas.

2. Review of Butterfish

2.1 Background

The butterfish is a short-lived fast growing schooling pelagic teleost that can be found along the eastern seaboard from Newfoundland to Florida, but occurs in relatively high abundance from the Gulf of Maine to Cape Hatteras, NC (Bigelow and Schroeder 1953). Few individuals survive to ages older than 3 years and most are sexually mature at 1-2 years of age (Cross et al. 1999). Butterfish overwinter along the edge of the continental shelf and migrate inshore during spring. In summer and fall months, butterfish are distributed over the entire mid-Atlantic shelf, including in bays and estuaries. Butterfish are broadcast spawners and demonstrate a protracted spawning season ranging from May to October with peak activity occurring in July (Cross et al. 1999). Various teleosts, elasmobranchs (sharks), marine mammals, and seabirds prey on butterfish (Duffy 1988, Cross et al. 1999). Management of the butterfish fishery falls under the auspices of the Mid-Atlantic Fishery Management Council's Atlantic Mackerel, Squid, and Butterfish Fishery Management Plan.

2.2 Evaluation of the Terms of Reference for Butterfish

ToR 1. *Characterize the commercial catch including landings, effort and discards by gear type. Describe the magnitude of uncertainty in these sources of data.*

This ToR was met. Estimates of commercial catches (directed landings + discards) were provided for 1887-2012, although the assessment was based on catch data from 1989-2012. The choice of 1989 as the starting year in the assessment was a notable change from the previous butterfish assessment, which had a starting year of 1973. There are pros and cons to this decision. On the one hand, excluding historical catch data ignored potentially valuable information about the stock, especially since the highest landings on record occurred in 1973. However, the estimated catches prior to 1989 were highly uncertain because of no observer information to estimate discards and no measures of precision of the foreign fleet catches. Ultimately, the Panel agreed with the CPWG's decision to truncate the catch series and use 1989 as the starting year.

During the specified time period for the assessment model, directed butterfish landings have been very low or nonexistent due to highly reduced quotas (~500 mt/yr) or the complete absence of a directed fishery. Consequently, contemporary landings are primarily the result of discards in the longfin squid (*Doryteuthis pealeii*) and other trawl fisheries. The coefficients of variation (CVs) associated with butterfish catches were reasonable (0.3-0.4) for several years in the time-series, with occasional high values (> 0.8). Since the last assessment, there has been a considerable increase in the sampling intensity of the observer program for both the New England and mid-Atlantic regions, and the Panel agreed that this trend should continue into the foreseeable future. Adequate observer coverage, particularly in the mid-Atlantic region, is vital to the assessment of butterfish. The Panel agreed that recreational catches of butterfish were negligible and that the use of commercial effort data was not appropriate given that recent landings are largely the result of discards.

Significant information was presented by the CPWG regarding the length and age composition of butterfish landings along with observer length composition of butterfish kept and discarded. Commercial catch-at-age data showed the butterfish catches were comprised of primarily age-0 and age-1 fish, with some age-2 fish and virtually no age 3+ fish. Thus, good recruitment and the propagation of strong year classes appear vital to the success of the butterfish fishery.

ToR 2. *Characterize the survey data that are being used in the assessment. Describe the magnitude of uncertainty in these sources of data.*

This ToR was met. The Panel noted that the primary sources of fisheries-independent butterfish data were from bottom trawls which may not be ideal for measuring changes in abundance of a pelagic species known to undergo diel vertical migrations. However, no other types of survey data were available and the Panel accepted that the survey data presented represented the best available information.

The CPWG examined the utility of several trawl data sets, including state surveys (CT, RI, MA, ME-NH), the NEAMAP spring and fall surveys, and the NEFSC spring and fall surveys. The state surveys were not used in the assessment due to their respective limited spatial coverage when compared to the distribution of butterfish. Although the Panel accepted this decision, there was general discussion that the state surveys may represent an important area of further inquiry. Estuarine and near coastal areas can support relatively high localized abundances of butterfish and survey data from these area could be informative. Overcoming the limited spatial coverage challenge could potentially be accomplished by developing standardized indices for each survey and then examining ways to combine those indices through time-series methods (e.g., dynamics factor analysis). The Panel encouraged the CPWG to more fully consider the state survey data for the next butterfish assessment.

Six survey data sets were included in the butterfish assessment: NEFSC spring & fall offshore, NEFSC spring & fall inshore, and NEAMAP spring & fall. Length based calibration coefficients for the NEFSC spring and fall offshore survey data were estimated internally within the assessment model to adjust for vessel/gear changes associated with the transition from the FSV *Albatross IV* (1989-2008) to the FSV *Henry B. Bigelow* (2009-2012). The CVs of the NEFSC spring indices of abundance were generally higher (0.4-0.6) than those for the fall indices (0.2-0.4). The CVs of the NEAMAP abundance indices were fairly low (generally < 0.25) for both spring and fall. The base assessment model configuration developed by the CPWG included all six surveys, however, the Panel noted that the respective trends in abundance of the spring indices differed from the fall indices, particularly in the most recent years.

Since 2008, the number of fish measured for length and processed for age determination from the NEFSC surveys has increased, which improved confidence that size and age composition of those surveys were well characterized. The NEFSC spring offshore survey data was almost exclusively comprised of age-1 fish while the NEFSC fall offshore survey data was largely age-0 fish. Therefore, it appears that the fall data provided a reasonable measure of butterfish recruitment. However, the Panel noted that the NEFSC age-0 fall indices and the age-1 spring indices were inversely correlated, which raised questions about overwintering mortality and the ability of the offshore survey to track cohorts. In general, correlation analyses indicated that the NEFSC spring surveys more strongly tracked cohorts through the age-structure than the complementary fall surveys, although the numbers-at-age underpinning these correlations were low. The Panel concluded that the spring surveys seemed to provide more robust measures of butterfish abundance than the fall surveys (based on age propagation), but it was acknowledged that butterfish are more widely distributed within the survey area during fall. Reconciliation of survey incongruences was explored more fully through alternative configurations of the base assessment model (see ToR 5 summary for details).

ToR 3. *Characterize oceanographic and habitat data as it pertains to butterfish distribution and availability. If possible, integrate the results into the stock assessment (TOR-5).*

This ToR was met. The Panel strongly agreed that the analytical efforts to understand butterfish habitat preferences, particularly in the context of how those preferences may affect survey catchability (q), were rigorous and highly innovative. The breadth, depth, and complexity of the modeling underlying the butterfish habitat preference analysis was substantial, however, the Panel generally felt that there was insufficient time allotted to fully understand many of the key analytical details. The supporting text in the assessment report could have been clearer.

