Pacific Ocean Perch Stock Assessment Review (STAR) Panel Report

NOAA Fisheries, Northwest Fisheries Science Center
2725 Montlake Blvd East
Seattle, WA 98112

26-30 June 2017

STAR Panel Members:

Dr. David Sampson	Oregon State University, Scientific and Statistical Committee (SSC), STAR Panel Chair
Dr. Panayiota Apostolaki	Center for Independent Experts
Dr. Norman Hall	Center for Independent Experts
Dr. Kevin Piner	NOAA Fisheries, Southwest Fisheries Science Center, SSC

Stock Assessment Team (STAT) Members:

Dr. Chantel R. Wetzel	NOAA Fisheries, Northwest Fisheries Science Center (NWFSC)
Mr. Lee Cronin-Fine	University of Washington
Mrs. Kelli F. Johnson	NOAA Fisheries, Northwest Fisheries Science Center (NWFSC)

STAR Panel Advisors:

Ms. Lynn Mattes	Oregon Dept. of Fish and Wildlife, Groundfish Management Team
Mr. Louie Zimm	Groundfish Advisory Panel
Mr. John DeVore	Pacific Fishery Management Council

Overview

A Stock Assessment Review (STAR) Panel met during 26-30 June 2017 at the Northwest Fisheries Science Center (NWFSC) auditorium in Seattle, Washington to review a draft stock assessment for Pacific ocean perch (*Sebastes alutus*). The assessment had been prepared by a stock assessment team (STAT) led by Dr. Chantel Wetzel of the NWFSC and was documented in Wetzel et al. (2017). The Panel operated under the Pacific Fishery Management Council's (PFMC) Terms of Reference for stock assessment reviews (PFMC 2016). This same panel also reviewed a draft assessment for lingcod (*Ophiodon elongatus*).

Pacific ocean perch (POP) were the target of distant-water foreign fishing fleets operating off the west coasts of U.S. and Canada during the mid-1960s to mid-1970s. This species also occurs off northern Japan but in the eastern Pacific the species is most abundant in the Gulf of Alaska and occurs as far south as Baja California. The portion of population off the U.S. West Coast, which has generally been treated as a single separate stock for assessment and management purposes, was declared overfished in 1999. The current assessment, as in the most recent previous assessment (Hamel et al. 2011), assumes that POP off the U.S. West Coast are a single, unit stock whose dynamics are independent of POP populations off Canada and in the Gulf of Alaska.

The stock assessment was conducted with Stock Synthesis (version 3.30.03.05). Data were compiled into three fishing fleets and several indices of relative abundance. The three fishing fleets were a commercial fleet (combining three gear-types: bottom trawl, midwater trawl and fixed gear), a historical foreign fleet, and the at-sea hake fleet. These fishing fleet definitions were based on discard practices, with unreported discards only assumed to occur in the commercial fleet. For this fleet the model included a retention curve and was informed by observer data on discard rates and length-compositions. Several indices of relative abundance were considered during development of the model, including fishery dependent CPUE indices and fishery independent surveys (POP survey, the Triennial survey, the AFSC slope survey, the NWFSC slope survey and the NWFSC shelf-slope survey). The STAT team identified the NWFSC shelf-slope survey as providing the most reliable information on population abundance and data from that source were prioritized in the model explorations. Selectivity was estimated for each modeled fleet using observations of age and/or length composition data, except for the historical foreign fleet, which assumed the same selectivity as the commercial fleet because no data were available for the foreign fleet to inform its own selection curve. The NWFSC shelf-slope survey age data were included as conditional age-at-length observations (CAAL). For some fleets both the age and length composition data were included. All fleets and surveys were modeled with double normal selectivity parameterizations except for the POP survey, which was forced to estimate an asymptotic selectivity pattern. The assessment model was structured to have two sexes and it started from an unfished non-equilibrium state in 1918 with annual recruitment deviations estimated to 2015.

