CIE Independent Peer Review Report

on

Pacific Ocean Perch and Petrale Sole STAR Review

Prepared by

Yong Chen

Professor of Fisheries Population Dynamics
School of Marine Sciences
University of Maine
Orono, ME 04469

July 25, 2011

Contents

I. Executive Summary

The STAR review for the Pacific ocean perch (POP) and petrale sole (PS), held in Seattle, WA from June 20-24, 2011, was aimed to evaluate the data and model, make recommendations for improvement, develop base case stock assessment, and identify sensitivity analyses and alternative scenarios that need to be considered. Future research needs to improve assessment were also identified.

This stock assessment is the first full POP assessment since 2003 and it is the first time the assessment has used Stock Synthesis (SS), the stock assessment framework suggested to be used for groundfish stock assessment on the west coast. For PS, this is the first stock assessment since 2009, which updates to the most recent version of SS from an old version of SS used in 2009. The Northwest Fisheries Science Center (NWFSC) and Pacific Fishery Management Council (PFMC) provided all the necessary logistics support, documentation, data, and background information requested. The scientists involved in the process were open to suggestions and provided additional information and stock assessment runs upon request. The whole process was open and constructive and all materials were sent to me in a timely manner. As a CIE reviewer on the STAR panel, I am charged to evaluate PS and POP stock assessment with respect to the Terms of Reference and help develop STAR Panel report.

I would like to commend the great efforts of all the participants in the STAR review for providing necessary background information on PS and POP life history, fishery-dependent and fishery-independent monitoring programs, stock assessment history, and development of management regulations. I was impressed by the breadth of expertise and experience of the participants, the openness of discussion for considering alternative approaches/suggestions, and the constructive dialogs between the STAT team and STAR Panel throughout the review.

Overall, I believe the PS and POP stock assessments were done based on the best science available. The assessments appeared to be scientifically sound and adequately addressed management requirements. In particular, I would like to commend the PS and POP STAT teams for their efforts and openness in addressing uncertainty in the assessment and in exploring alternative model configurations. However, I believe some important questions still need to be addressed and there is still room for improving the PS and POP stock assessments.

My specific recommendations/comments include (1) conducting in-depth analysis to identify and quantify uncertainty for a given set of data BEFORE the data are inputted in the SS model; (2) combining PS fisheries data between WA and OR in the assessment; (3) estimating uncertainty associated with catch estimates to develop a plausible range of catch estimates, which can be used in the sensitivity analysis to evaluate impacts of uncertainty associated with catch estimates on stock assessment; (4) evaluating reconstructed historical landing data more thoroughly prior to STAR review; (5) involving the STAT teams in analyzing discard data for future PS and POP stock assessment and making processing of discard data more transparent; (6) analyzing potential temporal changes in length at age and weight at age data outside the SS model for both PS and POP; (7) evaluating effects of bin width in data aggregations on residual patterns; (8) conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate

potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs; (9) exploring robust likelihood functions to remove impacts of outliers which are likely to exist in input data used in the assessment, given that the data were derived from different sources and were subject to different errors; (10) conducting a simulation study to generate a simulated fishery based on the current stock assessment results and apply the SS to assess this simulated fishery with different process and observation errors to evaluate the performance of SS model in quantifying the PS and POP stock dynamics; (11) evaluating residual patterns of age/size composition data more thoroughly to help identify possible temporal trends in growth and selectivity; (12) exploring dynamic binning approach to reduce the impact of numerous size classes without data; (13) evaluating interactions of h and M and uncertainty resulting from their high correlations between these two parameters in the SS; (14) developing competitive models of different complexities to compare with the SS; (15) evaluating current virgin biomass-based harvest control rule which depends on often-poorly-estimated virgin stock biomass because of large errors in historical data; (16) keeping assessment model structure relatively stable over time; (17) evaluating the performance of the projection done in the past assessment, retrospectively, to evaluate their performance in achieving management objectives; (18) conducting a transboundary stock assessment for PS and POP, given the evidence of interactions between the PS and POP stocks in the US and Canada; and (19) conducting habitat suitability modeling to identify suitable habitats for PS and POP, outline potential habitat maps, and help improve survey design.

Detailed comments and recommendations with respect to TORs can be found in Sections IV and V of this report.

II. Background

Petrale Sole (*Eopsetta jordani*)

Petrale sole (PS), a right-eyed flounder, is distributed from the western Gulf of Alaska to northern Baja California (Love et al. 2005). It prefers soft substrates of sandy and muddy bottoms and can be found in the depth range of 0-550 meters. Limited information is available with respect to the PS stock structure. There is no information suggesting that there are separate stocks within the US waters. Previous studies suggest that juveniles are subject to limited movement, but adults tend to move inshore and northward onto the continental shelf to feeding grounds in spring and summer, and offshore and southward in fall and winter to deep water spawning grounds (Love 1996). Corresponding to such spatial and temporal changes in the PS distribution, the PS fishery also has a strong seasonality. There are two fishing seasons off the US west coast: summer fishery catching majority of PS in the depth of 70-220 meters from March through October, and winter fishery targeting spawning aggregations in depth of 290-440 meters during November through February.

PS spawn in winter in deep water (270-460 m) off the US west coast from November to April. PS eggs and larvae are planktonic. The larval duration including egg stages spans about 6 months with larvae settling at about 2.2 cm in length on the inner continental shelf (Pearcy 1977). Juvenile PS are benthic and carnivorous foraging on worms, clams, brittle star and other juvenile flatfish (Casillas et al. 1998) and usually found on sandy or sand/mud bottom. Adult PS have more diverse diets and become more piscivourous at larger sizes (Allen et al. 2006) and often found on soft bottom. Different growth rates are found between females and males. The maximum report length is 70cm (Love et al. 2005) and maximum observed age determined using break and burn method is 31 years (Haltuch et al. 2011).

Ecological factors have not been explicitly modeled in the assessment although the PS population dynamics are clearly affected by these factors. For example, the California current, the Pacific Decadal Oscillation, and offshore Ekman transportation of eggs and larvae can all affect the PS abundance and distribution.

PS have been caught in the flatfish fishery off the US Pacific coast since the late 19th Century. The fishery was developed off California water prior to 1876. Although the flatfish fishery in California was well developed by the 1950s PS catch statistics were not reported until 1970. In the early CA catch statistics report, PS landings during 1916-1930 were not separated from the total flatfish landings. Trawling for PS off Oregon began early, but the fishery was not established until 1937 and fishing grounds moved to deeper water over the time. Catch off Washington was small until the late 1930s. The PS stock showed sign of depletion in the 1950s (Harry 1956). By the 1980s, winter catch became higher than summer catch. Trawl fleets dominate the PS fishery. Reconstructed catch history shows that catch was high in the 1940s and 1950s, but generally declined until the mid-2000s. The stock was defined as "overfished" in 2009 and catch in 2010 was limited to 701 mt.

The 2005 PS assessment assumed two stocks, northern (US Vancouver and Columbia INPFC areas) and southern (Eureka, Monterey and Conception INPFC areas) stocks. Because of strong

evidence of mixing stock from tagging studies, lack of study on stock structure, limited evidence for differences in growth, and difficulty in separating data in different states, the 2009 stock assessment combined the two previously separated areas to yield a coastwide stock assessment. The 2011 stock assessment also considers the coastwide PS as a unit stock using data through 2010. Various management regulations to limit fishing efforts including area closures have been implemented since 1998 and have altered the PS fishery substantially. This is especially true for the PS summer fishery. Fishery-dependent information collected from onboard observer and port sampling programs includes catch, discards, and their size compositions.

In addition to the PS fishery in the US, it also needs to note that there is a PS fishery in Canada. The PS stock in Canada is considered to mix with the US stock. However, no transboundary stock assessment has been done.

The data used in this 2011 stock assessment are: (1) fishery-dependent data including commercial landings from 1876 to 2010, length composition, mean weight and total biomass of discarded catch in the fishery from the West Coast Groundfish Observer Program (WCGOP) and the study by Pikitch et al. (1988), and summer and winter fishery CPUEs from 1987-2009 for WA, OR, and CA; (2) fishery-independent data including abundance indices, age/length data from the NWFSC bottom trawl survey from 2003-2010 and the AFSC bottom triennial bottom trawl survey from 1980 to 2004; and (3) biological information including fecundity, maturity, length-weight relationships and ageing errors of different sources.

The model used for this assessment is Stock Synthesis (3.21d) (SS). This new version of model allows the timing of the catches to be specified for different fleets. Thus, the 2011 model is a 12 month model with removals from fishery catch assigned to appropriate season. The iterative re-weighting was done manually to make the input weighting factors (Coefficient of variation CV and effective sample sizes) consistent with the outputs. The STAT team presented the base case and some sensitivity analyses at the STAR review. Model configuration and parameterization options were extensively discussed at the STAR Panel and STAT members, which resulted in more runs and changes in model configuration and parameterization for the base case and sensitivity analyses (see Appendix for the list of requests by the STAR Panel for further analyses in each review day).

Pacific Ocean Perch (*Sebastes alutus*)

Pacific ocean perch (POP) is widely distributed off the US and Canadian west coasts. They are most abundant in the Gulf of Alaska, and sparse south of Oregon. The POP was mainly harvested by Canadian and US fishing fleets prior to 1966. Landings were very low prior to 1940, but reached 1,000 mt in 1951, and over 7,000 mt in 1965. With the introduction of large distant-water fishing fleets from the former USSR and Japan after 1965, catch by foreign fishing fleet increased dramatically, reaching over 15,000 mt in 1966 (Rogers 2003). Catches declined quickly after the 1960s and averaged 1,500 mt over the period of 1977-1994 with the POP stock being considered to be severely depleted by 1969. Landings continued to decrease since 1994 because of restricted regulations (Gunderson 1978).