The primary result of the habitat preference modeling was an index of butterfish availability. For the assessment model, catchability for the NEFSC fall offshore survey was taken to be the product of availability (A) and efficiency (E): q = A*E. Therefore, the annual estimated habitat index values were used to adjust the annual availability of butterfish to the survey, which resulted in a time-series of q values. However, a key assumption to the habitat preference analysis was that butterfish distribution and localized abundance were structured by temperature. As a base hypothesis, this assumption seemed reasonable. However, the Panel noted that the GAM model used to quantitatively describe the butterfish-temperature relationship explained only 30% of the deviance in the observed data. Therefore, a potential disconnect was raised; if temperature was overridingly important, why did the GAM not explain more variance? Additionally, the noted failure of the GAM to converge when large survey catches were included in the analysis was concerning in terms of model robustness. Although these issues could not be fully reconciled during the meeting, the Panel strongly recommended continued refinement of the butterfish habitat model, consideration of other potential covariates, and exploration of its use for assessments of other species.

ToR 4. *Evaluate consumptive removals of butterfish by its predators. If possible, integrate results into the stock assessment (TOR-5).*

This ToR was partially met. The NEFSC is fortunate to have such a large-scale database of stomach contents information to evaluate predation impacts. The CPWG presented time-series of removals by the primary butterfish predators: smooth dogfish (*Mustelus canis*), spiny dogfish (*Squalus acanthias*), silver hake (*Merluccius bilinearis***)**, summer flounder (*Paralichthys dentatus*), bluefish (*Pomatomus saltatrix*), and goosefish (*Lophius americanus*). Trends in the abundances of those fishes were also presented along with correlations among butterfish and predator abundances. Estimated annual consumptive removals by these predators ranged from 20 to > 25,000 mt during the time period of the assessment model, with a general range of 1000 to 8000 mt per year. Clearly these estimates represent a minimum since many other predatory fishes, cetaceans, and seabirds were not included in the analysis. The Panel acknowledged that the predation information provided was important and foundational to considering butterfish stock dynamics in an ecosystem context, however, there was consensus that much more could be done. The Panel noted that the CPWG could have taken advantage of the rich ecosystem and multispecies modeling efforts that have been ongoing in the NW Atlantic over the past decade (e.g., Atlantis, EwE, MSVPA, MS Statistical Catch-At-Age) to gain deeper insights into butterfish stock dynamics. This observation is particularly germane given recently documented environmental changes in the NW Atlantic ecosystem, and evidence that the NEFSC fall offshore survey trend for butterfish has generally declined over the past 20 years despite very low to nonexistent direct fishery removals. The Panel recommended stronger integration among the Ecosystem Assessment and Population Dynamics

groups within the NEFSC, particularly for assessments of species known or hypothesized to be important food web components or highly affected by changing environmental conditions.

ToR 5. *Use assessment models to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a comparison with previous assessment results and previous projections.*

This ToR was met. The Panel supported the general philosophy of providing an input value for NEFSC fall offshore survey catchability (q) and allowing the model to estimate natural mortality (M). Although this approach is somewhat unusual, there was consensus that the CPWG was thoughtful, diligent, and comprehensive in how it generated the estimated q. The Panel also agreed that the approach taken was conservative since the estimated q hinged on the assumption that the survey cruises conducted from the *FSV Henry B. Bigelow* were 100% efficient at sampling butterfish. The externally defined q also includes the average measure of availability estimated under ToR 3.

The Panel acknowledged that the ASAP model within the NOAA Fisheries Toolbox has been peer-reviewed as a general assessment approach, and thus there was no concern about its use for butterfish. The general characteristics of the modified ASAP model (ASAP3) base configuration put forth by the CPWG were as follows:

- As mentioned above, the NEFSC fall offshore survey q was estimated externally to the assessment model and taken to be known; consequently, M was estimated by the ASAP3 model.
- The habitat index derived from the habitat preference modeling was included and used to modify the annual availability of butterfish to the NEFSC fall offshore survey (see ToR 3 summary for details).
- The ASAP3 model was fitted to the butterfish catch time-series and all six previously described survey indices (see ToR 2 summary for details).

After review of the base ASAP3 model results, the Panel concluded that it could not support the CPWG baseline model configuration for use in guiding butterfish management. Consensus points raised by the Panel were:

- The base model fit the catch time-series well; there were more positive residuals than negative residuals, but the magnitude of those residuals was quite small.
- The base model did not fit any of the indices particularly well, and there were strong contrasting patterns in the residuals associated with the fit to the NEFSC spring and fall offshore indices. For the spring survey, there were consistently large negative residuals early in the time-series followed by consistently large positive residuals in the more recent years. The residuals plot for the base model fit to the fall survey showed a reversal of this pattern, and the Panel inferred that these patterns in residuals were likely the results of contrasting trends in the fall and spring NEFSC offshore indices.
- The habitat index was generally flat through the time-series, and as expected, the base assessment model results were insensitive to its inclusion. Therefore, the principle of parsimony of the entire parameter space (within and 'outside' of the assessment model) would potentially favor excluding the habitat index.

The latter two points above were the Panel's basis for not accepting the base ASAP3 model configuration, and as a result, the Panel asked the CPWG to explore several alternative configurations:

- Alt. 1: Inclusion of spring only survey indices and the habitat index. This configuration was motivated from the perspective that these survey data more strongly tracked cohorts through the age-structure.
- Alt. 2: Inclusion of fall only survey indices and the habitat index. This configuration was motivated from the perspective that butterfish are more widely distributed across the survey area during fall and thus those survey trends more accurately represented patterns in overall abundance.
- Alt. 3: Inclusion of fall only survey indices without a temporally varying habitat index.

The Panel did not accept Alternative 1 because fitting that configuration required using the original base model estimated M as an input (somewhat circular logic). Alternative base models 2 and 3 were viewed as viable, and the Panel ultimately accepted Alternative 3 since, as expected, the model results were insensitive to inclusion of the temporally varying habitat index and the principle of parsimony would therefore favor not including it.

ToR 6. *State the existing stock status definitions for "overfished" and "overfishing". Given that the stock status is currently unknown, update or redefine biological reference points (BRPs; point estimates for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY, or their proxies) and provide estimates of their uncertainty. Consider effects of environmental factors on stability of reference points and implications for stock status.*

This ToR was met. Since the previous butterfish assessment model was not accepted (SARC 49), formal peer-reviewed BRPs were not available. Therefore, the CPWG proposed new BRPs which were: F_{MSY} proxy =2M/3 (Patterson 1992) and B_{MSY} proxy being the median 50yr spawning stock biomass (SSB) projection at the F_{MSY} proxy. The overfishing definition was $F_{Terminal} > F_{MSY}$ proxy, and the overfished definition was $B_{Terminal} < ½ B_{MSY}$ proxy. The Panel accepted these proxies and stock status definitions.

The Panel noted that the CPWG did not devote a great deal of effort to evaluating modeling sensitivities on BRP calculations and stock status determinations. That is, for most of the sensitivity runs examined (original base ASAP3 model configuration or accepted Alternative 3 configuration), the sensitivity runs were not carried forward to evaluate stock status with respect to BRPs. During the meeting, the Panel requested that the CPWG do this, particularly with respect to the configuration without the habitat index (thus evaluating effects of environmental variability). From the perspective of fully evaluating scientific uncertainty and model robustness, the Panel recommended carrying sensitivity model runs through to stock status determinations in future assessments.