Results for the base model developed during the STAR meeting are summarized as follows. The model estimates that the spawning output of POP at the start of 2017 was 6,966 million eggs and was at 96% of its unfished level (i.e., almost at the unfished level). The trajectory of spawning output has been increasing steadily since about 1990 and underwent a rapid increase during the last three years, largely due to exceptionally large recruitment in 2008 and strong recruitment in 2013. The assessment estimates that the stock's spawning output hovered slightly above the Council's target level (40% of unfished) for a period extending from about 1970 to 1990 but never dropped below the below the Council's the minimum stock size threshold (MSST, 25% of unfished, i.e., the stock was never overfished).

The estimated dynamics from this stock assessment are considerably different from the prior assessment, which indicated the stock was in a depleted state below the MSST for a period extending from 1980 to 2011, reaching a low point of 14% depletion in 1998. The differing views of the stock from the two assessments was partly due to the inclusion of new data, but changes in key assumed life-history parameters (natural mortality and steepness) and the elimination of data from the Triennial survey were more influential. Overall, this assessment appears to rely heavily on compositional data, especially the conditional age-at-length data from the NWFSC shelf-slope survey.

The Panel concluded that the base model is appropriate for use by management and would constitute the best available science. Although the biomass indices included in the model were uninformative regarding scale and the model is heavily dominated by compositional data, the panel considers that the use of surveys, composition and estimation of recruitment deviations makes this a Category 1 assessment. The Panel applauds the STAT team for their well-structured presentation of the assessment and the very competent work completed before and during the STAR meeting.

Summary of Data and Assessment Model

Catch series and fishing fleet structure

The fishing fleet definitions were changed from the 2011 assessment, which had a single fleet that combined catches from 1940-2010 from trawl, fixed gears, at-sea hake, foreign fleet, and surveys. In the new assessment there are three coast-wide fishing fleets defined in the model based on discarding practices. A commercial shore-based fishing fleet (1918-2016) was defined that aggregated catch from bottom trawl, midwater trawl and fixed-gear. A second fleet was defined as the historical foreign fleet (1966-1976, based on catch reconstructions from Rogers 2003), and the third fishing fleet was the at-sea fleet targeting Pacific hake (1975-2016). Only the first fleet was treated as producing unreported discarded catch of POP. The other fleets were assumed to retain (and report) their catches of POP. Significant catches in the commercial fleet occurred after 1940 but were greatly reduced after 2000 due to the determination that the stock was overfished. The largest annual catches, in the late 1960s by the foreign fleet, were three to four times the annual catches of the commercial fleet at that time. The at-sea hake catches are quite small in comparison to the other fleets. Over the period modeled by the assessment (1918-2016), the cumulative catch of POP by the commercial fleet (landings plus model-estimated dead discards) was 106.7 million metric tons (mmt, 71.6% of the overall cumulative estimated removals), whereas the cumulative catch by the foreign fleet a bit less than 41.3 mmt (27.7% of the overall cumulative estimated removals) and catches by the at-sea hake fleet were a bit less than 1.1 mmt (0.7% of the overall cumulative estimated removals).

Survey indices

Several indices of relative abundance covering different periods were considered by the STAT team including: the NWFSC shelf-slope trawl survey (2003-2016), the NWFSC slope survey (1999-2002), the AFSC slope survey (1996-2001, excluding 1998), the Triennial survey (1980-2004, every third year) and a Pacific Ocean Perch survey (1979 and 1985). The STAT Team indicated that the NWFSC shelf-slope survey was likely the most representative data source in the assessment. Data from the NWFSC shelf-slope, AFSC slope, Triennial and POP surveys had been analyzed using the VAST software to develop the biomass indices. Data from the NWFSC slope survey was analyzed using a Bayesian delta-GLMM. In fitting the assessment model to the survey indices an extra variance parameter was estimated for the Triennial and NWFSC shelf-slope surveys; the other surveys did not need additional variance based on standardization estimates of observation error and the initial assessment model runs. During the STAR the Triennial survey was removed from the base model.

Fishery-dependent indices

One commercial fishery CPUE index from the Vancouver and Columbia INPFC areas for the period (1956-1973) was considered. It was based on a published paper by Gunderson (1977) but little information was available on the development of this index. The coefficients of variation associated with the annual observations for this index were set to 0.4, which was similar to the fishery independent surveys. During the STAR Panel this commercial fishery index was shown to have little influence on the results and was removed from the base model.

