Center for Independent Experts (CIE) Independent Peer Review Report

on

STAR 3: Assessment of the U.S. Pacific West Coast Sablefish in 2019

Prepared by

Yong Chen, Ph.D.

Center for Independent Experts

August 2019

Contents

I. Executive Summary

The STAR 3 review, held in Seattle, WA in 2019 from July 8-12th, was aimed to evaluate the 2019 stock assessment for the U.S. Pacific West Coast Sablefish (*Anoplopoma fimbria*) stock. This includes an evaluation of stock assessment data quantity and quality, assessment model configuration and parameterization, assessment outputs, model projection and uncertainty, as well as recommendations for the use of the stock assessment in providing management advice and for the improvement of the stock assessment and assessment process.

Using fishery-dependent and fishery-independent data, an integrated statistical catch-at-age model implemented in Stock Synthesis (SS) (version 3.30.13 released 2019-03-09) was pre-configured and pre-parameterized by the Sablefish Stock Assessment Team (STAT) before STAR 3. After a week-long intensive discussion, numerous requests for models runs with different model configurations and parameterizations, and careful evaluations and analyses of all the modeling results, the STAR Panel and STAT came to an agreement for a base case model that can be used to provide management advice. STAT, the Northwest Fisheries Science Center (NWFSC) and the Pacific Fisheries Management Council (PFMC) provided all necessary logistics support, documentations, data, and other background information requested. The STAT analysts involved in the process were open to suggestions, provided additional information upon request, and conducted many additional model runs identified by the STAR Panel. The STAT analysts accommodated all requests made for different model runs as well as extra biological and fishery information. The whole process was open and constructive, and all materials were sent to me in a timely manner. As a CIE reviewer, I am charged with evaluating the U.S. Pacific West Coast Sablefish stock assessment with respect to a set of predefined Terms of Reference.

I would like to commend the STAT's efforts during the Sablefish stock assessment review. I was impressed by the breadth of expertise and experience of all participants in STAR 3, the amount of effort spent to compile information and input data for the model, the considerations of plausible scenarios in the development and identification of the base case model, the openness of discussion for considering alternative approaches and suggestions, and the constructive dialogues between the Review Panel, STAT and other participants during the review.

There are large uncertainties in the Sablefish stock assessment, although the relative stock trend over time tends to be more robust regarding the different model configurations compared to the absolute stock biomass estimates. The uncertainties resulted from different data weighting methods, inclusion/exclusion of certain data (e.g., discard length composition data, length and age composition data), different assumptions on survey and fishery selectivities (e.g., dome versus logistic curves), model parameterizations (e.g., sex-specific life history parameters like natural mortality and fixing some model data in parameters estimation), as well as possible spatio-temporal variations in life history parameters (e.g., growth). The large uncertainties and several unusual model behaviors observed during the review process resulted in extensive discussion and a large number of requests for different model runs which eventually led to the identification and finalization of the base case model towards the end of the review. This prevented the STAR Panel and STAT from conducting further sensitivity analyses to investigate the uncertainties for the "final" base case model. However, the large number of model runs

during the process of identifying the "final" base case model can be considered as sensitivity analyses, which greatly helps in understanding the behavior of the model, the possible interactions among different modeling components and among different data sources, and the sources of uncertainty in the Sablefish stock assessment.

Overall, although stock assessment modeling results varied among different model configurations and parameterizations, this assessment clearly suggests that the West Coast Sablefish is currently not undergoing overfishing and is currently not overfished. The stock was estimated to be above the depletion level that would lead to maximum sustainable yield.

Based on the materials presented, extensive discussions and additional model runs conducted during the review, I conclude that the "final" base case model and resultant stock assessment results represent the best available practice and information on the fishery status of the West Coast Sablefish stock assessment, although large uncertainties still exist in the assessment regarding the data and model. I believe the assessment is scientifically sound and adequately addresses management needs. However, there is a need for improved understanding of (1) spatio-temporal variations in life history parameters (e.g., growth and maturation), fishing fleet behavior (e.g., selectivity and discard/retention) and associated implications on stock assessment; (2) statistical interactions of different SS model components in parameter estimation; and (3) implications of various assumptions made on the model configurations and parameterizations (e.g., fixed values and choices of selectivities). These issues can be addressed in short- and medium-long-term research recommendations identified in Section IV-6.

II. Background

Sablefish (*Anoplopoma fimbria*, or blackcod) are a highly mobile, long-lived, and commercially valuable groundfish species. They are distributed throughout the North Pacific Ocean, from the Northeastern Pacific, between the north-central Bering Sea and Southern California, to the Northwestern Pacific Ocean, between Kamchatka and the northeastern coast of Japan (Hart 1973). The Sablefish stock assessment and management in the Northeastern Pacific Ocean are more so based on political/jurisdictional boundaries and are thus divided by region: the Alaska federal region, the Alaska state region, British Columbia, and the U.S. West Coast (Johnson et al. 2015, DFO 2016, Hanselman et al. 2016). Closed stock is assumed in the assessment for each of these stocks, although studies suggest that Northeastern Pacific Sablefish are not genetically distinct among these stocks (Jasonowicz et al. 2017). This lack of genetic structure within Sablefish populations in the Northeastern Pacific Ocean may suggest that a region-wide stock assessment may be needed, or at the least, potential movement/exchange between these four areas should be considered in the assessment of each stock (Fenske et al. 2018).

Sablefish experience age/size-dependent inshore-offshore distributions, with small and young fish distributed in shallow/inshore waters, and large and older fish more likely to be found in deep shelf-slope waters. Older Sablefish dominate the areas beyond the shelf-slope break. Previous studies suggest a large spatial variation in Von Bertalanffy growth parameters across the Northeastern Pacific Ocean, with increasing maximum body sizes and decreasing growth rates as latitude increases (Head et al. 2014, Gertseva et al. 2017). Geographic break points in Sablefish growth occur at 36N at the start of the southern California Bight and around 50N where the North Pacific Current bifurcates.

The West Coast Sablefish exhibit a protracted spawning season from December through March, with spawning peaking in February along the continental shelf-slope break in waters deeper than 300 m. This shortened spawning season may reduce Sablefish availability to the commercial fishery during winter. Eggs are buoyant, and juveniles are pelagic in offshore surface waters, settling to the benthos at depths of less than 250 m as age-0 recruits during late summer to fall.

Sablefish grow fast, reaching full size (i.e., asymptotic size) and maturity in their first decade of life. Females tend to be larger than males, but males are more prevalent among mature Sablefish, implying a higher natural mortality for females. However, the oldest recorded Sablefish, a 102 year-old found off Washington in 2006, was female.

Historical Sablefish landings since 1890 have been reconstructed from different sources, and the quality of the reconstructed landing data is considered reliable. Sablefish landings increased dramatically during the 1970s with the largest single-year removal of over 25,000 mt from U.S. West Coast waters. Around 14,000 mt of Sablefish were landed per year between 1976 and 1990, with a rapid increase in domestic pot and trawl landings. Annual landings have remained below 10,000 mt in recent decades with around 46% from hook-and-line, 21% from pot, and 33% from trawl. The decline in landings over time may be a result of weak recruitment, declined stock, reduced market and increasing management regulations. Prior to the catch share program, the trawl fishery was widely distributed across the continental shelf and fixed gear fisheries (i.e., pot and hook-and-line) were more patchily distributed. Since the catch share

program in 2011, the trawl fishery has shifted towards deeper waters with a large reduction in California, hook-and-line fishery has shifted north with little effort in waters south of 40^0N, and the pot fishery has expanded in waters south of 36^0 N.