ToR 7. *Evaluate stock status with respect to a newly proposed model and with respect to "new" BRPs and their estimates (from TOR-6). Evaluate whether the stock is rebuilt.*

This ToR was met. The Panel agreed that recent fishing mortality is quite low and that estimated stock biomass appears to be healthy (note, appropriate scaling of population size within the

assessment model hinges on accurate estimation of q). The current fishing mortality rate (F_{2012} = 0.02) was well below the overfishing reference point (2M/3 = 0.81 [CV = 0.05]). The current spawning stock biomass (SSB_{2012}=79,451 mt) was well above the accepted biomass reference point (B_{MSY} proxy =45,616 mt [CV = 0.25]). The Panel therefore accepted the conclusions that overfishing was not occurring, the stock was not overfished, and that the stock was rebuilt.

ToR 8. *Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).*

a) *Provide numerical annual projections (2 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment). Comment on which projections seem most realistic.*

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

This ToR was met. The Panel accepted the projection analysis based on the Alternative 3 base ASAP3 model configuration. Stock projections were made using a standard forward projection methodology, sampling recruitment from the entire time-series of model estimated values. Additionally, the Panel accepted the decision to structure the projections with the known 2013 butterfish catch, the 2014 butterfish Acceptable Biological Catch (ABC) and the F_{MSY} proxy thereafter. Collectively, results of the projection analysis suggested that the probability of overfishing in 2014 was < 1% and the probability of becoming overfished by 2017 was < 1%.

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

This ToR was met. The CPWG reported on progress made relative to research recommendations made in last assessment (SARC 49) and proposed new recommendations. Impressive advancements were made with regard to understanding the environmental effects on butterfish (habitat preference modeling – ToR 3), new BRPs for butterfish that acknowledge its potentially key role as a small pelagic prey fish (ToR 6), and enhanced observer coverage in both the New England and mid-Atlantic regions to improve discard estimation. The CPWG also provided four additional research recommendations that were accepted by the Panel:

- Field experiments to examine efficiency and catchability of survey gear for the purpose of improving assessments. Particular emphasis should be given to catchability of the FSV *Henry B. Bigelow* trawl net configuration.
- Explore the possibility of notable butterfish spawning south of Cape Hatteras, NC and potential contributions to the northern population.

- Continued development of a modified ASAP model that incorporates environmental covariates, particularly the addition of survey qs.
- Do not conduct additional assessments until a fishery has developed that holds potential to influence total stock biomass.

The Panel recommended augmenting the last point contingent upon not observing signs of concern in the survey indices (notable downward trends) and fishery catches (notable increases in estimated discards). Additionally, the Panel recommended:

- Calculating q for the NEFSC spring offshore survey.
- Expanding ecosystem considerations in the assessment of butterfish by drawing from and/or concurrently using ecosystem and multispecies models.
- Examine statistical modeling approaches that can make use of state survey data for the butterfish assessment.

3. Review of Golden Tilefish

3.1 Background

The golden tilefish is a relatively long-lived, slow-growing demersal teleost that can be found along the outer continental shelf and upper continental slope from Nova Scotia to South America, including the Gulf of Mexico, but is relatively abundant in the New England and mid-Atlantic regions (Bigelow and Schroeder 1953). This species is often found in and around deep water canyons where it creates burrows in the bottom sediment. Golden tilefish generally prefer colder water temperatures ranging from 10-15 °C where they feed on a variety of crustaceans, bivalves, and gastropods. Age and growth studies have confirmed that golden tilefish can attain sizes > 1 m and have a lifespan > 40 years (Turner et al. 1983, Turner 1986). Reproductive maturity occurs at ages 5-6 and spawning of takes place from March through November (McBride et al. 2013). Golden tilefish along the U.S. east coast are managed by the Mid-Atlantic Fishery Management Council.

3.2 Evaluation of the Terms of Reference for Golden Tilefish

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

This ToR was met. Estimates of total catches (directed landings) were provided from 1915-2012, although the beginning years for the assessment models were 1971 (SCALE, ASAP models) and 1973 (ASPIC model). The rationale for using starting years in the 1970s was because the directed golden tilefish longline fishery developed during that same time period. The Panel accepted the DWG's choices of start years, and the congruence between model start years and inception of the fishery implied that the assessment began when the golden tilefish stock was likely near virgin conditions.

Landings appear to come primarily from two relatively small areas within the continental shelf along the U.S. eastern seaboard (statistical areas 616 and 537). There is a temporal pattern to how these two areas have been fished. During the early years, fishers were based largely out of ports in New Jersey which implied landings came primary from the southern area (616). Part way through the fishery, effort shifted from originating out of New Jersey to originating out of New York. Thus,

catches later in the time-series came largely from the northern area (537). Discards appeared to be negligible and recreational landings seemed low. The Panel concluded that the DWG thoughtfully evaluated the sources of landings data and ultimately produced a useful and informative catch series.

Landings by market category (size delineations) were available from 1990-2012 and length data of landings by size category were available from 1995-2012. Aging of commercial landings has only recently begun such that age composition information was available from 2007-2012. Clearly, these age data were crucial to the application of the ASAP assessment model (more details below in ToR 4 summary) and efforts to age commercial catches should continue. From the available age composition data, the longline fishery appeared to target age 4-8 fish despite the reported golden tilefish longevity of approximately 40 years.

The Panel agreed with concern raised by the DWG about the apparent inconsistency in the market category definitions across dealers. Given that the bulk of the present day fishery is comprised of a selected few fishers and dealers, the prospect of standardizing market category designations seems likely. This represents an important area of future research.

ToR 2. *Characterize commercial LPUE as a measure of relative abundance. Consider the utility of recreational data for this purpose. Characterize the uncertainty and any bias in these sources of data.*

This ToR was met. In general, the use of commercial landings-per-unit-effort (LPUE) as a measure of population abundance is not desirable, however, the Panel acknowledged that the lack of available fishery-independent data left the DWG with very options. The spatial limitation of the fishery relative to the distributional range of golden tilefish is a continued concern, however, the apparent consistency in the LPUE indices with strong year classes propagating through the recent age composition data was encouraging. Recreational catches seemed low and MRIP data were not likely to yield an informative index of abundance. Estimated CVs for the nominal LPUE index were not provided and the Panel recommended such measures be routinely reported in future assessments. The Panel also raised concern about the effort definition used in developing the LPUE indices; trip length was used because a more specific definition that reflected variables such as number of hooks, longline length, or hours fished could not be developed. Refined trip level data would be informative to future assessments.