Compositional data – lengths

Sex-specific length-composition data were included for the commercial fishing fleet (1966-2016) and at-sea hake fleet (2003-2016), as well as for the surveys: Triennial (1980-2004); POP (1979-1985); AFSC slope (1997-2001); NWFSC slope (2001-2002); and NWFSC shelf-slope (2003-2016). All length-composition data were treated as observations of the catch-at-length and linked to the dynamics using length-based observation model processes (i.e., selection was strictly length-based). During the STAR the length-compositions from the Triennial survey were removed from the base model.

Compositional data – ages

Sex-specific age-composition data were included for the commercial fishing fleet (1981-2016) and at-sea hake fleet (2003-2014) as well as for the surveys: Triennial (1989-2004); POP (1985); NWFSC slope (2001-2002); and the NWFSC shelf-slope (2003-2016). The NWFSC shelf-slope survey age data were treated as conditional age-at-length (CAAL); the rest were used as observations of the catch-at-age (marginal age-compositions) and linked to the dynamics using length-based observation model processes (i.e., selection was strictly length-based). During the STAR the age-compositions from the Triennial survey were removed from the base model.

Discards Data

Modelling of the historical foreign fleet assumed no discarding. The at-sea-hake fleet reported catch that included discards, and thus discards did not need to be modelled for this fleet. The commercial fishing fleet included discard information from two sources: (1) discard weight ratios (1985-1987) and sex-specific length-composition data of discard (for 1986, from Pikitch et al. 1988) and (2) West Coast Groundfish Observer Program discard ratios (2002-2015) and length-composition data (2004-2015) of discarded fish that informed the retention curve for the commercial fishing fleet.

Maturity and weight-length relationships

Life history parameters that were fixed in the model included coefficients for the weight-length relationship, the maturity versus length relationship, fecundity per female body weight, natural mortality and steepness of the Spawner-Recruit relationship. The length-at-age (growth) relationship and associated process variability were estimated in the assessment model and primarily evaluated by fits to the CAAL composition data associated with the NWFSC shelf-slope survey.

Bridging analysis

The STAT presented a bridging analysis that examined the effects of changing the version of the Stock Synthesis software from version 3.24 (used to produce the 2011 assessment) to version 3.30. The analysis demonstrated that the version of the software had negligible effects on the assessment estimates of spawning output. The Panel endorsed the use of the newer software.

The assessment model

The assessment model was a two-sex, length- and age-based age-structured model that estimated dynamics starting in 1918 with the assumption of equilibrium with no fishing prior to the start of the model. Recruitment deviations were estimated from 1900-2016 with deviations prior to the start of the dynamic period used to increase uncertainty around the unfished state. Selectivity was modelled as double-normal (domed) for all fleets and surveys except the POP survey, which was forced to be asymptotic. Discarding by the commercial fleet was modelled using logistic retention curves. Several time blocks for the commercial fishing fleet retention curve were used that corresponded to changes in management that influenced fishing and or discard practices.

The final agreed base model was well structured, was thoroughly investigated by the STAT, and is the best currently available for the formulation of management advice.

Treatment of uncertainty

The base model included estimates of uncertainty for estimated parameters and derived quantities such as spawning output and depletion. The STAT also explored uncertainty of the base model results using likelihood profiles across the key parameters M, steepness (h), and ln($R0$). Although the fixed values for M and h in the base model were based on prior distributions from meta-analyses, the prior distributions were not included in the likelihood and therefore did not inform the model and contribute to the model's estimates of uncertainty. The likelihood profiles for the final base model indicated the available data were almost uninformative regarding the values for both M and h. Further, the likelihood profile for ln($R0$) indicated that the survey indices were essentially uninformative regarding population scale, with all the information coming from the compositional data.

Requests by the STAR Panel and Responses by the STAT

The pre-STAR draft assessment document was very complete and the STAT's opening presentation to the STAR Panel anticipated many questions regarding why the draft model's results were so different from the previous assessment. This allowed for an efficient and effective review that could quickly identify the most important questions and allocate review time accordingly. The STAT provided thorough responses to all requests.