The most recent benchmark stock assessment for the U.S. West Coast Sablefish was conducted in 2011, implemented using Stock Synthesis 3. The model was considered well developed and provided sound management advice, although the assessment suffered from problems with residual pattern and lack of fit. An updated stock assessment was completed in 2015 with the same model structure and new additional data (Haltuch et al. 2019).

The current (2019) stock assessment (STAR 3) is aimed to assess the status of the U.S. West Coast Sablefish distributed off the coast from southern California to the U.S.-Canadian border through 2018 (Haltuch et al. 2019). The possible movement/exchange among the four areas is not explicitly considered in this assessment. The STAR 3 review, held in Seattle, WA from July 8-12, 2019, is aimed to evaluate the 2019 stock assessment for the U.S. Pacific West Coast Sablefish. The Panel reviewed stock assessment data quantity and quality, assessment model configuration and parameterization, assessment outputs, model projection, and uncertainty associated with the assessment, and made recommendations for the use of the stock assessment in providing management advice and for the improvement of the stock assessment and assessment process.

Using fishery-dependent and fishery-independent data, an integrated statistical catch-at-age model implemented in Stock Synthesis (SS) (version 3.30.13 released 2019-03-09) was pre-configured and pre-parameterized by the STAT before the STAR 3 review. After almost a week-long discussion, numerous new model runs with different model configurations and parameterizations, and careful evaluations and analyses of all the modeling results, the STAR Panel and STAT come to an agreement for a base case model which can be used to provide management advice. The STAT, Northwest Fisheries Science Center (NWFSC) and Pacific Fisheries Management Council (PFMC) provided all the necessary logistics support, documentation, data, and background information for the review.

III. Description of the Individual Reviewer's Role in the Review Activities

My role as a CIE independent reviewer is to conduct an impartial and independent peer review of the STAR 3 West Coast Sablefish stock assessment regarding the guidelines in the pre-defined Terms of Reference.

I received the draft of the stock assessment report, relevant working papers and background materials two weeks prior to the STAR 3 review. I also received relevant files for the "initial" base case model proposed by the STAT prior to the STAR 3 review, including SS3 input data files, control files, report files and other relevant files.

I have read the draft West Coast Sablefish stock assessment report and all the working papers, informational papers and other relevant documents that were sent to me (see the list in Appendix I). I have also searched, collected and read references relevant to the topics covered in the reports and the Performance Work Statement (PWS) prior to my trip to the STAR 3 review.

The STAR 3 review was held from July 8 to 12, 2019 in the Northwest Fisheries Science Center (NWFSC) in Seattle, WA (see Appendix II for the schedule). The STAR Panel consists of

one PFMC SSC member (Dr. John Field, STAR 3 Review Panel Chair), a stock assessment scientist from the NOAA Fisheries AFSC (Dr. Jim Ianelli), and two CIE reviewers (Dr. Robin Cook and myself). In addition, the five days of review were also attended by STAR Panel Advisors, NWFSC fisheries stock assessment scientists, PFMC Council staff and SSC representatives, representatives from CA, OR, and WA state agencies, and other stakeholders (see the List of Participants in Appendix III).

Presentations were given during the STAR 3 review to provide the STAR Panel with stock assessment input data, information on model configuration and parameterization, stock assessment modeling outputs and results, sensitivity analysis scenario settings and results, results of additional model runs identified during the STAR 3 review, and model projections. Presentations also covered ecosystem considerations and modeling Sablefish growth in the Northwestern Pacific Ocean (see the list of presentations in Appendix I).

During the STAR 3 review, the STAR 3 Panel worked with the STAT to evaluate a series of model configurations and associated modeling results in an effort to develop a plausible base case scenario. The development of model configurations follows the following principle: changing one variable at a time so that we can ensure that changes observed in modeling can be solely attributed to the change we made. The daily requests for additional analyses and model runs can be found in Appendix III. These additional model runs and analyses were completed in order to improve our understanding of model fitting, impacts of uncertainty in data and models, robustness of the assessment results, and the roles of different life history and fishery processes in modeling, all of which helped lead to the identification and finalization of the base case scenario and possible sensitivity analysis scenarios for providing management advice.

I was actively involved in the discussion during the STAR 3 review by (1) questioning and asking for clarification on monitoring/sampling program designs, statistical analyses, assessment model configuration, assumptions, uncertainties of various sources, and interpretations; (2) commenting on the assessment and review processes; (3) making constructive comments and suggestions for alternative approaches and additional analyses; (4) interpreting analysis results and potential issues; and (5) contributing to the development of the STAR 3 Review Panel report. I had also been interacting with relevant scientists and other panel members regarding issues raised during the review process for further clarifications and discussion during the breaks and after the review.

IV. Summary of Findings

My detailed comments on each item of the ToRs are provided under their respective subtitles from the ToRs (see below).

1. *Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.*

   The draft stock assessment report, input data, control file, and all the output data for the "initial" base case model run were made available online about two weeks before the review. The relevant background information was also made available online including previous STAR panel reports, 2011 CIE reports, the 2018 International Sablefish Workshop report,

peer-reviewed publications, and previous Sablefish stock assessment reports. I downloaded all of the information provided and read through all documents prior to the STAR 3 review. Potential questions and possible model runs were identified and noted, and a preliminary plan was developed for developing sensitivity model runs for testing and modifying the "initial" base case model proposed by the STAT.

The information provided is adequate and in a timely fashion to familiarize me for the STAR 3 review.

2. *Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.*

Input data

Catch data: Landing data from 1890 were reconstructed after the last review (2011) using data from different sources. The quality of the estimates was questioned and discussed, and I am convinced that the West Coast Sablefish catch data quality is adequate for the assessment.

Discards: Limited information is available for recent years, but not early years. Discards were estimated in the model. Discard mortality was assigned different values for the trawl, hook-and-line and pot. Although these values are consistent with those assumed in the management, there is little field work to support these numbers. During the review, it was suggested that four years of size composition data for discards be included in the model run; the model was found to be very sensitive to the discard size composition data. More studies are needed to better quantify the discards, including estimation of discards and mortality of fish discarded in the three fisheries as well as quantification of the size compositions of discarded catch. The reasons for this substantial impact of discard size composition data on the stock assessment (the effect of which resulted in large changes in the scale of stock biomass estimates) remains unknown. To reduce this effect, the discard length composition data were initially included to estimate retention models, and then the estimated retention models were fixed at those estimated parameters, which were used in the "final" base case model with the exclusion of discard length composition data.

Size composition data: The data seem to be well estimated with sufficiently large sample sizes for both fisheries and surveys. Although spatial variabilities may exist for both fisheries and survey data, they were not explicitly examined. The cumulative size composition data were compared among the surveys and fisheries to advise the choices of selectivities.

Age composition data: Ages are not validated, but various verification studies have been done to estimate possible precisions. It is inconclusive whether ageing results are biased or not. Tagging data were only available for young fish. Studies (e.g., tagging study with old fish) are needed to validate the ageing results. Selectivities are assumed to be age-dependent, rather than length-dependent. Because age composition and size composition data were derived from the same data set in a survey program or fishery, the information may be replicated in the assessment. Interestingly, age composition and size composition data tended to provide contradicting information in the parameter estimation based on the analysis of likelihood profiles. Thus, we may want to avoid including both size- and age-composition data in the assessment except those data (i.e., WCGBTsurvey data) that are required to estimate growth parameters.