The DWG provided a GLM standardized LPUE index that included year and permit (vessel) effects. The Panel viewed this as an important first step, but ultimately concluded that much more could be done. In light of apparent recent dynamics changes in the NW Atlantic Ocean, there is likely significant promise in exploring the effect of other covariates on LPUE, particularly the potential relationship of larger scale climate drivers on tilefish recruitment. CVs of the standardized index were good (< 0.3).

ToR 3. *For the depth zone occupied by tilefish, examine the relationship between bottom temperature, tilefish distribution and thermal tolerance.*

This ToR was partially met. The DWG provided a golden tilefish preferred temperature profile based on literature and occasionally survey collections. The Panel acknowledged that few

ecological studies of golden tilefish have been conducted, however, as mentioned above, the GLM analysis of LPUE could have been expanded significantly to address hypotheses regarding the role of large scale environmental and climate drivers on golden tilefish abundance and recruitment. Formally merging ecological analyses with stock assessment modeling efforts allows for developing more holistic conclusions about stock dynamics.

ToR 4. *Use assessment models to estimate annual fishing mortality and stock size for the time series, and estimate their uncertainty. Include a historical retrospective to allow a comparison with previous assessment results.*

This ToR was met. The Panel greatly appreciated the efforts of the DWG to apply three independent assessment models (ASPIC, SCALE, and ASAP). Each modeling approach has strength and weaknesses, and the simultaneous consideration of all three facilitates a more informed evaluation of population dynamics, fishery impacts, and stock status.

ASPIC model: This model was applied primarily for continuity purposes with the previous peer-reviewed assessment (SARC 48). The Panel agreed with the DWG that this model should not be used for golden tilefish. The primary concerns were the significant retrospective patterns and the inability of the ASPIC model to adequately capture strong year classes as they propagated through the population. The presence of periodic strong golden tilefish year classes was confirmed via modal analyses of the commercial length and age composition data.

SCALE model: This model was also applied for the previous peer-reviewed assessment (SARC 48), and updated in support of the current benchmark review. The Panel shared the concerns raised by the DWG about the use of this model for golden tilefish. Specifically, the model did not fit the LPUE indices or the commercial length frequency data well in some years, the lack of length frequency information from the 1980s-1990s was problematic, and there were strong retrospective patterns. Heuristically, the SCALE model results provided some evidence for dome-shaped selectivity since estimated numbers-at-age in the older age classes were unrealistically low.

ASAP model: This model has not been applied for any previous golden tilefish assessment. The base ASAP model configuration developed by the DWG was as follows:

- Time period: 1971-2012.
- Ages 1-10+.
- Historic commercial age composition was resolved using pooled age/length information from 2007-2012 samples.
- M=0.15.
- Dome-shaped selectivity curve with the peak fixed at age-7 for time block 1 (1971-1982) and age-5 for time block 2 (1983-2012); other age-specific selectivity values were estimated within the model.

The base model slightly over-predicted the total catches (small but consistently positive residuals), but fit the yearly age-composition and LPUE indices well. As a result, the Panel accepted the base ASAP assessment model as configured by the DWG. There was good discussion about the choice of a dome-shaped selectivity function in the base model. Imposing a dome-shaped selectivity function is not risk averse since doing so implies older fish will not be experiencing the full annual fishing mortality rate. If this assumption is incorrect, then the assessment model results may be overly

optimistic regarding stock status. However, there appeared to be reasonably strong evidence supporting the dome-shape. Most notable was the low frequency of age-8+ fish in the recent commercial age composition data. Additionally, industry representatives provided statements indicating that they deliberately target younger/smaller fish and avoid areas where there are higher abundances of older/larger fish because of increased operational difficulties and gear conflicts with other fisheries.

The results of the base model indicated that full fishing mortality systematically increased during the 1970s to relatively high levels in the 1980s and early 1990s. Over the past decade, estimated fishing mortality has been variable but generally decreasing. Estimated SSB has decreased considerably since the 1970s, which is expected given that the golden tilefish longline fishery began when the stock was near virgin conditions. Since 2000, estimated SSB has increased slightly. Model predicted recruitment showed a fluctuating pattern which reflects the periodic emergence of strong year classes. Several of the CVs associated with the estimated recruitments were high (> 0.5) early in the time-series, but much improved over the last decade (< 0.3). All of the CVs for the estimated SSB and fishing mortality values were low (< 0.35). Although there was evidence of retrospective patterns in the ASAP base model, it was not possible to directly compared these results with those of previous assessments since the ASAP approach had not been applied in earlier assessments.

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

This ToR was met. The DWG took advantage of the fact that the golden tilefish stock has been managed under a constant quota since 2002 to provide a basis for defining BRPs. From the base ASAP model, the average fishing mortality was 0.37 which corresponds to $F_{25\%}$ in a yield-per-recruit analysis. Therefore, F=0.37 was assumed as an F_{MSY} proxy. Long term projections with fishing mortality held constant at the F_{MSY} proxy=0.37 value and recruitment sampled from the estimated values across the entire time-series were used to establish an SSB_{MSY} proxy of 5153 mt.

The Panel accepted these BRP definitions. Since the golden tilefish stock has produced strong year classes under the 2002-present constant quota management regime, and the base ASAP model results indicated no signs of recent stock decline or collapse, these BRPs seemed reasonable and possibly conservative. Estimates of uncertainty for the BRPs were not provided.

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

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

b) *Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-4).*

This ToR was met. The Panel accepted the base ASAP model configuration, subsequent BRPs, and conclusions that the overfishing was not occurring (F_{2012}=0.275, F_{MSY} proxy=0.37, F_{2012}/F_{MSY} proxy=0.74) and the stock was rebuilt (SSB_{2012}=5229, SSB_{MSY} proxy=5153, SSB_{2012}/SSB_{MSY} proxy=1.01). The average recruitment from 1971 to 2012 was estimated to be 1.24 million fish at age 1. Recent large year classes have occurred in 1998 (2.35 million), 1999 (2.39 million) and 2005 (1.85 million). The 2011 year class was estimated to be about 0.75 million age-1 fish.

The ASPIC and ASAP models are structurally very different and the ASPIC-based BRPs are estimates of F_{MSY} and B_{MSY}, so direct comparisons were not possible. The Panel agreed that the ASPIC model fit the observed data poorly and it cannot properly account for high variability in recruitment, which has been observed for golden tilefish. Consequently, stock status determination for golden tilefish should not be based on ASPIC assessment results.

ToR 7. *Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).*

a) *Provide numerical annual projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).*

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

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

This ToR was met. Stock projections were made using a standard forward projection methodology, sampling recruitment from the entire time-series of model estimated values. During the meeting, there was discussion that the estimated recruitment values in 2010-2012 were likely too low. The evidence supporting this conclusion came from results of the retrospective analysis. Commercial age composition data indicated that fish begin recruiting to the fishery at about age-4, and as a result, the assessment model has virtually no information to guide recruitment estimation for cohorts that have not yet entered the fishery. When the base model projection analysis was conducted, estimated catches and stock abundances were considerably lower than what had been previously observed and estimated by the base ASAP model. The stock also became significantly overfished in the first few years of the projections, which was inconsistent with the trends in stock abundance resulting from the assessment model. The Panel therefore agreed that the estimated recruitment values in the last three years of the assessment should be adjusted upward to the mean of the model estimated recruitments across the assessment time period. Operationally, this adjustment affected the 2013 population estimates for ages 2-4.