Requests below are provided sequentially by the day of the request. Responses from the STAT team (which were generally delivered the following day) are given below each request. The bolded sentences within each Response (if any) are major conclusions drawn by the STAR Panel that were considered important in the construction of the final base mode. Figures from those important requests are also often given.

Request 1.1: *Provide the figure from "2011 Model Data "Update" and "2017 Base Model sensitivities" from today's presentation and provide a written description of plots for our report.*

Rationale: To better describe what has changed from the 2011 assessment, specifically why the stock is not so depleted in the new assessment in the 1990-2000s.

Response: The STAT provided the requested figures and text (see below).

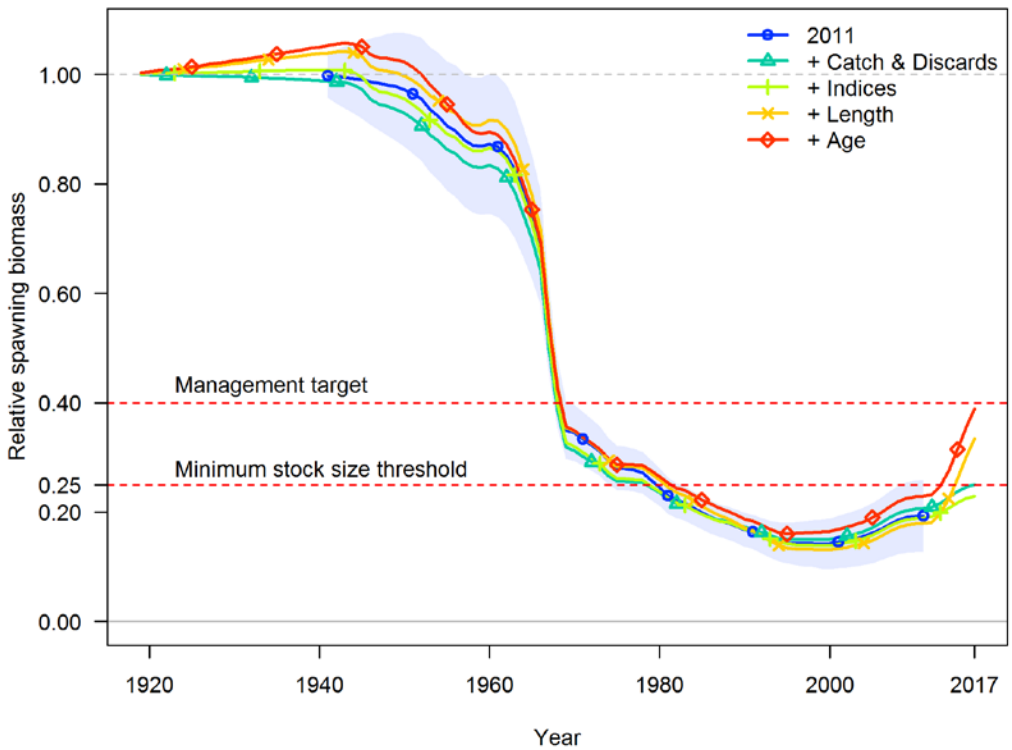

Figure 1. Estimated relative spawning biomass based on the 2011 model structure when new data are added to the model. No modifications were made from the assumed 2011 base model structure beyond extending the model end year to 2016 and extending the ITQ block for fishery retention to the final model year of 2016. Each colored line represents an additive data update to the 2011 model where "+ Catch & Discards" reflects the change in estimated status when catch and discard data were updated to the 2017 values, "+ Indices" reflects the estimate of status including the updated indices, catch and discard data, and so forth.

Request 1.2: *Estimate the number of POP caught in the shrimp fishery by year.*

Rationale: To gauge the impacts of the shrimp fishery catch, as catch weight is low but the fish are tiny.

Response:	The STAT provided estimates of the numbers caught recorded in the WCGOP (using mean/median length to calculated wt) for 2004-2015. The numbers caught ranged from 1300-310,000 fish. This information is not used in the assessment model but to gauge the numbers of fish taken in this fishery. The numbers appear to be trivially small.