In the US water, POP stock was managed by relevant state government prior to 1977. With implementation of the Magnuson Fishery Conservation and Management Act in 1977, management of groundfish, including POP, was the responsibility of Pacific Fishery Management Council and National Marine Fisheries Service (NMFS). The PFMC adopted a management strategy in 1981 to rebuild the depleted POP stock to B_{MSY} within 20 years. Various management regulations were introduced including Acceptable Biological Catch (ABC), landing limits and trip limits. However, the stock biomass is still low.

Although there is evidence of mixing of stocks between the US and Canada, no transboundary stock assessment has been done. Limited information is available for stock structure. Separate stock assessments were conducted in the USA and Canada.

Previous stock assessment tended to over-estimate natural mortality and productivity because of under-ageing POP using the surface ageing of otoliths which yielded the maximum age of around 20s. However, using break and burn methods, maximum age was considered over 100 years.

Several research survey programs were used to yield fishery-independent information on POP abundance, distribution and biological parameters. A coastwide rockfish survey was conducted in 1977 and was repeated every three years through 2004 (Gunderson and Sample 1980). A POP survey off the WA and OR coast was conducted in March-May 1979 and 1985. The NWFSC annual bottom trawl survey began in 1998 initially targeted Dover, thornyhead and sablefish, and expanded to other groundfish in 1999, and further expanded its spatial coverage to include the shelf in 2003.

The data used in this 2011 stock assessment are: (1) fishery-dependent data including commercial landings from 1940 to 2010, CPUE, fishery age and length composition, mean weight and total biomass of discarded catch and discard length composition in the fishery from the West Coast Groundfish Observer Program (WCGOP) and the study by Pikitch et al. (1988); (2) fishery-independent data including abundance indices, age/length data from the NWFSC bottom trawl survey from 2003-2010, the NWFSC slope survey from 1999-2002, the POP survey in 1979 and 1985, NMFS triennial shelf survey conducted every third year from 1977-2004, the AFSC slope survey for 1996-97 and 1999-2001; and (3) biological information including fecundity, maturity, length-weight relationships and ageing errors of different sources. Because of differences in survey timing, the NMFS triennial shelf survey series were split into two time periods 1980-1992 and 1995-2004.

The model used to conduct this assessment is SS (3.21d). This new version of model allows the timing of the catches to be specified for different fleets. An iterative re-weighting approach was used to make the input weighting factors (CV and effective sample sizes) consistent with the outputs. The STAT team presented the base case and some sensitivity analyses at the STAR review. Model configuration and parameterization options were extensively discussed among the STAR Panel and STAT team, which resulted in more runs and changes in model configuration and parameterization for the base case and sensitivity analyses (see Appendix IV for the list of requests made by the STAR Panel during the review).

The current model is different from the 2003 POP assessment model and represents the first POP assessment done in the SS framework. Major differences between this model and 2003 stock assessment model include (1) the current model is sex specific; (2) growth is estimated internally; (3) selectivity is length-dependent and can be quantified with double normal or logistic curves; (4) a time-varying retention curve to model discards; (5) catch series was extended to 1940, rather than 1956 in the previous assessment; and (6) survey indices were standardized using GLMM, rather than area swept estimates. A beta prior was also developed for steepness parameter based on a meta-analysis of west coast groundfish species

As one of the two CIE reviewers who were the members of the STAR Panel, I am charged to evaluate PS and POP stock assessment with respect to the Terms of Reference and help develop STAR Panel report. This report is my independent review and includes an executive summary (Section I), a background introduction (Section II), a description of my role in the review activities (Section III), my comments on each item listed in the Terms of Reference (ToRs, Section IV), a summary of my comments and recommendations (Section V), and references (Section VI). The final part of this report (Section VII) includes a collection of appendices including the Statement of Work (SoW).

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

My role as a CIE independent reviewer was to actively engage in STAR Panel review, make contributions to the development of STAR Panel report, and conduct an impartial and independent peer review of the PS and POP stock assessments with respect to the pre-defined Terms of Reference for the CIE reviewers.

Two weeks prior to the review in Hotel Deca in Seattle, WA, I received the draft 2011 PS and POP stock assessment reports and relevant background information including previous stock assessment reports for two fish stocks, previous STAR Panel reviews, and comments from the PFMC Scientific and Statistical Council (SSC). I also received relevant background documentations, instructions for SS3, an executable SS3 program, and a technical report about SS3 model structure prior to the review meeting.

I read the two draft stock assessment reports and all other relevant documents that were sent to me (see the list in the Appendix I). Also, I collected and read references relevant to the topics covered in the reports and the SoW prior to my trip to the STAR Panel review.

The STAR Panel review was held from June 20-June 24 in the Hotel Deca in Seattle, WA (see Appendix II for the schedule). The review was attended by the STAR Panel Technical reviewers and advisors, STAT team, NWFSC scientists, PFMC managers, and publics (see the List of Participants in Appendix III).

Presentations were given during the first two days of the review to provide the STAR Panel with background information on the fishery-dependent data, fishery-independent data, biological data, model, initial model configuration, and model runs for PS and POP (see the list of presentations in Appendix I). As a STAR Panel reviewer, I was actively involved in the discussion during the

presentation by (1) questioning and asking for clarification on monitoring/sampling program design, data collection methods, statistical analysis, and interpretations; (2) making observations of the process; (3) making comments and suggestions for alternative approaches and more analyses; and (4) developing a list of daily requests for more model runs and more background information to evaluate roles of data, model parameterization, and model assumptions and identify potential discrepancies in stock assessment modeling for PS and POP. I had also been interacting with relevant scientists who presented the talks and asked for further clarifications and references during the reviews. Different model configurations were developed at the end of each day in the review. More model runs were conducted each night based on the newly configured models, and results were presented and discussed during the following day. Such an interactive discussion between the STAT teams and STAR Panel led to the development of base case run and sensitivity model run for both PS and POP at the end of the review week. Detailed requests for more model runs, their justification, and responses from the STAT team can be found in the STAR Panel reports for PS and POP.

I was actively involved in developing test run scenarios, discussing outputs and their implications, and identifying issues related to test runs. I also discussed relevant issues with the fellow STAR reviewers.

IV. Summary of Findings

My detailed comments on each item of the ToRs are provided under their respective subtitles from the ToRs (see below), separately for the PS and POP.

IV-1. Petrale sole

IV-1-1. Become familiar with the draft stock assessment and background materials.

Two weeks prior to the review, I received the draft PS stock assessment report and relevant background materials as scheduled. The background materials include previous PS stock assessment reports and relevant comments from the STAR Panel and SSC. I read the draft report and the background materials, took the notes of the key issues in the previous and current stock assessment, evaluated major differences between the current and previous assessment reports, identified potential issues, and drafted a list of potential questions/concerns I would like to raise at the STAR Panel review. I also identified potential model runs we should consider at the review.

IV-1-2. Comment on the quality of data used in the assessments including data collection and processing.

A large set of fishery-dependent and fishery-independent data were compiled for the 2011 PS stock assessment from the beginning of the fishery through 2010. The fishery-independent data include abundance indices and age/length data from the annual NWFSC bottom trawl survey from 2003 to 2010 and the AFSC triennial bottom trawl survey from 1977 to 2004 with 1977 data being excluded because of incomplete coverage of PS depth range. Fishery-dependent data

available to the assessment include commercial landings from 1876 to 2010 in WA, OR, and CA (the OR historical data were reconstructed after the last STAR review), length composition data, mean weight and total biomass of discarded catch in the fishery from the West Coast Groundfish Observer Program (WCGOP) and the study by Pikitch et al. (1988). Summer and winter fishery CPUEs were also derived from 1987-2009 for WA, OR, and CA. Other information available to the assessment includes fecundity, maturity, length-weight relationships and ageing errors of different sources. A meta-analysis was also conducted to identify priors for natural mortality and steepness parameters. I commend the STAT team and relevant agencies for compiling such a comprehensive data set and for conducting extensive analyses to standardize data (e.g., fishery CPUEs and survey catch rates) and biologically justify the priors for the two of the most model parameters. Overall, I believe this set of the data available to the 2011 stock assessment represent the best data available for the PS stock assessment. However, I believe more studies can be done in the future to improve the input data.

Comprehensive reconstruction of catch data history has not been done in WA. Current reconstructed landing in OR may include portion of data double counted in OR and WA prior to 1996 (mainly in the 1960s and 70s). They might result from catch in WA but landed in OR. Reconstructed data were based on area, not fishing ports, and for annual, not seasonal, and they might mix data from the two states. The seasonal partitions were based on 2009 data, which might not be representative of previous fishing years. Thus, the OR reconstructed landing data do not represent the best data available, and was suggested by the STAT Team not to be included in the base case assessment. There were large spatial and temporal variabilities in fishing intensity and sampling efforts, implying that large variability in the quality of landing data during different time periods and different fishing ports. This suggests that the data reconstruction should have been done on a finer spatial and temporal scale to capture such variability. I did not see any measure for quantifying uncertainty associated with the estimate of landings. An analysis should be done to more thoroughly evaluate the quality of reconstructed data. Given the importance of the catch data in the assessment (assumed to be error-free in the current stock assessment), I suggest conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs. A study was done in 2003 to evaluate and analyze field sampling in North Pacific groundfish fisheries (MRAG 2003). Although the work was mainly focused on evaluating biological sampling protocols (MRAG 2003), a similar study can be done for evaluating quality of catch estimates. I suggest estimating uncertainty associated with catch estimates to develop a plausible range of catch estimates, which can be used in the sensitivity analysis to evaluate impacts of uncertainty associated with catch estimates on stock assessment.