Additionally, the Panel accepted the decision to structure the projections with the known 2013 tilefish catch and the 2014 tilefish ABC. For the post 2014 years, two landings scenarios were considered. The first allowed landings to vary through time and ultimately increase to 1028 mt by 2017, while the second held landings constant through time at 905 mt. In both cases, the probability of becoming overfished in any year up to 2017 was less than 3%. Under the constant landings projection, the probability of overfishing occurring in any year up to 2017 was less than 50%.

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

This ToR was met. The DWG provided two research recommendations:

- Development of an industry based survey that would follow a design similar to a fishery-independent survey.
- Increased maturity sampling.

The first research recommendation was made in last assessment (SARC 48), and thus very little progress has been made to date. Nevertheless, this is an important area of research and efforts should be directed at initiating and industry-based golden tilefish survey. The second recommendation seems to have been addressed by the recent golden tilefish maturity study where 688 specimens were collected by commercial fishers in June and July of 2008 (McBride et al. 2013). Although the temporal coverage of this study is limited, sample size is adequate and maturity was determined via histology. This study represents an important contribution to the reproductive biology of golden tilefish, so it is a bit perplexing why the DWG highlighted increased maturity sampling as a research need. Additional research recommendations articulated in the last assessment and reiterated by the Panel were:

- Refined longline trip level information.
- Standardization of dealer designations of market categories.

In addition to those listed above, the Panel recommended:

- Analyses to investigate potential large scale environmental and climate drivers on golden tilefish abundance and recruitment.
- Continued aging of the commercial landings with a focus toward age validation.
- Routine reporting of CVs for estimated quantities (e.g., nominal LPUE indices, point estimates of BRPs, etc.).

The most notable research advancement was the 2007-2012 aging information of the commercial golden tilefish catches. Clearly, these aging data facilitated the successful application of the ASAP assessment model, and as noted above, should continue into the future.

4. Review of Northern Shrimp

4.1 Background

The northern shrimp (*Pandalus borealis*) is a crustacean that is found eastern Canadian and U.S. waters. Within the U.S., northern shrimp primarily inhabit waters along the Maine, New Hampshire, and Massachusetts coastlines, and as such, that the Gulf of Maine (GOM) ecosystem is considered to be the southern limit of its range (Shumway et al. 1985). Northern shrimp are protandrous hermaphrodites where they first mature as males at about age-2 and then transform to females by about age-3 (Shumway et al. 1985, Clark et al. 2000). Maximum age is 5 years and maximum size is 7-10 cm. Spawning occurs in offshore waters during late summer. During fall, females brood embryos externally and migrate inshore to release larvae during winter (Richards et al. 2012). Larvae experience six pelagic developmental stages before settling as benthic juveniles in late spring or early summer (Richards et al. 2012). Juveniles remain inshore for approximately a year and then migrate offshore where they join the mature population (Shumway et al. 1985). Temperature has been shown to be an important structuring variable for northern shrimp year-class success (Richards et al. 2012). The GOM fishery for northern shrimp is managed by the Atlantic States Marine Fisheries Commission's (ASMFC) interstate fishery management plan.

4.2 Evaluation of the Terms of Reference for Northern Shrimp

ToR 1. *Present the Gulf of Maine northern shrimp landings, discards, effort, and fishery-independent data used in the assessment. Characterize the precision and accuracy of the data and justify inclusion or elimination of data sources.*

This ToR was met. Estimates of commercial catches (directed landings) were provided for 1960-2013, although the 2013 landings were considered preliminary. Annual catches over the time series were highly variable with the early 1970s, late 1990s, and early 2010s being periods of large catches and the late 1970s, early 2000s, and 2012 being periods of very low catches. The bulk of the landings are in the state of Maine, followed by New Hampshire and Massachusetts, respectively. From 2010-2012, commercial landings have exceeded recommended levels by the NSTC. The fishery for northern shrimp takes place from December-April, with the majority of fishing activity occurring in January-February. There are two primary gear types used to harvest northern shrimp, trawls and traps, with the former being responsible for over 85% of the landings on average (2000-2013). The spatial dynamics of the fishery resulted in landings being largely comprised of ovigerous and spent females (ages 4-5); very few small shrimp (males, juveniles, stage-1 females) were landed. Shrimp discards were deemed negligible.

Data from port interviews indicated that annual fishing effort is highly variable. From 2000-2013, the number of vessels in the fishery averaged 230, with a sharp increase to > 300 from 2011-2012, followed by 198 in 2013. Total trips peaked in the mid-1990s at over 10,000 annually, however, since that time the number of trips has ranged from 2-6,000 (not including 2013). Measures of commercial catch-per-unit-effort (CPUE; lbs/hr) were poorly correlated with total landings. The NSTC viewed commercial CPUE as an unreliable indicator of northern shrimp abundance and consequently excluded it from the assessment. The Panel concluded that northern shrimp landings and effort were well characterized, however, the Panel recommended revisiting the utility of the commercial effort time-series (see ToR 2 summary for more details).

In terms of fishery-independent data sources, the NSTC considered five surveys: the state-federal (ASMFC) summer shrimp survey, the NEFSC spring and fall trawl surveys, the Maine-New

Hampshire Inshore survey, and the Maine Summer Shrimp survey. Of those considered, only the ASMFC summer survey and the NEFSC fall survey were included in the assessment (note that the FSV *Albatross IV* [1968-2008] and FSV *Henry B. Bigelow* series [2009-2012] were treated separately due to the absence of northern shrimp calibration coefficients). The rationale for excluding the other surveys largely centered on their relatively limited spatial coverage (NEFSC spring survey not withstanding), and concern about confounding effects of interannual variation in the timing of post-hatch female northern shrimp. The Panel supported this decision, but also encouraged the NSTC to more fully explore the utility of the other surveys perhaps through application of model based approaches that can standardize for environmental effects on northern shrimp movement.

Confidence intervals for the ASMFC summer and NEFSC fall FSV *Albatross IV* indices were fairly narrow, with the exception of 2006 which was a very unusual year in that both surveys yielded incredibly high index values when compared to all other years. The FSV *Henry B. Bigelow* index was much less precise. All survey size composition data appeared to be informative and useful for tracking stage/age classes.

ToR 2. *Estimate population parameters (fishing mortality, biomass, and abundance) using assessment models. Evaluate model performance and stability through sensitivity analyses and retrospective analysis, including alternative natural mortality (M) scenarios. Include consideration of environmental effects where possible. Discuss the effects of data strengths and weaknesses on model results and performance.*

This ToR was not met. The Panel greatly appreciated the efforts of the NSTC to apply three independent assessment models (UME, CSA, and ASAP). Each modeling approach has strength and weaknesses, and the simultaneous consideration of all three facilitates a more informed evaluation of population dynamics, fishery impacts, and stock status.