Request 1.3: *Contact the NWFSC survey group with the question, are expansions based on numbers or weight.*

Rationale:	To make sure that the expansions of the compositional data are based on the numbers of fish (rather than their biomass).

Response:	The survey team clarified that they applied an expansion based on numbers per hectare. The STAR Panel agreed that this was the appropriate approach.

Request 1.4: *Provide a model run with all indices and associated composition data removed except the NWFSC shelf/slope survey. Provide alternative runs with VAST and design-based indices. There should be no estimated extra_SD on the survey index any of these runs.*

Rationale:	To avoid data conflicts associated with multiple indices. Alternative runs will help the Panel to understand the effect of using the model-based versus design-based survey indices.

Response:	A model fit to only the NWFSC survey gives higher population scaling than the draft base model (Fig.3, provided below in the response to Request 1.5). It was determined that the omission of the Triennial survey data was responsible for this change in scale (see response to Request 1.5).

The design-based index from the NWFSC survey had similar trends as the VAST model-based index but differed considerably in scale and the design-based index had wider confidence bands (Fig.2). Substituting the design-based estimates from the NWFSC survey had a moderate effect on the scale of the resulting biomass but not on the depletion estimates (see Fig.3). The use of the VAST model-based index was not influential.

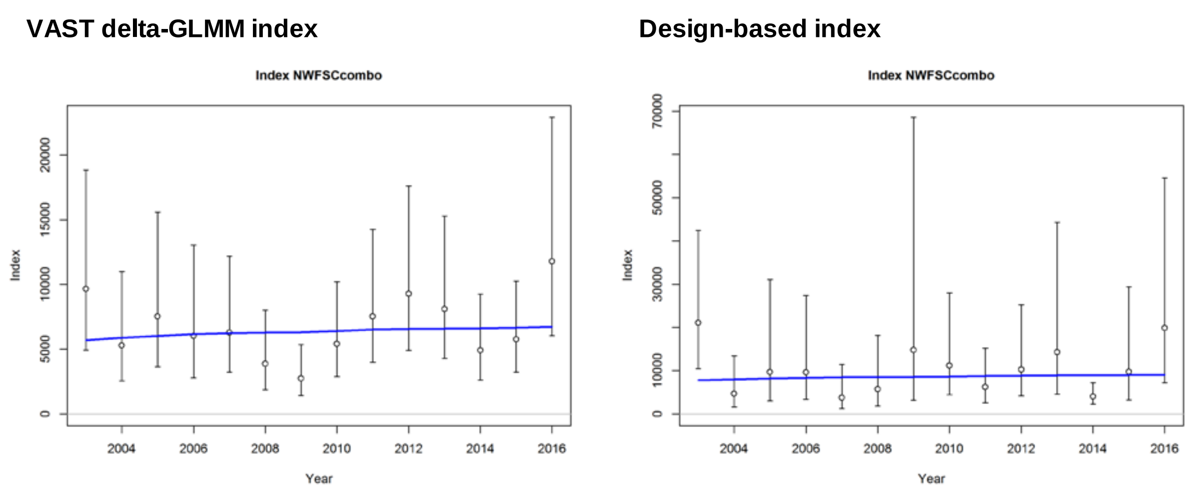

Figure 2. VAST model-based (left panel) and designed based (right panel) indices for the NWFSC shelf-slope survey.

Request 1.5: *Provide a model run with all indices and associated compositional data removed except from the NWFSC shelf/slope and Triennial surveys. Provide alternative runs with VAST and design-based indices. There should be no estimated extra_SD on the survey index any of these runs.*

Rationale: To avoid data conflicts associated with multiple indices as in Request 1.4 but extend the duration of survey information. Alternative runs will help the Panel to understand the effect of using the model-based versus design-based survey indices..

Response: A model fit using only the NWFSC shelf-slope and Triennial survey scales population lower than the draft base model (Fig.3). This run confirmed the importance of the Triennial index as a source driving model estimates of biomass to lower scale and depletion. Design-based estimates of the survey indices resulted in model estimates of spawning output with similar scale as the draft base model. With the designed-based estimates there was reduced influence of the Triennial survey due to the larger associated survey variance. Design-based surveys result in moderate differences from using VAST-based indices in estimates of assessment scale and depletion.