Given the strong seasonality of fisheries, I believe the current partition of fisheries landing and catch size composition by season is necessary and reasonable. However, more study is needed to evaluate annual variability in quality of fisheries catch size composition data. Size composition data for fisheries catch were derived from various sources and likely subject to various errors. In-depth analyses should be conducted to evaluate whether the quality of size composition data for fisheries catch vary with year, season and gear.

Discard data in 1986 and 1987 were obtained from Pikitch et al (1988), which covered the OR summer fishery. More recent discard data were collected from the WCGOP on-board observer program. Uncertainty estimates for discards appeared not right, and bootstrapping should be done consistently with how the observer program is designed to reflect the true variability of the process. Limited information was available about how the discard data were derived. Without a full understanding of the observer program that collects discard data, it is difficult to evaluate the quality of discard data. Currently, the PS STAT team was not involved in the derivation of discard data. I *STRONGLY* suggest that the STAT members be involved in analyzing and developing discard data for future PS stock assessment. Without the STAT team's involvement, it is difficult for them to evaluate the quality of discard data and its potential impacts on the stock assessment.

The AFSC contracted MRAG Americas, Inc. to conduct a study to evaluate biological sampling protocol in North Pacific groundfish fisheries (MRAG 2003). The empirical and computer simulation approaches developed in MRAG (2003) can be used to quantify variability associated with size composition data for PS.

Although not explicitly stated in the review, I believe that the current fishery-dependent sampling program has some overlaps in catch reporting from different sources. Thus, data from different sources can be compared and cross-validated. Such a study can yield some insights about potential errors in fisheries data from different sources.

Fishery CPUE data were derived for both summer and winter fisheries. The ability of summer fishery CPUE in describing stock biomass might be compromised by changes in regulations which greatly affected the fishermen's behavior and subsequent catchability. On the other hand, fishermen's behavior in the winter fishery was less affected by changes in regulations (comments of John Devore, Pacific Fishery Management Council). However, concerns were raised about the winter CPUE, because it targeted spawning aggregations and might have issues of hyperstability (i.e., CPUE tends to be stable even when stock biomass changes). The relationship between winter CPUE and stock biomass should be nonlinear. Large temporal contrast of the winter fishery CPUE was surprising given that the winter fishery targeted spawning aggregations. It appeared that the temporal trend of this CPUE was consistent with temporal trends of the two fishery-independent summer survey indices. Given lack of abundance index in winter season, this CPUE was valuable and should be used in the stock assessment. However, relevant hypotheses should be developed and evaluated to identify biological reasons in the behavior of temporal variability of this CPUE.

I commend the effort by the STAT team to standardize CPUEs. However, I observed that q-q plots showed lack of fit when CPUE was low and there was no measure presented to describe how well the model fitted the data in the CPUE standardization. More modeling diagnosis may be needed to improve the CPUE standardization.

Data derived from the two fishery-independent survey programs were incorporated in the stock assessment. These two survey programs differ in their survey design (systematic versus stratified random designs) and spatial coverage. Both of them are conducted in the summer season. Although together these two programs covered the PS stock area pretty well, lack of coverage of

shallow waters by the survey programs might miss some small fish, which might introduce extra uncertainty in estimating recruitment. Standardizing survey abundance indices using GLMM is likely to improve the quality of survey abundance index.

Age-at-length data were only available for the NWFSC survey. Ageing results were prone to ageing errors, in particular for old fish, in early time series because of use of the surface ageing method for some samples, rather than the burn-and-break method. This problem was recognized and corrected prior to this stock assessment. I believe that ageing data were well scrutinized prior to their use in this stock assessment.

I would like to commend the STAT team in their efforts to address data quality issues (e.g., ageing errors, historical landing, CPUE and survey abundance index standardization) raised in the last STAR Panel review (their detailed replies can be found in the 2011 PS stock assessment report).

IV-I-3. Evaluate and comment on analytic methodologies.

The base model developed for the 2011 PS stock assessment includes the following features: coast-wide model with seasonal fleet structure, sex specific, asymptotic selectivity, blocks on selectivity, internally estimated growth, informative priors on M with male M estimated as an offset from female, priors for steepness estimated with Meyer's priors, effective sample sizes tuned, and use of the most recent SS version SSv3.21e. CPUE and survey abundance index data were standardized using GLMM before being used in the assessment to remove effects of factors that may influence catch efficiency in the surveys and fisheries.

Growth is assumed to be time-invariate in the assessment. However, large changes in population over the time period covered by the stock assessment may result in changes in maturity and growth. It is, thus, desirable to evaluate temporal variability in growth before making the time-invariate growth assumption. It would be useful to fit length-at-age data of each year class to von Bertalanffy growth model outside the SS to evaluate if growth varied over year classes. Given changes in environment gradients in the area covered by the stock assessment, it is also desirable to evaluate if growth varied substantially among WA, OR and CA outside the SS.

I also believe choices of selectivity functions for the survey and fishery should be further evaluated and justified. This might be done through model diagnosis of size/age composition fitting.

Overall, I believe the analytic methods used in the analyses of data prior to their inclusions in the SS are sound and that the SS provides very flexible stock assessment platform to accommodate needs for the PS stock assessment. However, I noticed that most model diagnoses done were rather qualitative (often eyeballed). Even if likelihood function values were provided, they only provided relative performance among alternative model configurations. This makes it hard to evaluate the model performance. The complex model structure with data of different processes and sources is likely to lead to possible high correlations between different model components which quantify different life history and fishery processes. Thus, more efforts should be focused on the evaluation of the model performance and quality of parameter estimates. I suggest conducting a simulation study to generate a simulated PS fishery based on the current stock

assessment results and apply the SS to assess this simulated fishery with different process and observation errors. Such an exercise can be done in conjunction with efforts to develop management strategy evaluation (MSE) framework for the PS. It can provide insights about the model performance and identify key sources of uncertainty in modeling. Following the previous STAR Panel, I would suggest that a simpler age-structured model be developed to evaluate possible differences in stock assessment resulting from different stock assessment models.

IV-I-4. Evaluate model assumptions, estimates, and major sources of uncertainty and provide constructive suggestions for improvements if technical deficiencies or additional major sources of uncertainty are identified.

One of the most important assumptions implied in the PS assessment is that landing is free of errors. This assumption was certainly violated in modeling because of the quality of catch data. I did not see that a systematic approach was developed to evaluate impacts of violating this assumption. Two approaches may be possible to evaluate possible consequences of violating this assumption. One is to assume catch is subject to errors and develop a likelihood function for predicted and observed catches and give a small weight to this likelihood function to see the level of catch predicted by the rest of the data in the assessment. The other approach is to quantify the uncertainty associated with catch and develop a probable distribution of catch and conduct MCMC runs with each run randomly drawing catch from the defined distribution. This approach can directly incorporate uncertainty in catch in the stock assessment.

Given the SS model and input data used, I believe that one of the most important model configurations is to determine weighting factors for different components of the input data. This is reflected by effective sample size for size/age composition data and variance values for abundance indices. Some iterative runs were conducted in current stock assessment to iteratively adjust variances for some abundance indices, but this is not done for size composition data (although the predicted and observed effective sample sizes were evaluated). More analysis can be done to evaluate impacts of weighting factors including giving very large or very small weights to some data series to evaluate their impacts on the quantification of uncertainty in stock assessment.

A lot of bins in small and large sizes have zero (0) observations. They might affect model fitting. Dynamic binning can be used to reduce the impact. Limited runs were done during the review, but more work should be done to fully understand this modeling process.

Time block for fishery selectivity was developed based on the analysis from the last assessment and was consistent with changes in regulations. This approach seems to be reasonable.

Discard fraction data were estimated, but with inflated variance (std = 0.3) because the mean estimates were close to zero (0) which was considered biologically unrealistic. As I stated in section IV-1-2, the STAT team needs to be involved in the estimation of discard data used for the stock assessment and the estimation process needs to be more transparent. The STAT team and STAR Panel did not have enough information to evaluate validity of discard estimates with respect to assumptions implicitly and explicitly included in the discard estimation.

Functional S-R relationship was assumed in the assessment. Both Ricker and Beverton-Holt models were used in previous PS stock assessments. Although we did request a model run with steepness h being set at a very large value (close to 1) and recruitment variance being given a large value (thus essentially no functional relationship), I believe more work is needed. I suggest estimating SSB and recruitment freely in the model and then evaluate the relationship between SSB and recruitment. Alternatively, I suggest developing indices for recruitment and SSB from the survey data and evaluating their relationship to justify the functional relationship assumed in the model. If there is no relationship, such a functional relationship should not be assumed.

Retrospective analysis was done in the assessment. However, we only had limited discussion about retrospective errors in the review. No quantification of retrospective errors was done. Although it seems that the direction of retrospective errors might vary, there existed retrospective errors in a given year, resulting in extra uncertainty currently not considered in the stock assessment. Retrospective errors associated with recruitment were not discussed, which might have large impacts on stock projection. I suggest more study be done to improve understanding of retrospective errors associated with key stock parameters such as SSB, F and recruitment.

Canadian fishery landing is equivalent to the WA fishery in scale. Recent stock assessment suggests that biomass was low, which is consistent with the stock projection made in the US PS stock assessment. Both countries projected similar recruitment dynamics. However, differences existed between the two fisheries. For example, Canada has few young fish in fishery data, while US (WA) has many more young fish in landings; Canadian catch (females) was about 10 cm larger than WA (more fish larger than 60 cm); more male fish in Canada were larger than 50 cm, few in the US (WA) were larger than 50 cm; more large fish were caught in Canadian fishery-independent survey while US (WA) capture smaller fish in survey. Previous stock assessment attempted to (1) include Canadian landings in the WA landings in the assessment; and (2) treat Canadian fishery as a separate fishing fleet. Given lack of evidence of two separate stocks in the Canadian and US water, efforts should be made for conducting a transboundary stock assessment.