Natural mortality: Priors were assumed to follow a lognormal distribution with lower and upper boundaries of 0.01 and 0.11. Natural mortality was assumed to be sex-specific with no spatio-temporal variability. The similarity between these estimates led to a request to have a sex-combined natural mortality. This, however, resulted in relatively large changes in the natural mortality estimate and caused issues with model convergence. The sensitivity of the assessment on this assumption may need further evaluation. Age-specific natural mortality may also need to be explored in future assessments, given that it is very likely natural mortality may vary with ages/sizes for Sablefish.

Steepness: Previous stock assessments suggest that there is no information in the data to estimate this parameter for Sablefish, and a fixed value of 0.7 is used after some likelihood profile analyses, which may result in the underestimation of the uncertainty.

Other life history parameters (e.g., maturity, length at age, weight at age): no temporal and spatial variabilities were considered, which may not be realistic. For example, the observed average weight at age from the WCGBTS varies from year to year, but internally calculated weight-at-age data in the model were assumed to be constant over time; large systematic differences could be observed over time between the observed and calculated weight-at-age data, perhaps resulting in biased estimates of fish stock biomass.

Survey indices: Four survey indices were used in the assessment. The WCGBTS covers most of the stock area for the West Coast Sablefish, and plays a critical role in the assessment of this stock. This survey spans more than 5 months in the survey each year. Although two passes were incorporated in the survey to address potential issues regarding possible changes in the Sablefish population, possible fish movement and area-specific fishing activities during the survey may still affect the results of the survey. Possible changes in fish spatio-temporal distributions may affect the effectiveness of the survey design in capturing the stock dynamics over time. Because of how important this survey program is, these issues should be evaluated in future studies. The two slope surveys (Northwest Fisheries Science Center Slope survey and Alaska Fisheries Science Center Slope Survey) cover the part of stock area that is not covered by the WCGBTS and provide valuable information that may not be represented in the WCGBTS data. Triennial Shelf Survey, conducted by the Alaska Fisheries Science Center since 1980 (but ended in 2004), covers areas with depths from 55 to 500 m from 34.5N to the Canadian border. It provides information on younger Sablefish abundance. Overall, the survey indices were used properly and uncertainty structures assumed (i.e., only modelled errors for WCGBTS, but modelled errors and additional error terms estimated in the stock assessment for the other surveys) in the model are reasonable. However, the models used to derive the model-based abundance indices tended to be unstable with the addition of new data over time, which needs to be evaluated in future studies.

Analytical methods

Stock assessment model: An integrated statistical catch-at-age model implemented in Stock Synthesis (SS) (version 3.30.13 released 2019-03-09) was used for the West Coast Sablefish stock assessment. The model was pre-configured and pre-parameterized by STAT before the STAR 3 review. A pre-STAR 3 review base case model was proposed by STAT and was considered as an "initial" base case model. After almost a week-long discussion, numerous new models runs with different model configurations and parameterization, and careful

evaluations and analyses of all the modeling results, the STAR Panel and STAT come to an agreement for a "final" base case model which can be used to provide management advice. The SS3 model has been well-studied and widely used, and is appropriate and adequate to serve as the West Coast Sablefish stock assessment modeling framework. The "final" base case model identified and configured in the STAR 3 review is adequate for providing management advice for the West Coast Sablefish stock.

Recruitment dynamics: The Beverton-Holt stock-recruit function was used to quantify recruitment dynamics. The function was parameterized with two estimated quantities: the log of unexploited equilibrium recruitment (R0) and the steepness h. Sea level, which is found to be related to Sablefish recruitment dynamics in other studies, was also included as an index in the assessment model to improve the estimation of recruitments. However, the linear function assumed in the assessment model was inconsistent with the function identified in an independent ecosystem study. Sea level index was found to have limited contributions to the improvement of recruitment estimation, and tended to under-estimate strong year classes in recent years. Like other survey indices (except for the WCGBTS, which only has a fixed variance term estimated outside the stock assessment model), an additional estimable variance term was added to the sea level index to allow the stock assessment model to estimate the variance term (and thus weighting in modeling). I commend the STAT's effort to include the ecosystem consideration in the assessment. However, more studies are needed in order to identify appropriate functions that can ensure consistency between studies and also to evaluate and better understand possible interactions between the environmental drivers and other fishery and life history processes (e.g., natural mortality, growth and steepness) included in the stock assessment model.

Selectivity: Selectivity was assumed to be age-dependent for all surveys because Sablefish spatial distribution is more likely age-dependent than size-dependent. The pre-STAR 3 base case model assumes that WCGBTS follows asymptotic (logistic) selectivity and all other surveys follow dome-shaped selectivities. Age- and size composition data from different surveys were compared in the review and no substantial differences could be found among the surveys. After several model runs of different selectivity configurations for each survey, the STAR Panel and STAT agreed to allow all the surveys an age-based dome-shaped selectivity for the final base case model. For the fishery selectivities, time blocks were used to account for temporal changes in fishing fleet dynamics as a result of changes in management regulations. However, spatial variability, likely existing in fishing fleet dynamics, was not considered in configuring fishery selectivity in this assessment. Given the change in spatial dynamics of trawl, hook-and-line and pot fisheries in recent years, it is important to account for this change. Future studies can explore the possibility of developing selectivity functions based on the spatial locations of the fishery. A more or less ad hoc practice of fixing certain parameters in the assessment also needs to be further evaluated.

Growth model: The von Bertalanffy growth parameters were internally estimated in the model. The length-at-age data were then applied to the weight-length model to calculate weight-at-age data, which were used to calculate stock biomass in the model. No spatial or temporal variability was considered. However, a review of average weight-at-age data from the WCGBTS survey data suggests a temporal trend, which raises concern for systematic biases if constant weight-at-age data in estimating stock biomass were used. It may be

difficult to estimate yearly weight-at-age data internally with the current model configuration, but use of the average weight-at-age calculated from the WCGBTS survey may be a good approximation to account for temporal changes in weight-at-age for the stock biomass estimation. The growth model was internally estimated, and its estimation is likely not to be independent of the estimation of other life history and fishery processes included in the model. Thus, it is important to evaluate the biological realism of the estimated growth model. For Sablefish, this means evaluating the biological realism of estimated length at young ages given the exceptionally high growth rates in the first decade of their lives.

Assessment modeling duration: During the review, the necessity of starting the stock assessment in 1890 was questioned. A request was made to run the model from 1970 to evaluate the sensitivity of the estimation of unfished spawning biomass and SSB trends in the absence of early age composition data. The STAR Panel and STAT evaluated the changes in the SSB and depletion estimates for the model with a start year of 1970 compared to the model starting in 1890. Virgin SSB and recent SSB values were found to depend on recent data. Choice of start years tends to have limited impacts on the SSB and depletion values of start and end of time series. Although this suggests that modeling time may not be necessary to start in 1890 for the West Coast Sablefish stock, the availability of the reconstructed Sablefish catch time series data and increasingly widespread practice of assessing other groundfish stocks on the West Coast makes the start time of 1890 more desirable.

Model configurations: The data available may not be sufficient enough to estimate all parameters in the stock assessment model, making the model performance less stable and sensitive to changes in model configuration. Thus, there is a need to either fix some parameters and/or combine some processes based on our understanding of Sablefish biology and fishery. Combining fixed gear (i.e., hook-and-line and pot) makes the model more stable and improves the model fitting. Fixing some parameters (e.g., steepness h) may also improve the model performance. However, it is important to keep in mind that this type of adjustment makes implicit assumptions about these fixed parameters (e.g., steepness and selectivities), perhaps leading to the under-estimation of uncertainty in stock assessment.