The STAR Panel concluded that the assessment model's treatment of the Triennial survey is a major decision-point but that the use of the VAST approach to produce the indices was not.

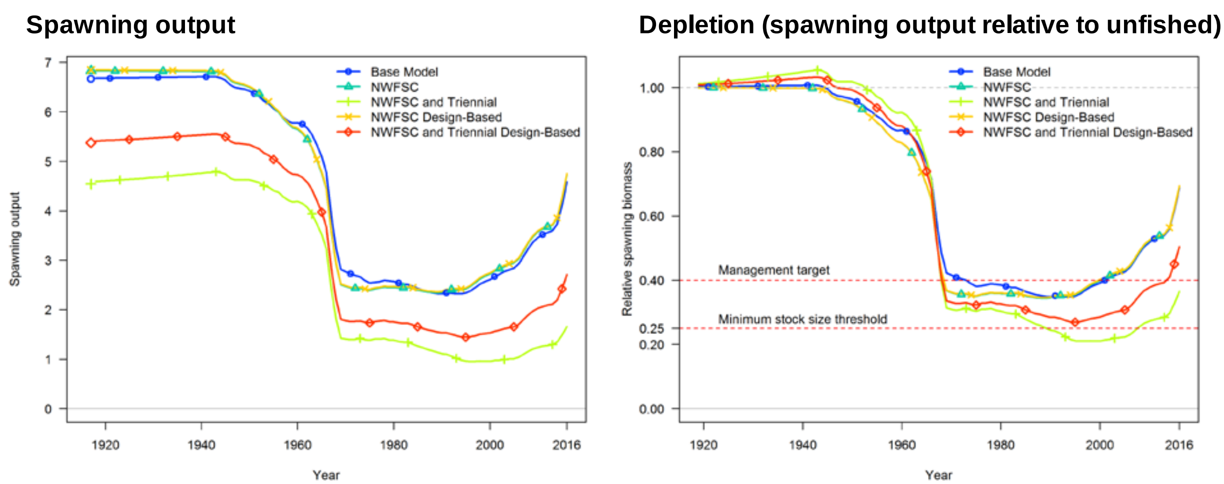

Figure 3. Spawning output trajectories for model runs responding to Requests 1.4 and 1.5. Spawning output (left panel) and depletion (right panel) from base model and models fitting NWFSC and Triennal indices that are design- and model-based (VAST). The "base model" shown here is not the final base model

Request 1.6: *Run an "age structured production model" where all estimated parameters from the candidate base model are fixed (specified and not estimated) at the base case estimates, with the exception of global scaling parameters, ln(R0) and survey catchability coefficients, which are estimated. Recruitment deviations are not estimated; instead recruitment is taken from the spawner-recruit relationship. Composition data are turned off in the model and only key indices of abundance contribute to the total likelihood. There should be no estimated extra SD for the survey indices for any of these runs.*

Rationale: Determine if fishing in the form of catch-at-age is causing the trends in the index or if index trends are determined by recruitment deviations (process variation). Connection between index trends and catch is evidence that the index is providing the information that sets the population scale.

Response: This run provides information similar to what is in a ln(*R*0) profile. The age structured production model (ASPM) documents the connection between catch and trends via the production function, while the *R*0 profile quantifies the degree of information from that relationship.

The ASPM using the NWFSC shelf-slope and Triennial surveys scaled the population to a very different unfished spawning output (higher) and depletion that was below the minimum stock size threshold (the green line with triangles in Figure 5, shown below with Request 1.8). It was expected that the population would scale lower due to the influence of the Triennial survey but the population scaled upwards instead. The Panel did not understand this surprising result.

The ASPM results do not capture trends from either of the two primary survey indices (Triennial and NWFSC shelf-slope surveys, Fig.4) and produce very different estimates of population scale. **This analysis suggests that modelled catch via the production function cannot explain the population trends and that process variability will be used in further models to explain the trend. Thus these survey indices together do not appear to provide much information on the scale of the stock and the base model's estimates of scale, implying that base-model results will be heavily influenced by the compositional data**.