UME model: Although this model has been used for the assessment of other species, the Panel raised concerns about whether or not the general model formulation and implementation had been peer-reviewed and/or subjected to adequate simulation testing. The absence of such testing renders it difficult to determine if patterns in model diagnostics are attributable to data in an assessment or are general emergent model properties. The Panel strongly encouraged the NSTC to conduct a general simulation analysis of the UME model and to have the model peer-reviewed prior to its next application.

The UME model is a statistical stage- and size-structured approach which conceptually is a good match for northern shrimp. However, the Panel raised the following concerns about the base model configuration and performance:

- The 'U-shaped' pattern for M-at-length seemed difficult to justify, especially the assumption of a terminal molt.
- The very poor fit to the survey indices, even when the fit to the 2006 data point was ignored.
- The very poor fit to the survey length composition data, which was especially problematic given that the strength of the UME model is its size-based structure.
- Modest retrospective patterns.
- Relative insensitivity of model output to M (constant M vs. highly variable M with size).

The poor fit to virtually all fishery-independent data led the Panel to reject this model for fishery management. To that end, the Panel did encourage additional refinement of the UME model as additional iterations of it may prove useful in future assessments.

CSA model: This model was applied and accepted during the last peer-reviewed assessment of northern shrimp (SARC 45). As a result, the NSTC indicated that this model was the preferred approach. Conceptually, a two-stage model also seems like a good match for northern shrimp. The Panel largely agreed with the choices underlying the base CSA model configuration but expressed concerns about the notable patterns in the residuals of fits to the survey indices (cyclic trends of negative and positive values through time). The Panel acknowledged that the 2006 survey data points were causing difficulties, but also surmised that the imposed CVs on the catches and survey indices along with the relative weighting scheme of the likelihood components could also be contributing to the patterns. In the base model, the various likelihood components were weighted equally and the imposed CV on the catch series was very low compared to the CV of the indices. In effect, a low catch CV forced the model to closely fit the observed landings series, and given the highly variable fishery dynamics (number of vessels in the fishery annually) and management policies (annual total allowable catches), it is reasonable to speculate that catches did not adequately represent true abundance. As a result, the Panel asked the NSTC to explore a variety of adjusted CVs and likelihood weighting schemes. The results of these alternative model configurations showed marginal improvements to the aforementioned residuals patterns, but more alarmingly, they revealed that vastly different results could be achieved depending on the likelihood weights (overfishing versus not overfishing). In effect, the chosen likelihood weights predetermined the resulting stock status determinations. The Panel therefore concluded that the CSA northern shrimp model was not robust and could not be used for fishery management.

The Panel acknowledged that its conclusions about the CSA model were inconsistent with the previously accepted northern shrimp peer-reviewed assessment. However, it is important to recognize that assessment model performance can break down with the addition of new data. The Panel felt strongly that NSTC's efforts were rigorous and that the analysts did everything possible to develop a reliable northern shrimp assessment. Unfortunately, the post 2006 data created unforeseen technical problems within the CSA and UME modeling frameworks.

ASPIC model: This model was applied primarily for heuristic and confirmatory purposes. However, recent northern shrimp recruitment has been highly variable, and the inability of ASPIC to adequately capture strong year classes as they propagate through the population led the Panel to recommend discontinued use of this assessment approach for the foreseeable future.

ToR 3. *Update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, SSB_{MSY}, F_{MSY}, MSY). Evaluate stock status based on BRPs.*

This ToR was not met. Given that the Panel did not accept any of the proposed assessment models, stock status based on BRPs could not be determined. The Panel also did not accept the proposed BRPs since these were based on new assessment models that were rejected. The Panel recommended that management advice be based on inferences drawn from patterns in the observed catches, commercial CPUE, and survey time-series in lieu of an accepted assessment model. It also noted that the recent survey indices indicated that the population abundance appears to be very low.

ToR 4. *Characterize uncertainty of model estimates of fishing mortality, biomass and recruitment, and biological reference points.*

This ToR was not met. Given that the Panel did not accept any of the proposed assessment models or BRPs, evaluation of uncertainty of model estimates and BRPs was not possible.

ToR 5. *Review the methods used to calculate the annual target catch and characterize uncertainty of target catch estimates.*

This ToR was not met. Given that the Panel did not accept any of the proposed assessment models, evaluation of the methods used to calculate annual target catch levels and associated uncertainty was not possible.

ToR 6. *Develop detailed short and long-term prioritized lists of recommendations for future research, data collection, and assessment methodology. Highlight improvements to be made before the next benchmark assessment.*

This ToR was met. The Panel agreed that the NSTC assembled a very thoughtful and comprehensive list of research priorities for northern shrimp. The following broad focal research areas were identified: fishery-dependent, fishery-independent, model/quantitative, life history/biological/habitat, and management/enforcement/socioeconomic. Within each of these groupings, numerous detailed research priorities were specified and organized into short-, moderate-, and long-term categories, with several highlighted for completion prior to the next benchmark assessment.

In addition to supporting those already specified by the NSTC, the Panel recommended three additional research priorities:

- Continued refinement of UME model.
- Evaluation of the adequacy of the current BRPs, perhaps through management strategy simulations.
- Potential inclusion of the effort time-series and/or a commercial CPUE time-series standardized for environmental effects in the CSA model.

ToR 7. *Based on the biology of species, and potential scientific advances, comment on the appropriate timing of the next benchmark assessment and intermediate updates.*

This ToR was met. As noted above (ToR 3 summary), in lieu of an accepted assessment model, the Panel recommended basing management advice on observed patterns in the northern shrimp catches, survey indices, and possibly commercial CPUE. In effect, the NSTC could conduct straightforward annual index based assessments while refining the UME and CSA models. Another benchmark assessment should be considered as soon as improvements can be made to either the UME model (better fits to survey, size composition data) and/or the CSA model (robustness to likelihood weights).

5. Literature Cited

Bigelow, H. B. and W.C. Schroeder. 1953. Fishes of the Gulf of Maine. U.S. Fish Wildlife Service Fisheries Bulletin 53.

Clark S.H., S.X. Cadrin, D.F. Schick, P.J. Diodati, M.P. Armstrong, and D. McCarron. 2000. The Gulf of Maine northern shrimp *(Pandalus borealis)* fishery: a review of the record. Journal of Northwest Atlantic Fisheries Science 27: 193–226.

Cross, J.N., C.A. Zetlin, P.L. Berrien, D.L. Johnson, and C. McBride. 1999. Essential fish habitat source document: butterfish, *Peprilus triacanthus*, life history and habitat characteristics. NOAA Technical Memorandum NMFS-NE-145.