Growth was assumed to be constant over time in the assessment model. However, given large changes in stock size and long time period covered in the assessment (thus, large temporal changes in environmental conditions), it is very likely that the PS growth patterns changed over time. Thus, potential temporal changes in length at age and weight at age need to be evaluated. I believe such an evaluation should be done outside the SS. For example, length-at-age data collected in different years can be fitted to von Bertalanffy growth function and a comparison can be done to evaluate differences in growth models among years. If there is evidence of temporal variability in growth, such variability should be included in the SS.

The inclusion of the winter fishery CPUE led to a more optimistic conclusion about stock status. The winter fishery, which targets spawning aggregation, might have an issue of hyperstability. However, fitting of a power function suggests that the relationship between stock biomass and survey abundance index is of hyperdepletion. This issue should be evaluated more thoroughly in the future.

Given large noises associated with data, some robust likelihood functions (e.g., Chen et al. 2003) should be considered in modeling to remove impacts of potential outliers in data.

IV-I-5. Determine whether the science reviewed is considered to be the best scientific information available.

The PS STAT team had made a lot of improvements since the last stock assessment in data quality, quantity, and model configurations. Although there is room for more improvement, given the constraints, I consider the science reviewed in this STAR Panel review to be the best scientific information available.

IV-I-6. Provide specific suggestions for future improvement in any relevant aspects of data collection and treatment, modeling approaches and technical issues.

Given the evidence of interactions between the PS stocks in the US and Canada, a transboundary stock assessment would be ideal. If this is impossible in the short term, informal exchanges of stock assessment information between the two countries may be a good first step to move towards a transboundary stock assessment.

Dome shaped selectivity function should be evaluated for different time blocks.

Recruitment tends to be over-estimated in retrospective analysis in most of years tested, but also under-estimated in some years. This may indicate that the retrospective errors are perhaps more like uncertainty than biased errors. More studies need to be done to evaluate nature and magnitude of retrospective errors.

Given the importance of the catch data in the assessment (assumed to be error free in the current stock assessment), I suggest conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs. I suggest estimating uncertainty associated with catch estimates to develop a plausible range of catch estimates, which can be used in the sensitivity analysis to evaluate impacts of uncertainty associated with catch estimates on stock assessment.

I *STRONGLY* suggest that the STAT members be involved in analyzing and developing discard data for future PS stock assessment. Without the STAT team's involvement, it is difficult for them to evaluate the quality of discard data and its potential impacts on the stock assessment.

Potential temporal changes in length at age and weight at age need to be evaluated. I believe such an evaluation should be done outside the SS.

More analysis can be done to evaluate impacts of weighting factors including giving very large or very small weights to some data series to evaluate their impacts on the quantification of uncertainty in stock assessment.

I suggest conducting a simulation study to generate a simulated PS fishery based on the current stock assessment results and apply the SS to assess this simulated fishery with different process and observation errors. Such an excise can be done in conjunction with efforts to develop MSE framework for the PS. I also suggest that a simpler age-structured model be developed to evaluate possible differences in stock assessment resulting from different stock assessment models.

Because of the difficulty in separating landing data between WA and OR (in particular historical information) and similarity between the two fisheries, I suggest combining the WA and OR PS fisheries data in the assessment.

The virgin biomass estimate might be subject to large errors because of lack of information on stock abundance and unreliable estimates of catch data in early years. Current stock biomass estimate might be subject to large uncertainty because of retrospective errors. Current virgin biomass-based harvest control rule may vary from year to year with changes in the estimates of virgin and current biomass. Large uncertainty associated with virgin and current stock biomasses makes such a control rule undesirable. I suggest MSE be developed to overcome the problem.

There is a need to evaluate effects of bin width in data aggregations on residual patterns. Dynamic binning and robust multinomial likelihood functions should also be evaluated in reducing impacts of large number of bins with zero observation on model fitting.

There is a need to evaluate possible temporal variability in growth. Because the growth process cannot be estimated independently from other processes in the model, such an evaluation should be done outside the SS model. Alternatively, time-varying growth may be considered in the stock assessment model if there are enough data.

Residual patterns of age/size composition data should be more thoroughly evaluated to help identify possible temporal trends in growth and selectivity.

There is a need to conduct habitat suitability modeling to identify suitable habitats for PS and POP, to outline potential habitat maps, and to help improve survey design (Chang et al. 2010).

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

The PS STAT Teams presented the STAR Panel with an initial compilation of input data, model configuration, and modeling results during the first day of the review. The STAR Panel asked questions about the quality and quantity of input data, model configuration and relevant justifications, model assumptions, and modeling results. At the end of the discussion, a list of requests was made to the STAT team for more information/clarification/analysis, alternative model configurations/parameterizations, and new model runs. The STAT team worked overnight and came back the next day to present their replies, followed by more requests for extra analyses. This iterative process was repeated from Monday (June 20) to Friday (June 24),

resulting in the development of the base case scenario and extra scenarios for sensitivity analysis for the 2011 PS stock assessment.

Reconstructed data should be reviewed prior to stock assessment because the STAR Panel did not have time and lacked the ability to review the data quality and the STAT team was not involved in and lack of knowledge about how some of the data were collected, processed and analyzed. Lack of transparency on how discard data were derived is also an issue that should be resolved prior to stock assessment.

I commend the STAT team for their hard work to address concerns raised by the STAR Panel in a timely manner during the review week. The whole process was open and rather efficient, and the discussion was very constructive. However, because of large amount of background information the STAR Panel needed to go through and the complexity of stock assessment model with a large number of input data and model configuration options, I feel we did not have enough time to cover all the alternative model runs we would like to do for the PS stock during the review week. Not all the information (e.g., details on the estimation of discards and their associated uncertainty) was available to the STAT Team. Because of the tight schedule during the review week, I felt that I did not have enough time to carefully read all the materials distributed during the review week.

IV-2. Pacific Ocean Perch

IV-2-1. Become familiar with the draft stock assessment and background materials.

Two weeks prior to the review, I received the draft POP stock assessment report (Hamel and Ono 2011) and relevant background materials as scheduled. The background materials include previous POP stock assessment reports and relevant comments from previous STAR Panel and SSC. I read the draft report and the background materials, took the notes of the key issues in the previous and current stock assessment, evaluated major differences between current and previous assessment reports, identified potential issues, and drafted a list of potential questions/concerns I would like to raise at the STAR Panel review. I also identified potential model runs we should consider in the review.

IV-2-2. Comment on the quality of data used in the assessments including data collection and processing.

The input data for the POP stock assessment include some key life history parameters, ageing error estimates, priors for natural mortality M and steepness parameter h, landing, discards, survey abundance indices, and size composition and conditional age at length data. Quality of these data varied.

The life history parameters were estimated outside of the SS model. Overall, their quality was pretty good. However, temporal variability of these parameters was not evaluated. For example, age at maturity was assumed to be constant for the period covered by the stock assessment. Given the removal of large proportion of the population in the 1950s and 60s in the fishery, it is unlikely that age at maturity was constant. For fecundity-at-weight estimates, it was not sure if

the weight used was somatic weight or whole body weight. Instead of using SSB, the STAT team used egg productions (spawning outputs) to replace SSB in the SS, which reflected the importance of size structure of the population. This is good, but it is important to make sure to be consistent with SSB used in stock-recruitment modeling (priors for steepness h were derived from SSB).

Ageing errors were estimated based on the results of a double-read analysis of 1,161 POP otoliths by the Cooperative Ageing Project at the Newport Lab of NWFSC. Because surface ageing method tends to under-estimate fish aged over 12-15 years, break-and-burn method was used. Standard deviation of observed ages was assumed to increase linearly with the age. This approach to estimating ageing errors implicitly assumed that ageing results of the reference reader was unbiased. This cannot be validated. The assumption of linearly increased standard deviation of observed ages with age also needs to be more carefully evaluated.

Priors were developed for M and h based on meta-analyses of data reported in the relevant studies. Log-normal priors were assumed for M based on data from Hoenig (1981) and McCoy and Gillooly (2008). Beta priors were assumed for the steepness h based on a meta-analysis of the west coast rockfish assessments. Different distributions were weighted and combined to form empirical distribution which was then fitted to log-normal or beta density function to derive a log-normal function for priors. Although such an approach can justify the choice of priors for M and h, the data compiled from these previous studies for the meta-analyses are questionable in their quality, which could compromise the quality of priors derived for M and h. Parameters M and h played a critical role in driving the POP population dynamics. Model runs during the review suggested that uncertainty associated with them tended to have great impacts on the results.

Catch was negligible prior to the 1940s, but increased quickly from under 300 mt in 1948 to over 2,000 mt in 1952, and peaked in the mid 1960s (including large foreign catch from 1967 – 1976). Catch was reduced with increased regulations since the 1980s. Catch was revised for OR landing since the last assessment, leading to higher landings in early years. However, the revised catch was not thoroughly reviewed, raising questions of its quality. Large foreign landings were reported by foreign fishing fleets and there was no quality control on this part of the landing data. Because the foreign landings were so high, this uncertainty may influence the results. The Canadian fishery was not included in the assessment, which might introduce some extra uncertainty. It is likely that large errors were associated with landing data. Because of implicit assumption of landing data being error free in the SS, impacts of errors in landing should be carefully evaluated and understood.