3. *Evaluate model assumptions, estimates, and major sources of uncertainty.*

During the STAR 3 Review, the STAR 3 Panel worked with STAT to evaluate a series of model configurations and associated modeling results in an effort to develop a plausible base case scenario. The development of model configurations followed the principle of adjusting one variable at a time so that observable changes in modeling can be attributed solely to the specific variable change. The daily requests for additional analyses and model runs can be found in Appendix III. These additional model runs and analyses were completed to improve our understanding of the model fitting, impacts of uncertainty in data and models, robustness of assessment results, as well as the roles of different life history and fishery processes in modeling. This comprehensive understanding helps lead to the identification and finalization of the base case scenario and other possible sensitivity analysis scenarios for providing management advice.

The West Coast Sablefish stock assessment, like its previous assessments, assumes a unit Sablefish stock in the assessment area with no linkage to other stocks in the Northwestern Pacific Ocean. This assumption on stock structure is likely not true given the genetic and tagging studies. The uncertainty caused by this assumption in the stock assessment has yet to

be evaluated and thus remains unknown. An international workshop was held in 2018 to develop and condition a management strategy evaluation (MSE) framework for Sablefish in the Northwestern Pacific Ocean. This MSE, if developed and conditioned on the whole region Sablefish, will be an ideal platform to evaluate uncertainty resulting from different assumptions on the Sablefish stock structure.

No spatial variability was considered in Sablefish life history and fishing fleet dynamics (and fishermen's behaviors). However, previous studies provided strong evidence of latitudinal variations in life history parameters and variation in the behavior of fishing fleets (e.g., selectivity, retention/discard) among different areas.

Temporal variability was considered in the fishery by applying time blocks in the model to reflect changes in fishing fleet behavior that were a result of regulations adjustments over time. However, temporal variability was not considered in life history processes and key life history processes such as growth and maturation were assumed constant over time. An examination of average weight-at-age data calculated from the WCGBTS suggests a temporal trend, which, if not accounted for in the estimation of stock biomass, may introduce biases in the assessment.

Using sex-combined natural mortality resulted in the failure of certain model configurations to converge. The reason for this is unclear, but the fact that the model performance was so sensitive to this change needs to be further investigated.

Many runs of models with different configurations and parameterizations were conducted prior to and during the review. Although temporal trends of stock biomass and exploitation rates tend to be robust to changes in model configurations, absolute values of stock biomass and exploitation rate over time are sensitive. This suggests that there is inadequate information to scale the West Coast Sablefish stock dynamics in the assessment, which may be improved with the addition of the more data from the WCBGTS program.

The assumptions presumed in the survey and fishery selectivities had large impacts on the stock assessment, which is to be expected. The choice of selectivity functions was mainly determined by evaluating cumulative age/size distribution data from the surveys and fisheries in the review. Given time constraints, this was an adequate approach. However, a more careful evaluation of possible impacts of spatial variability in the distributions of Sablefish and fishing fleets may be needed after the development of the MSE framework for the Northwestern Pacific Sablefish population.

The assessment uncertainty was quantified with asymptotic uncertainty estimates together with sensitivity analyses and likelihood profile analyses. Although SS includes Monte Carlo Markov Chain (MCMC) methods, time constraints did not permit the use of MCMC in this study. It remains an area to be explored in the future.

Almost all the model parameters were estimated internally in the model. Because of the complexity of the model configuration and confounding nature of length-based and age-based processes, it is difficult to interpret and understand changes resulting from adjustments in the model configurations. It might be useful to estimate some parameters outside the SS or fix the parameters estimated in the SS model runs in order to improve SS modeling efficiency and performance as well as identify key processes/parameters in the stock

assessment. This modeling strategy was used in the review to help identify and finalize the base case model.

The "final" base case model identified at the end of the STAR 3 review differed greatly from the "initial" base case model proposed prior to the review. After extensive discussions and many structured model run requests, the STAR Panel and STAT agreed upon a model that could be used as the base case model to provide management advice. This model includes the following key configurations: (1) allowing dome-shaped age-based selectivity for all surveys; (2) use of the WCGBTS age and length composition data to estimate growth model using CAAL data; (3) exclusion of length composition data of all other sources; (4) estimation of retention model with the initial inclusion of discard length composition data and then excluding discard length composition data (to avoid their impacts on the assessment which was identified as inappropriate); (5) pooling fixed gears with key selectivity parameters being estimated; (6) the use of the Francis weighting method for composition data; and (7) inclusion of sea level with an additional variance term being estimated (like other indices).

Almost no sensitivity analysis was done for the "final" base case model because it was only finalized late in the last day of the review. However, the numerous model runs completed during the review period for the purpose of identifying the "final" base case model could be considered as sensitivity analyses, which greatly improves understanding of the model's behavior as well as any possible interactions between different components of the model.

Overall, although stock assessment modeling results varied among different model configurations and parameterizations, this assessment clearly suggests that the West Coast Sablefish is currently not undergoing overfishing and is currently not overfished. The stock was estimated to be above the depletion level that would lead to maximum yield.

4. *Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.*

The approach used to estimate retention curves with the inclusion of discard length composition data and then fix the estimates and exclude discard length composition data in the "final" base case model is interesting. This approach was identified at the end of the review and was not carefully evaluated regarding possible impacts on the estimation of other parameters. Further studies are needed to evaluate its impacts.

An examination of average weight-at-age data calculated from the WCGBTS suggests a temporal trend, which, if not accounted for in the estimation of stock biomass, may introduce biases in the assessment. Because of a large number of years and ages defined in the stock assessment for Sablefish, estimating yearly weight-at-age data internally would be too time consuming and could also confound the estimation of other parameters. An adequate approach would be to calculate average weight-at-age matrices or estimate yearly weight-at-age model externally from the WCGBTS, which can capture temporal variability in growth.

Retrospective analysis was conducted in the late afternoon of the last day of review for the "final" base case model. Although we only briefly evaluated its results because of time constraints, it was clear that stock biomass tended to be under-estimated in the retrospective

analysis. Despite not being risk-prone bias, the under-estimation of stock biomass (and over-estimation of exploitation rates) shown in the retrospective analysis could have negative impacts on the development of management advice. Mohn's rho should be calculated to determine the magnitude of the retrospective errors. There is a need to determine if the retrospective errors need to be corrected for the determination of stock status and projection of future catch.

5. *Determine whether the science reviewed is considered to be the best scientific information available.*

The lack of robustness of the modeling results regarding some alternative model assumptions and configurations as well as several unusual model behaviors observed during the review process raised doubt on the suitability of the "initial" base scenario, resulting in a large number of requests for different model runs and extensive discussions. This led to the late identification and finalization of the base case scenario towards the end of the review period, leaving little time for a detailed, in-depth discussion on possible sensitivity analyses for the defined "final" base case model, model projection, and decision table development.

Overall, based on the materials presented, extensive discussions and additional model runs conducted during the review, I conclude that the "final" base case model and resultant stock assessment results represent the best available practice and information on the fishery status for the West Coast Sablefish stock assessment. However, large uncertainties still exist in the assessment, especially with regard to the data and models. I consider the information reviewed scientifically sound and representative of the best available information we have.

6. *When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.*

Short-term research recommendations:

- Temporal variability in growth (i.e., weight-at-age) needs to be further evaluated using the WCGBTS data. Yearly average weight-at-age vector or yearly weight-at-age model externally estimated from the WCGBTS data can be estimated and used in the estimation of stock biomass.
- Separate fixed gear fleets between north and south of 36^0N and other possible spatial differences in fleet dynamics need to be evaluated to evaluate and identify possible causes for strong tensions observed between age and length composition data in model fitting.
- Retrospective analysis results need to be carefully evaluated to determine whether retrospective errors should be corrected for the determination of stock status and the projection of future catch.
- The importance of WCGBTS program can be evaluated by conducting a sensitivity analysis without abundance indices from this program and/or only with portion of the time series in the assessment.
- Age-specific selectivities fixed at plus group age should be evaluated.