Duffy, D.C. 1988. Predator-Prey Interactions between Common Terns and Butterfish. *Ornis Scandinavica* 19:160-163.

McBride, R.S., T.E. Vidal, and S.X. Cadrin. 2013. Changes in size and age at maturity of the northern stock of Tilefish (*Lopholatilus chamaeleonticeps*) after a period of overfishing. Fishery Bulletin 111:161-174.

Patterson, K. 1992. Fisheries for small pelagic species: an empirical approach to management targets. Reviews in Fish Biology and Fisheries 2:321–338.

Richards, R.A., M.J. Fogarty, D.G. Mountain, and M.H. Taylor. 2012. Climate change and northern shrimp recruitment variability in the Gulf of Maine. Marine Ecological Progress Series 464:167-178.

Shumway S.E., H.C. Perkins, D.F. Schick, and A.P. Stickney. 1985. Synopsis of biological data on the pink shrimp, *Pandalus borealis* Krøyer, 1838. NOAA Technical Report NMFS 30, Seattle, WA.

Turner, S.C., C.B. Grimes, and K.W. Able. 1983. Growth, mortality, and age/size structure of the fisheries for tilefish, *Lopholatilus chamaeleonticeps*, in the Middle Atlantic-Southern New England region. Fishery Bulletin 81:751-763.

Turner, S.C. 1986. Population dynamics of and impact of fishing on tilefish, *Lopholatilus chamaeleonticeps*, in the Middle Atlantic-Southern New England region during the 1970s and early 1980s. Ph.D. dissertation, Rutgers University, New Brunswick, N.J.

6. Description of SAW Supporting Materials

Working paper	Title	Author(s)/Publisher
A1	Butterfish model diagnostics for the ASAP3 base model	Coastal/Pelagic Working Group
A2	49th SAW Assessment Summary Report	NEFSC
A3	49th SAW Butterfish Assessment Report	Coastal/Pelagic Working Group
A4	SARC 49 Reviewers Summary Report	Robert J. Latour
A5	SARC 49 Individual CIE Review	John Cotter
A6	SARC 49 Individual CIE Review	Henrik Sparholt
A7	SARC 49 Individual CIE Review	Michael Smith
A8	58th SAW Butterfish Assessment Summary Report (Draft)	Coastal/Pelagic Working Group
A9	58th SAW Butterfish Assessment Report (Draft)	Coastal/Pelagic Working Group
B1	The 1882 tilefish kill – a cold event in shelf waters off the north-eastern United States	Robert Marsh & colleagues
B2	Changes in size and age at maturity of the northern stock of Tilefish *(Lopholatilus chamaeleonticeps*) after a period of overfishing	Richard McBride & colleagues
B3	48th SAW Assessment Summary Report	NEFSC
B4	48th SAW Tilefish Assessment Report	Demersal Working Group
B5	58th SAW Tilefish Assessment Summary Report (Draft)	Demersal Working Group
B6	58th SAW Tilefish Assessment Report (Draft)	Demersal Working Group
C1	A hierarchical approach to determining reference points for Pandalid shrimp	Steve Cadrin & colleagues
C2	Developing a growth transition matrix for the Maine sea urchin stock	Yong Chen & colleagues
C3	Developing and evaluating a size-structured stock assessment model for the American lobster, *Homarus americanus*, fishery	Yong Chen & colleagues
C4	The Gulf of Maine northern shrimp (*Pandalus borealis*) fishery: a review of the record	Steve Clark & colleagues
C5	Estimates of predator consumption of the northern shrimp *Pandalus borealis* with implications for estimates of population biomass in the Gulf of Maine	Jason Link & Josef Idoine
C6	Climate change and northern shrimp recruitment variability in the Gulf of Maine	Anne Richards & colleagues
C7	45th SAW Northern Shrimp Assessment Report	ASMFC Northern Shrimp Technical Committee

C8	58th SAW Northern Shrimp Assessment Summary Report (Draft)	ASMFC Northern Shrimp Technical Committee
C9	58th SAW Northern Shrimp Assessment Report (Draft)	ASMFC Northern Shrimp Technical Committee

7. Appendices

Task Order T42-04, final 25 October 2013

Statement of Work

58th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for butterfish, tilefish, and northern shrimp

Statement of Work (SOW) for CIE Panelists

(including a description of SARC Chairman's duties)

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

SCOPE
Project Description: The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC 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. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fishery management in the northeast region.
Brief description of the science to be peer reviewed, and its relevant importance:
The purpose of this meeting will be to provide an external peer review of benchmark stock assessments for butterfish, tilefish, and northern shrimp. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results form the scientific basis for fishery management in the northeast region.

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

Duties of reviewers are explained below in the "Requirements for CIE Reviewers", in the "Charge to the SARC Panel" and in the "Statement of Tasks". The stock assessment Terms of Reference (ToRs) are attached in Annex 2. The draft agenda of the panel review meeting is attached in Annex 3. The SARC Summary Report format is described in Annex 4.

Requirements for the reviewers: Three reviewers shall conduct an impartial and independent peer review of the striped bass and summer flounder stock assessments, and this review should be in accordance with this SoW and stock assessment ToRs herein. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include statistical catch-at-age, state-space and index methods. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points that includes an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 58 will address fishery stock assessments of butterfish, tilefish, and northern shrimp. For shrimp and butterfish, experience in the following is desirable: assessment of short-lived species, stocks where the environment and environmental change can impact recruitment and availability in research surveys. Specifically for tilefish: experience with assessments based on commercial catch per unit of effort.

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

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

PLACE OF PERFORMANCE AND TRAVEL
Each reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during dates of January 27-31, 2014.

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

If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and $MSY_)$, the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
Each reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting at the Woods Hole, Massachusetts scheduled during January 27-31, 2014.
3) Conduct an independent peer review in accordance with this SoW and the assessment ToRs (listed in **Annex 2**).
4) No later than February 14, 2014, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment ToR in **Annex 2**.

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

December 16, 2013	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
January 13, 2014	NMFS Project Contact will attempt to provide reviewers the pre-review documents
January 27-31, 2014	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
January 31, 2014	SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA
February 14, 2014	Reviewers submit draft independent peer review reports to the contractor's technical team for independent review
February 17, 2014	Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *
February 21, 2014	SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)
February 28, 2014	Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements

March 7, 2014	The COR distributes the final reports to the NMFS Project Contact and regional Center Director

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

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

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

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

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

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

Upon the acceptance of each independent peer review report by the COR, the reports will be distributed to the NMFS Project Contact and pertinent NMFS science director, at which time the reports will be made publicly available through the government's website.
The contractor shall send the final reports in PDF format to the COR, designated to be William Michaels, via email William.Michaels@noaa.gov

Support Personnel:

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

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

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)

22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com, Phone: 571-223-7717

Key Personnel:

Dr. James Weinberg, NEFSC SAW Chairman, NMFS Project Contact
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov, Phone: 508-495-2352, FAX: 508-495-2230

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

Annex 1: Format and Contents of Independent Peer Review Report

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

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

    a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.
    b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.
    c. Reviewers should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification.
    d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.
    e. The independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, regardless of whether or not others read the SARC Summary Report. The independent report shall be an independent peer review of each ToR, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:
   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of this Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: 58th SAW/SARC Stock Assessment Terms of Reference

(file vers.: 8/2/2013)

A. Butterfish

1. Characterize the commercial catch including landings, effort and discards by gear type. Describe the magnitude of uncertainty in these sources of data.