Discard data were estimated as 16% for the 1980s (Pikitch et al. 1988), nearly 16% from 2002-2007 (WCGOP data), and 36% and 50% respectively for 2008 and 2009. Discard rate was assumed to be 36% in 2010. No uncertainty was estimated for these estimates, and the quality of these estimates was unclear. The STAT team needs to be involved in deriving these estimates to better understand the quality of discard data.

CPUE data were available from the domestic POP fishery from 1956-1973 (Gunderson 1977). However, the quality of these data was likely to be low. The CV associated with this CPUE data

set was doubled to reduce its weights in the SS modeling. Stock abundance indices were derived from the following fishery-independent survey programs: POP survey (1973, 1985), triennial shelf survey (1980-1992, 1995-2004), AFSC slope survey (1996-97, 1999-2001), NWFSC slope survey (1999-2002), and NWFSC slope/shelf survey (2003-2010). Spatial coverage of these programs differed. Little information on the stock abundance was available for the time period prior to the big removal by foreign fleets, making the estimation of early stock biomass subject to large uncertainty. Because of differences among programs in their design, objectives, and spatial and temporal coverage, there might be large differences in their quality with respect to their precision and accuracy in quantifying stock dynamics.

Length/age composition data were derived from trips (fishery) or tows (survey). However, limited information was available to evaluate the quality of the data. The empirical and computer simulation approaches developed in MRAG (2003) can be used to quantify uncertainty associated with size composition data for POP (see detailed discussion in Section IV-1-2).

Conditional age-at-length composition data were only available for those surveys with length data. Each sample was considered as a random sample of age within a size bin. The size bin width was 1 cm. It was suggested in the review that the 2004 age at length data for NWFSC shelf/slope were subject to large errors because sample for 2004 ageing data collected from one vessel was messed up and the data from that vessel were discarded.

IV-2-3. Evaluate and comment on analytic methodologies.

This stock assessment is the first time of using the SS model for the POP. For continuity, a bridge analysis was conducted. An SS model was developed to include all the input data used in the old model with the addition of recent data and other extra information required in the SS in the 2011 base case model. Differences between current and old models were then compared for discrepancy between the two models in quantifying the POP stock dynamics. I consider this analysis very important in the crossover of the old model to the SS. Transition from the simpler older model to SS resulted in similar conclusion about stock historical trend. However, the depletion estimates in recent years had large differences. The STAT team and STAR Panel spent some substantial time to discuss the discrepancy between the two models, but more studies need to be done to identify reasons and understand why there were such differences between the two models.

Overall, I consider the analytic methods used in the analyses of data prior to their inclusions in the SS are sound and that the SS provides very flexible stock assessment platform to accommodate needs for the POP stock assessment. However, I noticed that most model diagnoses done are rather qualitative (often eyeballed). Even if likelihood function values were provided, they only provided relative performance among alternative model configurations. This makes it hard to evaluate the model performance. More efforts should be focused on the evaluation of the model performance and quality of parameter estimates. I suggest conducting a simulation study to generate a simulated POP fishery based on the current stock assessment results and apply the SS to assess this simulated fishery with different process and observation errors. Such an excise can be done in conjunction with efforts to develop MSE framework for

the POP. It can provide insights about the model performance and identify key sources of uncertainty in modeling. Following the previous STAR Panel, I would suggest that a simpler age-structured model be developed to evaluate possible differences resulting from different stock assessment models.

IV-2-4. Evaluate model assumptions, estimates, and major sources of uncertainty and provide constructive suggestions for improvements if technical deficiencies or additional major sources of uncertainty are identified.

Three time blocks were used for fishery retention based on the analysis of changes in fishing gears and regulations over the time period covered in the stock assessment. Time varying selectivity was considered in the sensitivity analysis. It is concluded that time varying selectivity may result in risk of overfitting. Time blocks developed based on the analysis of temporal changes in regulations and fishing gear seemed to be reasonable for the POP. Foreign fishing fleet might be an issue because it might use different gears (fine mesh sizes).

The triennial survey only covered partial distribution of the POP in the early year, and as a result was assumed to have a dome-shaped selectivity. All other surveys were assumed to have logistic selectivity because they tended to cover the whole spatial ranges of POP. This seems to be a reasonable assumption. However, alternative selectivity functions such as double logistic models could be explored to evaluate impacts of this assumption on the stock assessment. During the review week, different selectivity functions were assigned to different survey programs, but fitting of size composition did not seem to be improved or in some cases become even worse for most survey length frequency data. More studies are needed.

Retrospective patterns were evaluated for 2000 recruitment deviation, virgin biomass estimate, 2006 depletion and 2005 SPR ratio. However, retrospective errors were not evaluated. For example, retrospective errors could be estimated for the stock biomass in 2004. I also believe that the retrospective errors were not explicitly dealt with in developing management parameters such as ABC.

Given possible errors associated with input data (e.g., size composition and abundance index data), use of robust likelihood function (e.g., robust multinomial distribution and t distribution; Chen et al., 2000) may be more desirable.

During the review, it was realized that major uncertainty in the stock assessment was likely to come from uncertainty associated with M and h. Parameters M and h tended to be correlated negatively. To evaluate their impacts on the assessment, varying combinations of h and M were evaluated with respect to likelihood profiles. Likelihood profile analysis showed that h and M could not be estimated precisely because the likelihood profile was very flat over a range of values of h given M or over an arrange of M given h. It was realized that the M and h could not be estimated independently in the assessment. Fixing one parameter and estimating the other one were considered a good option. After a long discussion, the STAR Panel recommended that M be fixed at 0.05 for both males and females while h was estimated. However, the flat likelihood

profiles over a large range of h values led to large uncertainty in the status of stock ranging from "overfished status" to "recovered status".

Because of needs to estimate B_0 for management parameters, the 2011 stock assessment started in the 1940s when there was little fishing activity. However, little information on population abundance and recruitment was available during such an early time period. Thus, the recruitment and stock biomass estimated for the early time period were questionable. Sensitivity analysis was done to evaluate impacts of starting the assessment from 1960 when there was information about recruitment and stock abundance. It appears that the current stock status was rather robust to the start year in the assessment.

Different selectivity functions were evaluated. Asymptotic selectivity for the fishery tended to yield optimal results about the status of population. More studies need to be done to identify/justify the choice of selectivity functions for the fishery.

Functional S-R relationship was assumed in the assessment. Both Ricker and Beverton-Holt models were used. I suggest estimating recruitment freely in the model and then evaluating the relationship between SSB and recruitment. Alternatively, I suggest developing indices for recruitment and SSB from the survey data and having their relationship evaluated to justify the functional relationship assumed in the model. If there is no relationship, such a functional relationship should not be assumed.

Retrospective analysis was done in the assessment. However, we only had limited discussion about retrospective errors. No quantification of retrospective errors was done and no retrospective errors were considered in harvest control rule. I suggest more study be done to improve understanding of retrospective errors associated with key stock parameters such as SSB, F and recruitment.

The assessment yielded a rage of management parameters with different measures, such as B_{MSY}, B_{MSY} proxies, and they should be explicitly distinguished in labeling so that comparison among different model scenarios can be done consistently.

One of the most important assumptions implicitly assumed in the SS-based POP assessment was that landing was free of errors. This assumption was certainly violated in modeling because of large uncertainty associated with landing data, when the data collection began, in particular, landing of the foreign fishing fleet. I did not see that a systematic approach was developed to evaluate impacts of violating this assumption. Two approaches may be possible to evaluate uncertainty resulting from violating this assumption. One is to assume catch is also subject to errors, develop a likelihood function for predicted and observed catches and give a small weight to this likelihood function to see the level of catch predicted by the rest of the data in the assessment. The other approach is to quantify the uncertainty associated with catch and develop a probable distribution of catch and conduct MCMC runs with each run randomly drawing catch from the defined distribution. This approach can directly incorporate uncertainty of catch into stock assessment.

IV-2-5. Determine whether the science reviewed is considered to be the best scientific information available.

This stock assessment represents the first effort of using the SS model to assess the POP stock. The POP STAT team developed a comprehensive data set for the SS-based stock assessment. Although there is still room for more improvement, given all the constraints, I consider the science reviewed in this STAR review represents the best scientific information available.

IV-2-6. Provide specific suggestions for future improvement in any relevant aspects of data collection and treatment, modeling approaches and technical issues.

Given lack of evidence of two independent POP stocks between the US and Canada, a transboundary stock assessment would be ideal. If this is impossible in a short term, informal exchanges of stock assessment information between the two countries may be a good first step to move towards a transboundary stock assessment.

Given the importance of the catch data in the assessment (assumed to be error free in the SS stock assessment), I suggest conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs. I suggest estimating uncertainty associated with catch estimates to develop a plausible range of catch estimates, which can be used in the sensitivity analysis to evaluate impacts of uncertainty associated with catch estimates on stock assessment. Landing in the early fishery, in particular catch of foreign fishing fleet, should be carefully evaluated for uncertainty.

Reconstructed landing data were often made by a few individuals and should be thoroughly evaluated to ensure the data quality and their impacts on the assessment should be evaluated in the sensitivity analysis.

I *STRONGLY* suggest that the STAT members be involved in analyzing and developing discard data for future PS stock assessment. Without the STAT team's involvement, it is difficult for them to evaluate the quality of discard data and its potential impacts on the stock assessment.

Potential temporal changes in length at age and weight at age need to be evaluated. I believe such an evaluation should be done outside the SS. For example, length-at-age data collected in different years can be fitted to von Bertalanffy growth function and a comparison can be done to evaluate differences in growth models among years.

More analysis can be done to evaluate impacts of weighting factors including giving very large or very small weights to some data series to evaluate their impacts on the quantification of uncertainty in stock assessment.

Residual patterns of age/size composition data should be more thoroughly evaluated to help identify possible temporal trends in growth and selectivity.