- The use of double-normal selectivities for the fixed gear fishery may need to be evaluated because of its unexpected behaviors observed in the model parameter estimation. Parameters fixing in the estimation needs to be better justified.
- Length-based selectivities should be explored for the WCGBTS to reduce the likelihood of having biased growth estimates for young fish (age 0).
- Differences between observed and predicted sex ratio should be carefully evaluated to improve our understanding of the needs for sex-specific parameterization and life history/fishery processes in modeling.

Medium- and long-term research recommendations:

- The models used to derive the model-based abundance indices tended to be unstable with the addition of new data over time, which needs to be evaluated in future studies.
- There is a need to develop a spatially explicit and process-based MSE for improved understanding of stock dynamics in its distributional range and identify key processes/factors that may influence the performance of the stock assessment models. Such a framework can also be used to evaluate the potential impacts, including spatial structure of the West Coast Sablefish stock and fishery on the stock assessment and management.
- The distribution of Sablefish has been observed to change over time, which is likely linked to changes in the ecosystem (e.g., climate change). Such a distributional change may influence the effectiveness of the existing survey program and may also result in changes in the spatio-temporal distribution of fishing fleets, which may influence the performance of various fisheries monitoring programs (e.g., sea sampling and port sampling program). Thus, research is needed to evaluate the effectiveness of current monitoring programs as well as determine whether the monitoring program designs should be adjusted given perceived changes in the ecosystem.
- Ecosystem consideration is included in the STAR 3 stock assessment. However, its role was relatively limited. Sea level was used as an ecosystem dynamic proxy in the modeling recruitment dynamics. However, the linear function assumed between the sablefish recruits and sea level was inconsistent with the relationship identified in the ecosystem study. More studies are needed to improve the process of incorporating ecosystem considerations in the stock assessment.
- There are large numbers of size/age bins and plus groups for the Sablefish. Many age/size bins may not have any observation or may play limited roles in population dynamics. Model fitting may put unnecessary weight in these zero-observation bins. A dynamic binning approach may be necessary to avoid overfitting those bins or groups.
- Even though the current asymptotic uncertainty may be adequate for defining the uncertainty associated with the stock assessment results, I recommend that MCMC be explored to better capture this uncertainty.
- Age validation study should be done to improve estimation of ageing errors.

7. *Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations*.

Using fishery-dependent and fishery-independent data, an integrated statistical catch-at-age model implemented in Stock Synthesis (SS) (version 3.30.13 released 2019-03-09) was pre-configured and pre-parameterized by the STAT before STAR 3. After almost a week-long discussion, numerous new models runs with different model configurations and parameterization, as well as careful evaluations and analyses of all the modeling results, the STAR Panel and STAT come to an agreement for a base case model that can be used to provide management advice. The STAT, Northwest Fisheries Science Center (NWFSC) and Pacific Fisheries Management Council (PFMC) provided the necessary logistics support, documentation, data, and background information I requested. The stock assessment analysts involved in the process were open to suggestions, provided additional information upon request, and conducted many additional model runs identified by the STAR Panel. Although almost no sensitivity analysis was done for the base case because it could only be finalized at the end of the last day of the review, the numerous model runs during the review period can essentially be considered as sensitivity analyses, which helped greatly in understanding the behavior of the model and possible interactions of different components in the model. The STAT analysts accommodated all of the requests made for different model runs as well as additional biological and fishery information. The review process was open and constructive, and all materials were sent to me in a timely manner.

My only recommendation for future improvement would be to include presentations to cover the (1) design and history of all survey programs included in the stock assessment; (2) management structure and dissemination of the STAR review and stock assessment; and (3) dynamics of fishing fleets (both spatial and temporal dynamics). These background presentations would help those not from the area who may not be familiar with the monitoring and management systems better understand key background information necessary to improving the review process.

V. Conclusions and Recommendations

After a week of discussion, numerous new model runs with different model configurations and parameterizations, as well as careful evaluations and analyses of modeling results, the STAR Panel and STAT have come to an agreement for a base case model, which is considered by the STAR Panel and STAT adequate for providing management advice.

There are large uncertainties in the Sablefish stock assessment, although the relative stock trend over time tends to be more robust regarding different model configurations compared with the estimation of absolute stock biomass. These uncertainties come from different data weighting methods, inclusion/exclusion of certainty data (e.g., discard length composition data, length and age composition data), different assumptions on survey and fishery selectivities (e.g., dome versus logistic curves) as well as model parameterizations (e.g., sex-specific life history parameters like natural mortality and fixing some model data in parameters estimation), and possible spatio-temporal variations in life history parameters (e.g., growth). The large uncertainties and unusual model behaviors observed during the review process resulted in

extensive discussions and a large number of requests for different model runs, which culminated in the late identification and finalization of the base case model towards the end of the review period. Thus, conducting a sensitivity analysis for the "final" base case model in the review is essentially impossible. However, numerous model runs with different model configurations and parameterizations have improved our understanding of the model performance and can be considered as alternative sensitivity analyses (despite not having been done with the base case model).

Overall, based on the materials presented, extensive discussions as well as additional model runs conducted during the review, I conclude that the "final" base case model and resultant stock assessment results represent the best available practice and information on the fishery status for the West Coast Sablefish stock assessment. However, there is a need for improved understanding of (1) spatio-temporal variations in life history parameters (e.g., growth and maturation), fishing fleet behaviors (e.g., selectivity and discard/retention) and associated implications on stock assessment; (2) statistical interactions of various SS model components in parameter estimation; and (3) implications of various assumptions made on the model configurations and parameterizations. These issues can be addressed in short- and medium-term research recommendations identified in Section IV-6.

VI. References cited

DFO. 2016. A revised operating model for sablefish (*Anoplopoma fimbria*) in British Columbia, Canada. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2016/015.

Fenske, K. H., A. M. Berger, B. Connors, J.M. Cope, S. P. Cox, M. A. Haltuch, D. H. Hanselman, M. Kapur, L. Lacko, C. Lunsford, C. Rodgveller, and B. Williams. 2019. Report on the 2018 International Sablefish Workshop. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-387, 107 p.

Gertseva, V., Matson, S. and Cope, J. 2017. Spatial growth variability in marine fish: example from Northeast Pacific groundfish. ICES Journal of Marine Science 74(6): 1602–1613.

Haltuch, M.A., K. F. Johnson, N. Tolimieri, C. A. Castillo-Jordán, and M. R. Kapur. 2019. Status of the sablefish stock in U.S. waters in 2019.

Hanselman, D., Heifetz, J., Echave, K. and Dressel, S. 2015. Move it or lose it: movement and mortality of sablefish tagged in Alaska. Canadian Journal of Fish and Aquatic Sciences 72:238–251. doi:10.1139/cjfas-2014-0251.

Hart, J. 1973. Pacific Fishes of Canada. Tech. Rep. 180, Fisheries Research Board of Canada Bulletin, St. Andrews, NB, Canada.

Head, M., Keller, A. and Bradburn, M. 2014. Maturity and growth of sablefish, Anoplopoma fimbria, along the U.S. West Coast. Fisheries Research 159: 56–67.

Jasonowicz, A., Goetz, F., Goetz, G. and Nichols, K. 2017. Love the one youâˇA ´ Zre with: genomic evidence of panmixia in the sablefish (Anoplopoma fimbria). Canadian Journal of Fisheries and Aquatic Science 74: 377–387.