2. Characterize the survey data that are being used in the assessment. Describe the magnitude of uncertainty in these sources of data.

3. Characterize oceanographic and habitat data as it pertains to butterfish distribution and availability. If possible, integrate the results into the stock assessment (TOR-5).

4. Evaluate consumptive removals of butterfish by its predators. If possible, integrate results into the stock assessment (TOR-5).

5. Use assessment models to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a comparison with previous assessment results and previous projections.

6. State the existing stock status definitions for "overfished" and "overfishing". Given that the stock status is currently unknown, update or redefine biological reference points (BRPs; point estimates for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY, or their proxies) and provide estimates of their uncertainty. Consider effects of environmental factors on stability of reference points and implications for stock status.

7. Evaluate stock status with respect to a newly proposed model and with respect to "new" BRPs and their estimates (from TOR-6). Evaluate whether the stock is rebuilt.

8. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

    a. Provide numerical annual projections (2 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment). Comment on which projections seem most realistic.

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

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

B. Tilefish

1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the magnitude of uncertainty in these sources of data.

2. Characterize commercial LPUE as a measure of relative abundance. Consider the utility of recreational data for this purpose. Characterize the uncertainty and any bias in these sources of data.

3. For the depth zone occupied by tilefish, examine the relationship between bottom temperature, tilefish distribution and thermal tolerance.

4. Use assessment models to estimate annual fishing mortality and stock size for the time series, and estimate their uncertainty. Include a historical retrospective to allow a comparison with previous assessment results.

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

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

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

    b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-4).

7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

    a. Provide numerical annual projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

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

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

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

C. Northern shrimp

1. Present the Gulf of Maine northern shrimp landings, discards, effort, and fishery-independent data used in the assessment. Characterize the precision and accuracy of the data and justify inclusion or elimination of data sources.

2. Estimate population parameters (fishing mortality, biomass, and abundance) using assessment models. Evaluate model performance and stability through sensitivity analyses and retrospective analysis, including alternative natural mortality (M) scenarios. Include consideration of environmental effects where possible. Discuss the effects of data strengths and weaknesses on model results and performance.

3. Update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, SSB_{MSY}, F_{MSY}, MSY). Evaluate stock status based on BRPs.

4. Characterize uncertainty of model estimates of fishing mortality, biomass and recruitment, and biological reference points.

5. Review the methods used to calculate the annual target catch and characterize uncertainty of target catch estimates.

6. Develop detailed short and long-term prioritized lists of recommendations for future research, data collection, and assessment methodology. Highlight improvements to be made before the next benchmark assessment.

7. Based on the biology of species, and potential scientific advances, comment on the appropriate timing of the next benchmark assessment and intermediate updates.

Annex 2 (cont.):

Appendix to the SAW Assessment TORs:

Clarification of Terms

used in the SAW/SARC Terms of Reference

Appendix to the Assessment TORs:

Explanation of "Acceptable Biological Catch" (DOC Natl. Standard Guidelines, Fed. Reg., vol. 74, no. 11, 1/16/2009):

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

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

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

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

Explanation of "Vulnerability" (DOC Natl. Standard Guidelines, Fed. Reg., vol. 74, no. 11, 1/16/2009):

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

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

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

Annex 3: Draft Agenda

58th Northeast Regional Stock Assessment Workshop (SAW 58)

Stock Assessment Review Committee (SARC) Meeting

January 27-31, 2014

Stephen H. Clark Conference Room – Northeast Fisheries Science Center

Woods Hole, Massachusetts

DRAFT AGENDA* (version: 25 October 2014)

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR
Monday, Jan. 27			
10 – 10:30 AM			
Welcome	**James Weinberg**, SAW Chair		
Introductions	**Robert Latour**, SARC Chair		
10:30 – 12:30	Assessment Presentation (A. Butterfish)		
	Charles Adams, John Manderson, Timothy Miller		
12:30 – 1:30 PM	Lunch		
1:30 – 3:00	Assessment Presentation (A. Butterfish)		
	Charles Adams, John Manderson, Timothy Miller		
3:00 – 3:15	Break		
3:15 – 5:15	SARC Discussion w/ Presenters (A. Butterfish)		
	Robert Latour, SARC Chair		
5:15 – 5:45	Public Comments (A. Butterfish)		

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, Jan. 28

8:45 – 11 AM	Assessment Presentation (B. Tilefish)
	Phil Nitschke
11:00 - 11:15	Break
11:15 – 12:15	SARC Discussion w/presenters (B. Tilefish)
	Robert Latour, SARC Chair
12:15 – 12:30	Public Comments (B. Tilefish)
12:30 – 1:45 PM	Lunch
1:45 – 4:15	Assessment Presentation (C. Northern shrimp)
	Kelly Whitmore, Anne Richards, Katie Drew
4:15 - 4:30	Break
4:30 – 5:45	SARC Discussion w/presenters (C. Northern shrimp)
	Robert Latour, SARC Chair
5:45 – 6:00	Public Comments (C. Northern shrimp)
7:00	**(**Social Gathering**)**

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Wed. Jan. 29

9:00 – 11:15 AM	Revisit with presenters (A. Butterfish)
	Robert Latour, SARC Chair
11:15 – 11:30	Break
11:30 – 12:30	Revisit with presenters (B. Tilefish)
	Robert Latour, SARC Chair
12:30 – 1:30 PM	Lunch

1:30 -2:30 (cont) Revisit with presenters (B. Tilefish)

Robert Latour, SARC Chair

2:30 – 2:45 Break

2:45 – 5:15 Revisit with presenters (C. Northern shrimp)

Robert Latour, SARC Chair

TOPIC PRESENTER(S) SARC LEADER RAPPORTEUR

Thur. Jan. 30

8:30 – 11:30 Review/edit Assessment Summary Report (A. Butterfish)

Robert Latour, SARC Chair

11:30 – 12:30 PM Lunch

12:30 – 2:45 Review/edit Assessment Summary Report (B. Tilefish)

Robert Latour, SARC Chair

2:45 – 3:00 Break

3:00 - 5:30 Review/edit Assessment Summary Report (C. Northern shrimp)

Robert Latour, SARC Chair

TOPIC PRESENTER(S) SARC LEADER RAPPORTEUR

Friday, Jan. 31

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

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

The NMFS Project contact will provide the final agenda before the meeting.

Reviewers must attend the entire meeting.

Annex 4: Contents of SARC Summary Report

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

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

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

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

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