I suggest conducting a simulation study to generate a simulated POP fishery based on the current stock assessment results and apply the SS to assess this simulated fishery with different process and observation errors. Such an exercise can be done in conjunction with efforts to develop MSE framework for the POP. I also suggest that a simple age-structured model be developed to evaluate possible differences in stock assessment resulting from different stock assessment models.

Retrospective errors associated with key fisheries and management parameters should be evaluated more thoroughly. I believe that retrospective errors for the POP assessment were biased errors, and their impacts on the assessment were not fully evaluated in the STAR review.

The virgin biomass estimate might be subject to large errors because of lack of information on stock abundance index and unreliable estimates of catch data in early years. Current stock biomass estimate might be subject to large uncertainty because of retrospective errors. Current virgin biomass-based harvest control rule might vary from year to year with changes in the estimates of virgin and current biomass. Large uncertainty associated with virgin and current stock biomasses makes such a control rule undesirable. I suggest MSE be developed to overcome the problem.

There is a need to conduct habitat suitability modeling to identify suitable habitats for PS and POP, to outline potential habitat maps, and to help improve survey design (Chang et al., 2010).

IV-2-7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations

The POP STAT Team presented the STAT Panel an initial compilation of input data, model configuration, and modeling results during the first day of the review. The STAR Panel asked questions about the quality and quantity of input data, model configuration and relevant justifications, model assumptions, and modeling results. At the end of the discussion, a list of requests was made to the STAT team for more information/clarification/analysis, alternative model configurations/parameterizations, and new model runs. The STAT team worked overnight and came back the next day to present their replies, followed by more requests for extra analyses. This iterative process was repeated from Monday (June 20) to Friday (June 24), resulting in the development of the base case scenario and extra scenarios for sensitivity analysis for the 2011 POP stock assessment.

Large uncertainty in some key parameters and sensitivity of results to some model assumptions (different parameterizations) were expected, given issues related to sparse data, lack of survey data in early fishery, and potential issues with quality of fisheries data. Because of the time limit, uncertainty of different sources was not thoroughly evaluated in this review.

Reconstructed data should be reviewed prior to stock assessment because the STAR Panel did not have time and lacked the ability to review the data quality and STAT members were not involved in and had a lack of knowledge about how data were collected, processed and analyzed. Lack of transparency on how discard data were derived was also an issue that should be resolved prior to stock assessment.

I commend the STAT team for their efforts to address concerns raised by the STAT Panel in a timely manner during the review week. The whole process was open and rather efficient, and the discussion was very constructive. However, because of large amount of background information the STAR Panel needs to go through and the complexity of stock assessment model with a large number of input data and model configuration options, I feel we did not have enough time to cover all the alternative model runs we would like to run for the POP stock during the review week. Not all the information (e.g., details on the estimation of discards and their associated uncertainty) was available to the STAT Panel. Because of the tight schedule during the review week, I felt that I did not have enough time to carefully read all the materials distributed during the review week.

V. Conclusions and Recommendations

I would like to commend the PS and POP STAT teams, NWFSC and PFMC for providing necessary background information on the PS and POP life history, fishery-dependent and fishery-independent monitoring programs, stock assessment history, and management issues. I was impressed by the breadth of expertise and experience of the participants, the amount of effort spent to compile the data, the constructive discussion alternative approaches/suggestions, and the constructive dialogs between the reviewers and other participants throughout the review. Most materials were sent to me in a timely manner and almost all my requests for extra information and extra runs were addressed promptly.

Overall, I believe that both the PS and POP stock assessments provide rather robust assessment results, in particular on temporal trends, for the PS and POP stocks with respect to various uncertainties in data and models. The assessment revised after the STAR Panel review adequately addresses management requirements and represents the best science available. However, I believe some important questions still need to be addressed, and there is still room for improvement. Although I have provided detailed comments and recommendations under each TOR, I would like to re-iterate the following recommendations:

- In-depth analysis should be conducted to identify and quantify uncertainty for a given set of data BEFORE the data are inputted in the SS model. Trying to resolve all uncertainties within the SS model may complicate parameter estimation, resulting in difficulty in the model converging.

- Outliers are likely to exist in input data used in the assessment, given that the data are derived from different sources and are subject to different errors. They may bias parameter estimation in stock assessment if normal or log-normal distribution was assumed for likelihood functions. Robust likelihood functions can reduce impacts of outliers in size composition and survey abundance index (Chen et al., 2003).

- I recommend that the performance of the projection done in the past assessment be evaluated, retrospectively, to evaluate their performance in achieving the management objectives.

- I suggest that assessment model structure and parameterization be kept relatively stable over time. If a new model needs to be used, it should be run in parallel to the old model to identify changes in stock assessment results ensuing from changes in model configurations.

- Given the evidence of interactions between the PS and POP stocks in the US and Canada, a transboundary stock assessment would be ideal. If this is impossible in short term, informal exchanges of stock assessment information between the two countries may be a good first step to move towards a transboundary stock assessment.

- Given the importance of the catch data in the assessment (assumed to be error free in the current stock assessment), I suggest conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs. I suggest estimating uncertainty associated with catch estimates to develop a plausible range of catch estimates, which can be used in the sensitivity analysis to evaluate impacts of uncertainty associated with catch estimates on stock assessment.

- Reconstructing historical landing data was often made by a few individuals and should be thoroughly evaluated to ensure the data quality and their impacts on the assessment should be evaluated in the sensitivity analysis.

- STAR Panel did not have enough time and background knowledge to review the quality of all the input data during the week of the review. Evaluating the quality of the input data should be done prior to the STAR review with experts who have a good knowledge of particular fishery-dependent and fishery-dependent monitoring programs. A data workshop held prior to the STAR review may address this issue.

- I *STRONGLY* suggest that the STAT members be involved in analyzing and developing discard data for future PS stock assessment. Without the STAT team's involvement, it is difficult for them to evaluate the quality of discard data and its potential impacts on the stock assessment.

- Potential temporal changes in length at age and weight at age need to be evaluated for both PS and POP. I believe such an evaluation should be done outside the SS.

- More analysis needs to be done to evaluate impacts of weighting factors including giving very large or very small weights to some data series to evaluate their impacts on the quantification of uncertainty in stock assessment.

- Likelihood profile analysis for the POP showed that h and M could not be estimated precisely because the likelihood profile was very flat over a range of values of h given M or over an arrange[ment?] of M given h. It was realized that the M and h could not be estimated independently in the assessment. Fixing one parameter and estimating the other

one was considered a good choice. After a long discussion, the STAR Panel recommended that M be fixed at 0.05 for both males and females while h was estimated. However, the flat likelihood profiles over a large range of h values led to large uncertainty in the status of stock ranging from "overfished status" to "recovered status". More study is needed to evaluate this uncertainty.

- For both PS and POP, I suggest conducting a simulation study to generate a simulated fishery based on the current stock assessment results and apply the SS to assess this simulated fishery with different process and observation errors. Such an exercise can be done in conjunction with efforts to develop MSE framework for the PS and POP.

- Although the SS is very flexible and has been tested and used in the assessment of many fisheries stocks, the results derived still need to be cross-validated to enhance the confidence in the assessment. I believe some competitive models of different complexities should be developed to compare with the SS. A comparative study of stock assessment, begot from different models, can help improve understanding of fish population dynamics modeled by the SS.

- Because of the difficulty in separating PS landing data between WA and OR (in particular historical information) and similarity between the two fisheries, I suggest combining the WA and OR PS fisheries data in the assessment.

- The virgin biomass estimate might be subject to large errors because of lack of information on stock abundance and unreliable estimates of catch data in early years. Current stock biomass estimate might be subject to large uncertainty because of retrospective errors. Current virgin biomass-based harvest control rule may vary from year to year with changes in the estimates of virgin and current biomass. Large uncertainty associated with virgin and current stock biomasses makes such a control rule undesirable. I suggest MSE be developed to overcome the problem.

- There is a need to evaluate effects of bin width in data aggregations on residual patterns. Dynamic binning and robust multinomial likelihood functions should also be evaluated in reducing impacts of large number of bins with zero observation on model fitting.

- Retrospective errors for the POP assessment were biased errors and the nature of retrospective errors was unclear for the PS assessment. However, their impacts on the assessment were not fully evaluated. More studies need to be done to evaluate retrospective errors associated with key fisheries and management parameters.

- Residual patterns of age/size composition data should be more thoroughly evaluated to help identify possible temporal trends in growth and selectivity.

VI. References

Allen, L.G., D.J. Pondella II, and M.H. Horn (eds). 2006. The ecology of marine fishes: California and adjacent waters. University of California Press, Los Angeles. 660 pp.

Casillas, E., L. Crockett, Y. deReynier, J. Glock, M. Helvey, B. Meyer, C. Schmitt, M. Yoklavich, A. Bailey, B. Chao, B. Johnson, and T. Pepperell. 1998. Essential Fish Habitat West Coast Groundfish Appendix, National Marine Fisheries Service, Seattle, Washington. Located at website http://www.nwr.noaa.gov/1sustfsh/efhappendix/page1.html

Chang, J. H., Y. Chen, D. Holland, and J. Grabowski. 2010. Estimating spatial distribution of American lobster *Homarus americanus* using habitat variables. Marine Ecological Progress Series 420: 145–156.

Chen, Y., Y. Jiao, and L. Chen. 2003. A general approach to developing robust frequentist and Bayesian stock assessment methods in fisheries. Fish and Fisheries 4:105-120.

Gunderson, D.R. 1977. Population biology of Pacific ocean perch, Sebastes alutus, stocks in the Washington-Queen Charlotte Sound region and their response to fishing. U.S. Natl. Mar. Fish. Serv., Fish. Bull. 75:369-403.