Johnson, K., Rudd, M., Pons, M., Akselrud, C., Lee, Q., Hurtado-Ferro, F., Haltuch, M. and Hamel, O. 2016. Status of the U.S. sablefish resource in 2015. Tech. rep., Pacific Fishery

Management Council, Portland, OR. Available from http://www.pcouncil.org/groundfish/stock-assessments.

VII-1: Appendix 1: Bibliography of materials provided for review and list of presentations

Fenske, K. H., A. M. Berger, B. Connors, J.M. Cope, S. P. Cox, M. A. Haltuch, D. H. Hanselman, M. Kapur, L. Lacko, C. Lunsford, C. Rodgveller, and B. Williams. 2019. Report on the 2018 International Sablefish Workshop. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-387, 107 p.

Gertseva, V., Matson, S. and Cope, J. 2017. Spatial growth variability in marine fish: example from Northeast Pacific groundfish. ICES Journal of Marine Science 74(6): 1602–1613.

Haltuch, M.A., K. F. Johnson, N. Tolimieri, C. A. Castillo-Jordán, and M. R. Kapur. 2019. Status of the sablefish stock in U.S. waters in 2019.

Haltuch, M. A., A'mar, Z. T., Bond, N. A., and Valero, J. L. 2019. Assessing the effects of climate change on US West Coast sablefish productivity and on the performance of alternative management strategies. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsz029.

Hanselman, D.H., C. J. Rodgveller, K. H. Fenske, S. K. Shotwell, K. B. Echave, P. W. Malecha, and C. R. Lunsford. 2018. Assessment of the Sablefish stock in Alaska.

Head, M.A., A.K. Keller, M. Bradburn. 2014. Matyurity and growth of sablefish, Anoplopoma fimbria, along the U.S. West Coast. Fisheries Research 159:56-67.

Johnson, K.F., M. B. Rudd, M. Pons, C. A. Akselrud, Q. Lee, F. Hurtado-Ferro, M. A. Haltuch, and O. S. Hamel. 2016. 2015 Status of the US sablefish resource in 2015. Northwest Fisheries Science Center.

Kupschus, S. 2011. Independent Peer Review Report on the STAR Panel for Sablefish and Dover Sole, held 25–29 July 2011, in Newport, Oregon.

Miller, S., A. Stephens, C. Whitmire, and J. Hastie. Overview of West Coast Groundfish Fishery-Independent surveys. Northwest Fisheries Science Center.

Stewart, I., J. Thorson, and C. Wetzell. 2011. Status of the U.S. sablefish resource in 2011. Final SAFE version. Northwest Fisheries Science Center.

Stokes, K. 2011. Report on the Stock Assessment Review (STAR) Panel for Dover Sole and Sablefish. The Center for Independent Experts.

Tolimieri, N., Haltuch, M., Lee, Q., Jacox, M. and Bograd, S. 2018. Oceanographic drivers of sablefish recruitment in the California Current. Fisheries Oceanography 27: 458–474.

Wespestad, K. Stokes, S. Kupschus, and D. Sampson. 2011. Sablefish STAR Panel Report.

List of presentations

- 2019 Sablefish stock assessment: Data
- 2019 Sablefish stock assessment: Model
- Ecosystem and socioeconomic considerations for sablefish
- Oceanographically determined growth variation for NE Pacific Sablefish
- 2019 Sablefish stock assessment: Day 1 requests
- 2019 Sabelfish stock assessment: Day 2 requests
- 2019 Stablefish stock assessment: Day 3 requests
- 2019 Sablefish stock assessment: Day 4 requests
- 2019 Stablefish stock assessment: Day 5 requests

VII-2: Appendix II: Performance Work Statement

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

Stock Assessment Review (STAR) Panel 3

Background

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

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

Scope

The National Marine Fisheries Service and the Pacific Fishery Management Council will hold four stock assessment review (STAR) panels and potentially one mop-up panel if needed, to evaluate and review benchmark assessments of Pacific coast groundfish stocks. The goals and objectives of the groundfish STAR process are to:

1) ensure that stock assessments represent the best scientific information available and facilitate the use of this information by the Council to adopt Overfishing Limits (OFLs), Acceptable Biological Catches (ABCs), Annual Catch Limits (ACLs), harvest guidelines (HGs), and annual catch targets (ACTs);
2) meet the mandates of the Magnuson-Stevens Fisheries Conservation and Management Act (MSA) and other legal requirements;

3) follow a detailed calendar and fulfill explicit responsibilities for all participants to produce required reports and outcomes;
4) provide an independent external review of stock assessments;
5) increase understanding and acceptance of stock assessments and peer reviews by all members of the Council family;
6) identify research needed to improve assessments, reviews, and fishery management in the future; and
7) use assessment and review resources effectively and efficiently.

A benchmark stock assessment will be conducted and reviewed for Sablefish. The sablefish stock was identified as the top ranked candidate for assessment during the Pacific coast groundfish regional stock assessment prioritization process, which was based on the national stock assessment prioritization framework (http://www.st.nmfs.noaa.gov/Assets/stock/documents/PrioritizingFishStockAssessments_FinalWeb.pdf.

Sablefish is one of the most important groundfish stocks on the West Coast and the most commercially valuable groundfish stock on a per pound basis. Sablefish is a major target species in commercial trawl and non-trawl fisheries and is readily caught with trawls, longlines, and sablefish pots/traps on the shelf and slope is an important component of the west coast groundfish fishery. The last full assessment of sablefish was in 2011 with an update completed in 2015. The update assessment indicated spawning biomass to be 34.5 percent of its unfished level in 2015. Following the review of the 2011 update assessment, the SSC recommended the next assessment of this stock be a full assessment.

An assessment for the sablefish stock will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. including providing scientific basis for setting OFLs and ABCs as mandated by the Magnuson-Stevens Act. The technical review will take place during a formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements
Two CIE reviewers will participate in the stock assessment review panel. One CIE reviewer shall conduct an impartial and independent peer review of the assessments described above and in accordance with the Performance Work Statement (PWS) and ToRs herein. Additionally, one "consistent" CIE reviewer will participate in all STAR panels held in 2019 and the PWS and ToRs for the "consistent" CIE reviewer are included in **Attachment A**.

The CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers

shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age-and size-structured models, use of Markov Chain Monte Carlo (MCMC) to develop confidence intervals, and use of Generalized Linear Models in stock assessment models. The CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Tasks for Reviewers
The CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

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

Documents to be provided to the CIE reviewers prior to the STAR Panel 3 meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available (including previous stock assessments and STAR panel reports).
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

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

Contract Deliverables - Independent CIE Peer Review Reports: The CIE reviewers shall complete an independent peer review report in accordance with the PWS. Each CIE reviewer shall complete the independent peer review according to required format and content as described

in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

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

Timeline 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 STAR Panel 3 review meeting in scheduled in Seattle, WA during the dates of July 8-12, 2019 as specified herein, and conduct an independent peer review in accordance with the ToRs.
3) No later than July 26, 2019, each CIE reviewer shall submit their draft independent peer review report to the contractor. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**

Foreign National Security Clearance

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

Place of Performance

The place of performance shall be at the contractor's facilities, and in Seattle, WA.

Period of Performance

The period of performance shall be from the time of award through September 2019. The CIE reviewers' duties shall not exceed 14 days to complete all required tasks.

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

Within two weeks of award	Contractor selects and confirms reviewers
At least two weeks prior to the panel review meeting	Contractor provides the pre-review documents to the reviewers
July 8-12, 2019	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 26, 2019	Contractor receives draft reports
August 9, 2019	Contractor submits final reports to the Government

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

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract.

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

NMFS Project Contacts:
Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive
Newport, OR 97365
Stacey.Miller@noaa.gov
Phone: 541-867-0535

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Jim.Hastie@noaa.gov
Phone: 206-860-341

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

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, 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 Summary Report that they feel might require further clarification.

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

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

Stock Assessment Review (STAR) Panel 3

8. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
9. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
10. Evaluate model assumptions, estimates, and major sources of uncertainty.
11. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
12. Determine whether the science reviewed is considered to be the best scientific information available.
13. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
14. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Annex 3: Tentative Agenda

Final Agenda to be provided two weeks prior to the meeting with draft assessments and background materials.

Stock Assessment Review (STAR) Panel 3
NMFS Northwest Fisheries Science Center
2725 Montlake Blvd, NE
Seattle, WA 98112
NOAA Fisheries, Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle, Washington 98112
July 8-12, 2019

To join the webinar in listen-only mode, click: https://nwfscfram.webex.com/nwfscfram

1. Enter the Webinar Meeting Number: 623 016 027
2. Enter your name and email address (required).
3. AFTER logging in to the webinar, connect audio by computer or dial the TOLL number: 1-650-479-3208. *NOTE: You will be muted upon entry to the webinar.*

Monday, July 8

8:30 a.m.	Welcome and Introductions
8:45 a.m.	Review the Draft Agenda and Discuss Meeting Format (Chair) - Review the Terms of Reference (TOR) for assessments and STAR panel responsibilities - Assign reporting duties - Agree on time and method for accepting public comments
9:15 a.m.	STAT Presentation of the Sablefish Assessment - Overview of data and modeling
12:30 p.m.	Lunch (Onsite)
1:30 p.m.	Continue STAT Presentation of the Sablefish Assessment
4:00 p.m.	STAR Panel Discussion - Panel develops written requests for first set of model runs / analyses
5:30 p.m.	Adjourn for day

Tuesday, July 9

8:30 a.m.	Review of Agenda Topics for the Day and STAR Panel Discussion
10:30 a.m.	STAT Presentation of first set of model runs and analyses
12:30 p.m.	Lunch on your own
1:30 p.m.	STAR Panel Discussion - Panel develops written requests for second set of model runs / analyses
5:30 p.m.	Adjourn for day

Wednesday, July 10

8:30 a.m. Review Agenda for the Day

9:00 a.m. STAT Presentation of second set of model runs and analyses

12:30 p.m. Lunch on your own

1:30 p.m. STAR Panel Discussion
- Panel develops written requests for third set of model runs / analyses

5:30 p.m. Adjourn for day

Thursday, July 11

8:30 a.m. Review Agenda for the Day

9:00 a.m. STAT Presentation of third set of model runs and analyses

12:30 p.m. Lunch on your own

1:30 p.m. STAR Panel Discussion
- Panel and STAT agree on final base model, develop decision table

5:30 p.m. Adjourn for day

Friday, July 12

8:30 a.m. Consideration of Remaining Issues
- Review decision tables for all assessments

11:00 a.m. Review First Draft of the STAR Panel Report
- Panel Agrees to Process for Completing the Final STAR Report for Council's September Meeting Briefing Book (Requested by August 15th)

12:30 p.m. Lunch on your own

1:30 p.m. Continue Drafting Report as needed

4:00 p.m. Review Panel Adjourns

VII-3: Appendix III: Panel Membership or other pertinent information from the panel review meeting.

Panel Membership

Panel Members

John Field, National Marine Fisheries Service Southwest Fisheries Science Center (Chair)
Jim Ianelli, National Marine Fisheries Service Alaska Fisheries Science Center
Yong Chen, Center for Independent Experts
Robin Cook, Center for Independent Experts

Stock Assessment Team (STAT) Members

Melissa Haltuch, National Marine Fisheries Service, Northwest Fisheries Science Center
Kelli Johnson, National Marine Fisheries Service, Northwest Fisheries Science Center
Nick Tolimieri, National Marine Fisheries Service, Northwest Fisheries Science Center
Maia Kapur, School of Aquatic and Fishery Sciences, University of Washington
Claudio Castillo-Jordán, School of Aquatic and Fishery Sciences, University of Washington

STAR Panel Advisors

Patrick Mirick, Oregon Department of Fish and Wildlife, Groundfish Management Team representative
Gerry Richter, B&G Seafoods, Groundfish Advisory Subpanel representative
John DeVore, Pacific Fishery Management Council representative

List of requests for additional model runs during the STAR 3 review

Day one requests

The following were requested

1) Show the SS weight-at-age (over time) results and compare with NWFSC WCGBTS weight-at-age data, if possible.
   Rationale: The growth model is embedded in the assessment model and the variability (or lack thereof) may differ from the data. Also, to see if there's a temporal / year-effect pattern (e.g., due to strong year class(es) that may have cohort effect / density dependence).
2) Plot cumulative size distribution for WCGBTS using the AKSLP survey footprint (N of 36º and deeper than 100 fathoms) and compare with the AKSLP cumulative length frequencies (over all years).
   Rationale: The issue of setting the WCGBTS selectivity to be asymptotic is a change from past assessments and data supporting this specification, external to the model might be useful. Also, this may provide some justification for specifying asymptotic selectivity for the AKSLP survey data.
3) Examine recruitment estimates from the Base model and compute ratio of the sea level (SL) index to derive a q variability (CV) estimate (prior variance). Compare this with the assumed CV.
   Rationale: For the more informed period (e.g., 1980-2017 when survey are available), the recruits are based on age data and SL data may have little impact. The variability estimated from this period could be used for the prior for the variability used when recruitment data are less commonly available. This is to provide a more objective approach to specify the level of process error that might exist between SL and actual recruitment.
4) Starting with base-run model including WCGOP composition data (added post May 29), document July 1st and July 6th model changes, incrementally by characteristic (cumulative):
   - Free young fish CV at age 0.5
   - Free selectivity and retention parameters (the P6) and
   - Include time-varying sea level catchability (SL q) deviation vector (and note assumed CV/prior)
   - Combine HKL and POT fisheries into one. This reduces complexity and parameter estimation issues

   Include figures reflecting changes to each of these aspects. Specifically

   - How did CV change? Distributions of length at age in growth plots
   - Selectivity curves changed
   - Retention curves
   - For SL q deviations, examine the time series of the values to evaluate variability in q
   - Fits to length composition data to the new combined, HKL+POT fleet (residuals of combined compared to when split)

   Rationale: STAT made changes prior to meeting and this will aid in understanding the impact of the changes.

Request if time allows (perhaps day 2 or later)

5) Test a case and just assume all bycatch/discards are dead rather than the current management values.

Rationale: This issue arose in discussions of discard mortality estimates being poorly determined/estimated. This is just a sensitivity to highlight the relative importance of a field study to better estimate/revise given all the management changes (tow duration etc).

6) Test a run a shorter period. E.g., post 1970 and examine the B_0 and SSB trends.
Rationale: This was intended as a way to evaluate the sensitivity on unfished spawning biomass in the absence of early age composition data.
7) Examine run omitting length data
Rationale: The profile likelihoods seem to suggest that this data component differs substantially from most other components, there are a large number of size bins (with nearly no data) may be affecting the likelihoods, and when a sensitivity with "no age data" run was performed, the spawning biomass crashed after the removals from the 1970s.