Gunderson, D.R. 1978. Results of cohort analysis for Pacific ocean perch stocks off British Columbia, Washington, and Oregon and an evaluation of alternative rebuilding strategies for these stocks. Unpubl. rep., 20 p. Available Pacific Fishery Management Council, Suite 420, 2000 S.W. First Avenue, Portland, OR 97201.

Gunderson, D.R., and T.M. Sample. 1980. Distribution and abundance of rockfish off Washington, Oregon, and California during 1977. U.S. Natl. Mar. Fish. Serv., Mar. Fish. Rev. 42(3-4):2-16.

Haltuch, M. A., A. Hicks, and K. See. 2011. DRAFT Status of the U.S. petrale sole resource in 2010. Northwest Fisheries Science Center, NOAA, Seattle WA, 98112.

Hamel, O. S. and K. Ono. 2011. DRAFT Stock Assessment of Pacific Ocean Perch in Waters off of the U.S. West Coast in 2011. Northwest Fisheries Science Center, NOAA, Seattle WA, 98112.

Harry, G. 1956. Analysis and History of the Oregon Otter-Trawl Fishery. PhD Thesis, Universityof Washington, Seattle. 328 p

Love, M. 1996. Probably More Than You Want To Know About The Fishes Of The Pacific Coast. Really Big Press, Santa Barbara, California, 381 p.

Love, M.S., C.W. Mecklenburg, T.A. Mecklenburg, and L. Thorsteinson. 2005. Resource inventory of marine and estuarine fishes of the west coast and Alaska: A checklist of north Pacific and arctic ocean species from Baja California to the Alaska-Yukon border. USGS, Seattle, WA. OCS study MMS 2005-030 and USGS/NBII 2005-001.

MRAG Americas, Inc. 2003. Evaluation and analysis of current field sampling in North Pacific groundfish fisheries: Task 1: Biological sampling protocols. A report submitted to National Marine Fisheries Service Alaska Fisheries Science Center, Seattle, Washington (Contract No. 50ABNF900054).

Pearcy, W.G., M.J. Hosie, and S.L. Richardson. 1977. Distribution and duration of pelagic life of larvae of Dover sole, Microstomus pacificus; rex sole, Glyptocephalus zachirus; and petrale sole Eopsetta jordani, in waters off Oregon. Fish. Bull. 75:173-183.

Pikitch, E.K., D.L. Erickson, and J.R. Wallace. 1988. An evaluation of the effectiveness of trip limits as a management tool. Northwest and Alaska Fisheries Center, National Marine Fisheries Service NWAFC Processed Report No. 88-27.

Rogers, J.B. Unpublished document (2003). Species allocation of Sebastes and Sebastolobus species caught by foreign countries off Washington, Oregon and California, U.S.A. in 1965-1976.

VII-1. Appendix I: List of materials received prior to and during the STAR review (files can be downloaded from ftp://ftp.pcouncil.org/pub/GF_STAR_2_2011_POP_PetraleSole/)

- 1_CoverLetter_STARPanel_POP_PetraleSole_June20-24_2011.doc
- 2_Tentative Agenda_POP_PetraleSole.doc
- 3_Participants_STAR2_POP_PetraleSole_2011.doc
- 4_GF_Stock Assessment_ToR_2011-12.docx
- 5_PreSTARDraftPOPassessment2011.pdf
- 6_PreSTARDRAFT_PetraleSole_Assessment_2011-0607.pdf
- BOTH_File_naming_conventions.docx
- Background_Final Summary Report from NWFSC Bott?Trawl Survey Workshop.pdf
- Background_POP_Assessment_2007_Final.pdf
- Background_POP_Assessment_2009_Final_SAFE_version.pdf
- Background_PetraleSole_2005_Assessment_FINAL_102405.pdf
- Background_PetraleSole_2005_STARReport_Mop-UpReport_9-05.pdf
- Background_PetraleSole_2005_STARReport_South_31Aug05_final.pdf
- Background_PetraleSole_2009_Assessment_SAFE_11_17_09.pdf
- Background_PetraleSole_2009_STARReport_final_v2.pdf
- Background_STAR_Panel_Rpt_POP_2003.doc
- Canadian-POP-2010-Assessment-Draft-101110.pdf
- DRAFT_Oregon_commercial_landings_1889-1986_v1.0.042611-1.pdf
- Mprior.ppt
- POP_Presentation_POP Requests1 2011.ppt
- POP_Presentation_POPSTARdata2011.ppt
- POP_Presentation_POP_Decisiontable2011.pptx
- POP_Presentation_POP_Decisiontable_2011.doc
- POP_Presentation_POPrequests2.2011.pptx
- POP_Presentation_POPsarDraft9acc.pdf
- POP_Presentation_STAR062111_model_results.pptx
- POP_Requests_Tuesday.docx
- POP_Requests_Wednesday.docx
- POP_reports_Canada assessment 2009.pdf
- Petrale_2009_Responses_to_SSC_E2a_ATT3_0909.pdf
- Petrale_Assessment_1999.pdf
- Petrale_BombRadiocarbon.DRAFT.MS.03.21.11.doc
- Petrale_Canada_2009_070_e.pdf
- Petrale_Presentation_Day1.Requests.ppt
- Petrale_Presentation_Panel.Day2.ppt
- Petrale_Presentation_Panel.Day3.ppt
- Petrale_Presentation_Panel.Day4.ppt
- Petrale_Presentation_Panel.Day5.ppt
- Petrale_Presentation_PetSTAR.062011.ppt
- Petrale_Presentation_Retrospectives.pptx
- Petrale_Presentation_contourMap.nTows.year.pdf
- Petrale_Presentation_contourMap.nTows.year.zoom.pdf

- Petrale_Presentation_summerContourMap.catchRate.allMonths.pdf
- Petrale_Requests_Monday_STAR 2 P sole requests.docx
- Petrale_Requests_Thursday_STAR 2 P sole requests.docx
- Petrale_Requests_Tuesday_STAR 2 P sole requests.docx
- Petrale_Requests_Wednesday_STAR 2 P sole requests.docx
- Petrale_Sole_STAR_09_final_v2.docx
- Petrale_meanLengthAtAgeOverTime.pdf
- Petrale_presentation_Modeling nonlinear CPUE with a power function.pptx
- Prior steepness.xls
- SS3-OUTPUT-X.xlsm
- SS_Change_Log.xlsx
- SS_Data_Input.xlsx
- SS_User_Manual_3.21.pdf
- all_profiles2011.xls
- Density plot natural mortality.pdf
- petrale_STAR_99.docx
- ss3.exe

VII-2. Appendix II: Attachment A: Statement of Work for Dr. Yong Chen

External Independent Peer Review by the Center for Independent Experts

Stock Assessment Review (STAR) Panel for Pacific Ocean Perch and Petrale Sole

Scope of Work and CIE Process: 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 Technical Representative (COTR), 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 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.

Project Description: Both Pacific ocean perch and Petrale sole are rebuilding species and are being considered for benchmark assessments in the upcoming assessment cycle. A benchmark assessment for Petrale sole is necessary to address unresolved data and modeling issues, as well as to explore the development of commercial CPUE indices. The last benchmark assessment for Pacific ocean perch was conducted in 2003 and has been updated during each assessment cycle. It is the only species with an individual ABC whose recent assessments have not been conducted using Stock Synthesis. 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 for CIE Reviewers: Two CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. One of the CIE reviewers will participate in all STAR panels held in 2011, except for the than Pacific hake, to provide a level of consistency between the STAR panels. Reviewers should have expertise in fish population dynamics, with experience in the integrated analysis modeling approach, using age- and size-structured models, use of MCMC to develop confidence intervals, and use of Generalized Linear Models in stock assessment models. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Seattle, Washington during the dates of 20-24 June 2011.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also 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 COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE 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 CIE reviewers who are non-US citizens. For this reason, the CIE 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/sponsor.html).

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 Lead Coordinator 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 SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review. Documents to be provided to the CIE reviewers prior to the STAR Panel meeting include:

- The current draft stock assessment reports;
- Previous stock assessments and STAR Panel reports for Pacific ocean perch and Petrale sole;
- 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.
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE**

Lead Coordinator. 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 Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. 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: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each 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.

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 tentatively in Seattle, Washington during the tentative dates of 21-23 June 2011.
3) In Seattle, Washington during the tentative dates of 20-24 June 2011 as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than 7 July 2011, 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 Die, CIE Regional Coordinator, via email to ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

17 May 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
7 June 2011	NMFS Project Contact sends the CIE Reviewers the pre-review documents
20-24 June 2011	Each reviewer participates and conducts an independent peer review during the panel review meeting
7 July 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
21 July 2011	CIE submits CIE independent peer review reports to the COTR
28 July 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

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 COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR 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 CIE 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: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,
(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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:

NMFS Project Contact:

Stacey Miller
National Marine Fisheries Service, 2032 SE OSU Drive, Newport OR 97365
Stacey.Miller@noaa.gov Phone: 206-437-5670

Michelle McClure
National Marine Fisheries Service, 2725 Montlake Blvd. E, Seattle WA 98112
Michelle.McClure@noaa.gov

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 Statement of Work
   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 for Pacific Ocean Perch and Petrale Sole

1. Become familiar with the draft stock assessment and background materials.
2. Comment on the quality of data used in the assessments including data collection and processing.
3. Evaluate and comment on analytic methodologies.
4. Evaluate model assumptions, estimates, and major sources of uncertainty and provide constructive suggestions for improvements if technical deficiencies or additional major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. Provide specific suggestions for future improvement in any relevant aspects of data collection and treatment, modeling approaches and technical issues.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations

VII-3. Appendix III: STAR Panel review schedule

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 for
Pacific Ocean Perch and Petrale Sole

June 20-24, 2011
Hotel Decca
4507 Brooklyn Avenue NE
Seattle, WA 98105

Monday, June 20, 2011

9:00 a.m.	Welcome and Introductions
9:15 a.m.	Review the Draft Agenda and Discussion of Meeting Format (Panel Chair)
	- Review Terms of Reference for Assessment and Review Panel
	- Assignment of reporting duties
	- Discuss and agree to format for the final assessment document
9:45 a.m.	Stock Assessment Team (STAT-1) Presentation of Species 1 (Authors)
	- Overview of Data and Stock Synthesis Modeling
12:30 p.m.	Lunch (On Your Own)
1:30 p.m.	Q&A session with the STAT-1 & Panel discussion
3:30 p.m.	Coffee Break
3:45 p.m.	Panel develops request for additional model runs / analyses for STAT 1
4:30 p.m.	Panel provides written requests for additional model runs / analyses to STAT 1
5:00 p.m.	Adjourn for day.

Tuesday, June 21, 2011

9:00 a.m.	Stock Assessment Team (STAT-2) Presentation of Species 2 (Authors)
	- Overview of Data and Stock Synthesis Modeling
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Q&A session with the STAT-2 & Panel discussion
3:00 p.m.	Coffee Break
3:15 p.m.	Panel develops request for additional model runs / analyses for STAT 2
4:00 p.m.	Panel provides written requests for additional model runs / analyses to STAT 2
4:30 p.m.	Panel check in with STAT-1 if needed
5:00 p.m.	Adjourn for day.

Stock Assessment Review (STAR) Panel for Pacific Ocean Perch and Petrale Sole

Wednesday, June 22, 2011

9:00 a.m. STAT-1 Presentation of first set of model runs for Species 1
- Q&A session with the STAT-1 & Panel discussion
- Panel develops written request for second round of model runs / analyses for STAT 1

12:00 p.m. Lunch (On Your Own)

1:30 p.m. STAT-2 Presentation of first set of model runs for Species 2
- Q&A session with the STAT-2 & Panel discussion
- Panel develops written request for second round of model runs / analyses for STAT 2

3:30 p.m. Coffee Break

3:45 p.m. Continue Panel discussion with STAT-2

5:00 p.m. Adjourn for day.

Thursday, June 23, 2011

9:00 a.m. STAT-1 Presentation of Second Set of Model Runs for Species 1
- Q&A session with the STAT-1 & Panel discussion
- Identification of preferred model and elements for the decision table.
- Panel develops third list of model runs for decision table and begins drafting STAR report.

12:00 p.m. Lunch (On Your Own)

1:00 p.m. STAT-2 Presentation of Second Set of Model Runs for Species 2
- Q&A session with the STAT-2 & Panel discussion
- Identification of preferrred model and elements for the decision table.
- Panel develops third list of model runs for decision table and begins drafting STAR report.

3:30 p.m. Coffee Break

3:45 p.m. Panel discussion or report drafting continues

5:00 p.m. Adjourn for day.

Friday, June 24, 2011

9:00 a.m. Consideration of remaining issues
- Review decision tables for Species 1 and Species 2

11:00 a.m. Panel agrees to process for completing final STAR report by Council's September meeting Briefing Book deadline

5:00 p.m. Review Panel Adjourn.

VII-4. Appendix IV. Composition of the PS and POP STAR Review

Participants of the
Stock Assessment Review (STAR) Panel for
Pacific Ocean Perch and Petrale Sole

June 20-24, 2011
Hotel Decca
4507 Brooklyn Avenue NE
Seattle, WA 98105

Technical Reviewers
Ray Conser, Panel Chair, Scientific and Statistical Committee (SSC)
Kevin Stokes, Center for Independent Experts (CIE)
Yong Chen, Center for Independent Experts (CIE)
Jim Ianelli, NMFS, Alaska Fisheries Science Center (NMFS/AFSC)

Panel Advisors
John DeVore, Pacific Fishery Management Council (PFMC) Staff
Daniel Erickson, PFMC Groundfish Management Team (GMT)
Pete Leipzig, PFMC Groundfish Advisory Subpanel (GAP)

Stock Assessment (STAT) Teams
Owen Hamel, Northwest Fisheries Science Center, Pacific ocean perch STAT
Kotaro Ono, University of Washington, Pacific ocean perch STAT
Melissa Haltuch, Northwest Fisheries Science Center, Petrale sole STAT
Allan Hicks, Northwest Fisheries Science Center, Petrale sole STAT
Kevin See, University of Washington, Petrale sole STAT

VII-5. Appendix 5: STAR Panel Requests for Extra Analyses During the Review

Requests Made for Petrale Sole

Requests made on June 20 2011 (Monday)

1. REQUEST Canadian picture: Any maps of catches, assessment results (status, relative stock size) etc. for comparison
2. REQUEST FOR PRESENTATION on meta analysis on M by Hamel – likely as part of POP
3. REQUEST explanation (provide draft ms on ftp) re bomb calibration and ageing error work
4. REQUEST plot of unfiltered CPUE for winter and summer using 80, 90 and 100% criteria. Depending on outcome decide on CPUE sensitivity run(s)
5. REQUEST spatial plots of unfiltered effort by year for winter fishery.
6. REQUEST for aggregate fits (normalised sums of obs and pred across all years for each survey and fishery – look for gross evidence of variation between surveys and fisheries)
7. REQUEST note no continuity run so try run with parameters fixed from 2009 (to look at what effects on Likelihoods etc) (stepwise fix M, h, growth params,)
8. REQUEST explore further downweighting of comp data, better fitting of survey series
9. REQUEST compare spatial extent of survey strata by areas relative to CPUE areas
10. REQUEST plot of priors and fits (for existing draft base case and any further runs)
11. REQUEST sensitivity test using old vs new OR catches (when have Base case)
12. REQUEST sensitivity to discard rates (2.8% versus 8.8%) (when have Base case)
13. REQUEST check on dynamic binning options in SS3 (after base case)

Requests made on June 21 (Tuesday)

1. Exchange conditional age-length data for marginal age compositions; fix growth at current base case values.
2. Do runs adding winter CPUE and beta fixed at 1 and 0.5 (and look at residuals).
3. Do run with summer CPUE and no change in q.

Requests made on June 22 (Wednesday)

1. Do a run with common selectivity for WA and OR fleets (note future exploration of joining data)
2. Look at data to explore growth changes through time (Note a likely RECOMMENDATION on investigating changes in growth, how to include in future assessments, and possibly issues of BRP definition);
3. Produce maps of survey and fishing coverage by year.
4. Increase robustification constant from 0.0001 to 0.01 on candidate base case (Note that in final runs may combine also with dynamic binning as being explored from Day 2 requests).

5. Do additional run estimating beta or winter CPUE and plot fits for this and run with beta = 1 (already done)

<u>Requests made on June 23 (Thursday)</u>

To get to base Model:

1. Fix SD of L@A (old), so that constant across ages
   Add 0.001 constant to all proportions in comps
   Reanalyzed Pikitch ratios from Dan
   Betas = 0
2. Tune surveys (with adjustment from the above base)
3. Tune CPUE (using an initial CPUE CV of 0.35)

Sensitivities

1. Estimate beta (all 3 independently)
2. Use 2009 OR catch series
3. Mirror OR/WA selex

Exploration

1. Dome shaped selex by block
2. Dropping conditional A@L (how much?)
3. Residuals on A@L
4. Look at tuning of comps

<u>Requests made on June 24 (Friday)</u>

No request was made.

Requests Made for Pacific Ocean Perch

Requests made on June 20, 2011 (Monday)

No request was made because the first day was focused on PS assessment.

Requests made on June 21 (Tuesday)

1) Use discard rates over time from Pikitch data
    a. Better data exists than what was assumed in base case as presented
2) Check discard sample size used
    a. Seems like actual number of fish and different than survey and fishery approaches used
3) Omit 2004 age data from the survey (perhaps it may be okay for the marginal age compositions) unless it can be corrected
4) Compare mean weights at age from 2009 assessment to this year
    a. Need a way to compare growth
5) Exchange conditional age-length data for marginal age compositions
    a. In the bridge analysis and elsewhere it was apparent that the composition data had a large impact—fix growth if needed
6) Check old model numbers over time (say age 3) with stock synthesis cross (A). Investigate what may be causing the difference in recent trend and in B_{msy} and other reference point estimates.
    a. To try to better understand the difference between old and new assessments
7) Try a run with R_1 specified
    a. See if that improves the behavior of the single year class
8) Do a run with and without the Oregon catch reconstruction
    a. A sensitivity to this has not been completed?
9) Try a run with higher s_R (say 2.0 or 3.0) and steepness fixed at 1.0
    a. See if M estimates change
    b. Compare dynamic B_0 relative to the base case
10) Show pairwise diagnostic plots of MCMC chain
    a. May show correlations among parameters and if there are parameters that are poorly determined
11) Summarize results from recent Canadian assessment
12) Show plots of priors on M and h relative to previously used values
13) Provide table and summary of the meta-analysis used for steepness prior
14) Provide maps showing coverage of the surveys relative to the fishery

Requests made on June 22 (Wednesday)

1. Drop 2004 age data from NWFSC survey
2. Look at mean weight at age data from survey; compare to old model input and new SS output
3. Make run with no S/R run using age comps for yrs available; plot S/R results

4. Refine base model then continue investigation of the principal factor(s) that caused differences in depletion and h from the 2009 model

Requests made on June 23 (Thursday)
Decision table was presented
Requests made on June 24 (Friday)

No request was made.