Day 2 STAR Panel Requests

1. Selectivity sensitivity analyses starting with the working base model agreed to this morning:
    a. Change the age-based selectivity curve to an asymptotic pattern for the NW slope and AK slope surveys.
    b. Leave the two age-based slope survey curves asymptotic and allow the WCGBTS to be domed shaped.
    c. Allow all age-based surveys to be domed shaped.

Show model results as well as a comparison table and likelihoods across these alternatives. Split out the sea level index likelihood from the other surveys in these comparisons.

Rationale: When evaluating the length and age data from these surveys, these data were comparable among all surveys with some indication of proportionally older ages in the slope surveys.

2. Re-run request 1c above with the length data removed (except for lengths in the discards). Show model results and the likelihoods in the comparisons requested in #1. If time allows, do a small number of jitters for requests 1c and 2.

Rationale: The profile likelihoods seem to suggest that this data component differs substantially from most other components, there are a large number of size bins (with nearly no data) that may be affecting the likelihoods, and when a sensitivity with "no age data" run was performed, the spawning biomass crashed after the removals from the 1970s. It's possible that small errors in assumed constant growth curves affect length frequency predictions which may impact selectivity.

3. Provide a comparison of the working base model, with and without the sea level influence on recruitment, as well as with or without fishing (e.g., dynamic B_0 estimates). Provide model results, including comparison of recruitment and recruitment deviation estimates, and include a plot of the cumulative sum of the recruitment deviation vectors over time (not necessary for the dynamic B_0 runs). Also include a table of the changes in likelihood for the two runs with the sea level index specified in this comparison.

Rationale: To understand the influence of sea level on recruitment over time, and to explore whether the cumulative values of recruitment deviation estimates indicate regime-like behavior in productivity.

4. Plot the recruitment values and deviations from the working base model without the sea level index and compare to the recruitment values and deviations in the 2015 assessment.

Rationale: The 2015 relationship informed the sea level index used in the current working base model.

5. Provide two simple plots of growth estimates and mean lengths for ages 0 - 30 with factors being Regions (colors) for females and males with sex being on two panels.

Rationale: To understand differences in growth within and outside the assessment area.

6. Using the working base model, do $F_{45\%}$ projections for 2021 and beyond catch assuming:
   a) fixed gear catch only
   b) trawl gear catch only

   and display the relative catch values (or F_{SPR}s given equal catches) given a) and b) (no other model comparison needed).

Rationale: To understand that there may be some future variability in catch between actual gear types (irrespective of current "fishery" allocations), this will provide a baseline (extreme) range of the impact that will aid in future management considerations of future catch by gear scenarios.

Day 3 STAR Panel Requests

1. Attempt to get a model to converge with dome-shaped age-based selectivity for all surveys with the fixed gear fishery selectivity pattern estimated, if possible (for example, by constraining some parameters); otherwise, fixed at a reasonable pattern from a previous run. No sex-specific M for this run.

Rationale: The STAT explored a wide range of selectivity patterns and has not found an optimal model that converges. However, the STAT thinks additional effort towards this approach may lead to a base model. Additionally, there is evidence that trawl survey selectivities would not be logistic. Further, the prior distribution for sex-specific M did not suggest a difference in M and the data do not appear to be informative between the sexes.

2. Complete requests #2-4 from day 2 if the STAT is able to develop a better model that converges. If request #1 is not successful, fall back to yesterday's working base model with no sex-specific M. Compare the results with and without sex-specific M. Complete requests #2-4 from yesterday. Show model results and the likelihoods in the comparisons requested in request #1 on day 2. If time allows, do a small number of jitters for requests 1c and 2 from day 2.

Rationale: See the rationale from day 2 requests #2-4.

3. Given the challenges in getting a base model to converge, postpone the earlier day 3 requests. Starting with the day 2 "working base model" (with WCBTS age- based

selectivity logistic) and the day 3 model with all age-based selectivity curves dome-shaped, sequentially explore models that:

a. Fix M at the median of the prior for males and females
b. Use conditional age-at-length data from the WCBTS to estimate growth
c. Fix growth at a credible estimate

Show model results and the likelihoods in the comparisons requested in request #1 on day 2. If time allows, do a small number of jitters for requests 1c and 2 from day 2.

Rationale: There are challenges in developing a base model that has reliable convergence.

Day 4 STAR Panel Requests

1. Provide a run in which growth is estimated with CAAL data from the WCBTS, the length data are removed from all fleets except for WCBTS and the discards, and natural mortality is estimated as a single value for both sexes. Provide an additional run with the above changes in which the model begins in 1970 with an estimated initial F.

Rationale: Based on the results of the day 3 requests that were presented, there is tension in the age and length data influencing the growth curve. This may be a result of regional differences in growth that could interact with shifts in the distribution of fisheries effort, leading to greater tension in the model. The proposed base model is informed with age-based selectivities and the age data are thought to be the more important data to retain. Further, developing a model based on length data would require additional effort.

2. Fix the retention curve for the discard length data for the fixed gear and trawl fisheries at their estimated values from the working base model and remove the compositional data from the likelihood estimation. Provide a comparison plot and table of likelihood and key parameter results for the two models.

Rationale: These data are intended to estimate the retention curve rather than year class strength. As presently configured, the magnitude of the sample sizes from the discard lengths is substantial and may conflict with age composition data, which are more directly related to fishing mortality. This change should further simplify the model and reduce any remaining tension between length and age data.

3. With the working base model, try once more to estimate age-based selectivity for the fixed gear fishery. If a model that converges is found, provide a comparison plot and table of likelihood and key parameter results to this model relative to the models in the previous request.

Rationale: To investigate whether reduced tension between age and length data may facilitate the estimation of the fixed gear fishery selectivity curve.

4. Run a retrospective analysis.

Rationale: Earlier runs suggested an unexpectedly strong influence of recent length data from discards. A retrospective analysis will help confirm that the model is not overly sensitive to recent data.

5. Try to estimate an asymptotic age-based selectivity curve for the fixed gear fishery. Provide a comparison with the previous base model result, likelihood values and key parameter values.

Rationale: This fishery catches the largest fish among all fisheries and surveys. Earlier efforts to estimate asymptotic selectivity were not successful, but length data have been removed and were thought to be a major source of tension.

6. Do a run with the aging error turned off for ages 0-5. Provide a comparison with the previous base model result, likelihood values and key parameter values.

Rationale: To ensure that the aging error is not influencing the ability to fit the age data for recent strong year classes, as there is an indication of underfitting in the age composition data.

7. Drop the last three years of sea level data (2016-2018). Provide a comparison with the previous base model result, likelihood values and key parameter values.

Rationale: To ensure that these data are not drawing down the age and length composition data with respect to the strength of the 2016 year class.

8. Do likelihood profiles on the working base model, with any of the above changes that the STAT finds to be improvements, for ln(R0), M and steepness (in that rank priority).

Rationale: To ensure no surprises in the current working base model.

Day 5 STAR Panel Requests

1. Do a weighting sensitivity (Dirichlet multinomial, Francis, Harmonic Mean) and report the results.

Rationale: To ensure that the model is insensitive to data weighting.

2. Do a retrospective analysis of both the current working base and the single M sensitivity run. The STAT is free to report a subset of retrospective years (e.g., -2, -4).

Rationale: The previous retrospective analysis did indicate retrospective patterns.

3. Provide a first pass at a possible major axis of uncertainty for the decision table. Use the $\ln(R_0)$ point estimate that results in an ending spawning biomass consistent with the upper limit from the working base model, and the $\ln(R_0)$ associated with the ending spawning biomass from the lower 1.15 asymptotic confidence limit for the single M sensitivity model.

Rationale: We need an axis of uncertainty for the decision table.