ASPM fit to the Triennial index **ASPM fit to the NWFSC shelf-slope survey**

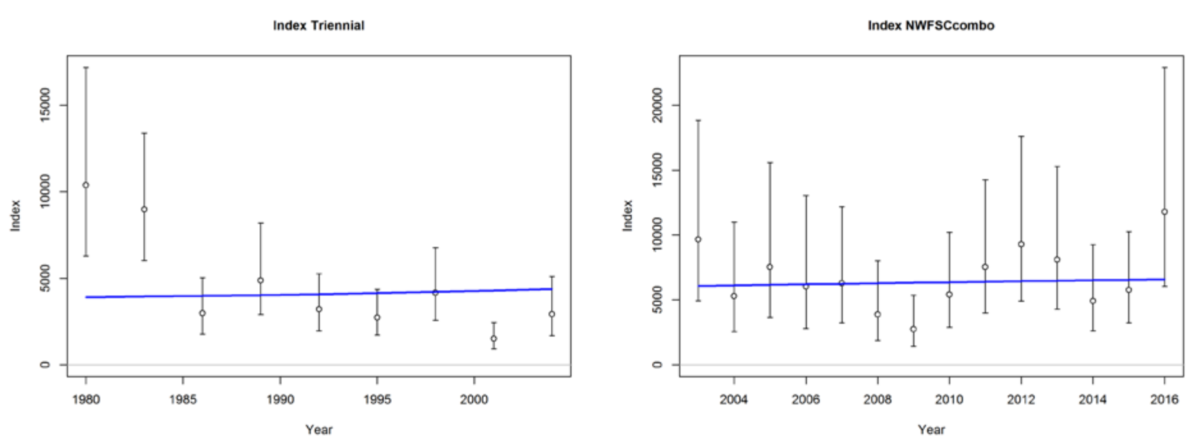

Figure 4. Fits of an age-structured production model (ASPM, the solid lines) to the Triennial (left panel) and NWFSC shelf-slope (right panel) indices (the open circles), which indicate that a

production function version of the POP model driven by the POP catch-at-age histories does not produce reasonable fits to the trends in either of the primary tuning indices. As noted in earlier runs, only the Triennial survey appears to have informative contrast, but that contrast does not appear to be consistent with the catch and production function. The later NWFSC shelf-slope does not contain informative contrast.

Request 1.7: *Provide a new Retrospective run without the adjustment to age-composition sample size for 2016 from the NWFSC shelf-slope survey.*

<u>Context</u>: For the draft base model brought to the STAR the STAT had down-weighted the input sample sizes for the 2016 NWFSC shelf-slope age-at-length data due to greater perceived variability in the data relative to previous years and greater model sensitivity to these data.

<u>Rationale</u>: To see how the STAT's ad hoc adjustment to input sample size for 2016 affected the model results.

<u>Response</u>: When addressing this request, the STAT detected an error in how the draft base model treated the NWFSC shelf-slope treatment of composition data. A lambda value of 0.50 (instead of 1.0) had been incorrectly applied to the length-compositions from the NWFSC shelf-slope survey. Correcting this error eliminated the need for this request.

Request 1.8: *Provide output from the model run with growth parameters fixed at the base model values but without including the Conditional Age-at-Length (CAAL) component likelihood in the total likelihood.*

<u>Rationale</u>: To see effects of the CAAL component beyond estimating growth.

<u>Response</u>: Removing the CAAL data from the assessment model scales the biomass estimates lower with greater level of depletion than the draft base model (Fig.5). This analysis demonstrates that the CAAL data in the base model has strong influence on the model estimates of scale and depletion. A plausible conceptual model for the life-cycle for POP (larger fish are deeper) suggests that there could be spatial patterns in the CAAL data due to ontogenetic movements motivated by age. Un-modeled age-based processes in the population that are not properly accounted for in the CAAL data would lead to biases. **The STAR panel concluded that another major decision point in the final base model is the treatment of the CAAL data**.

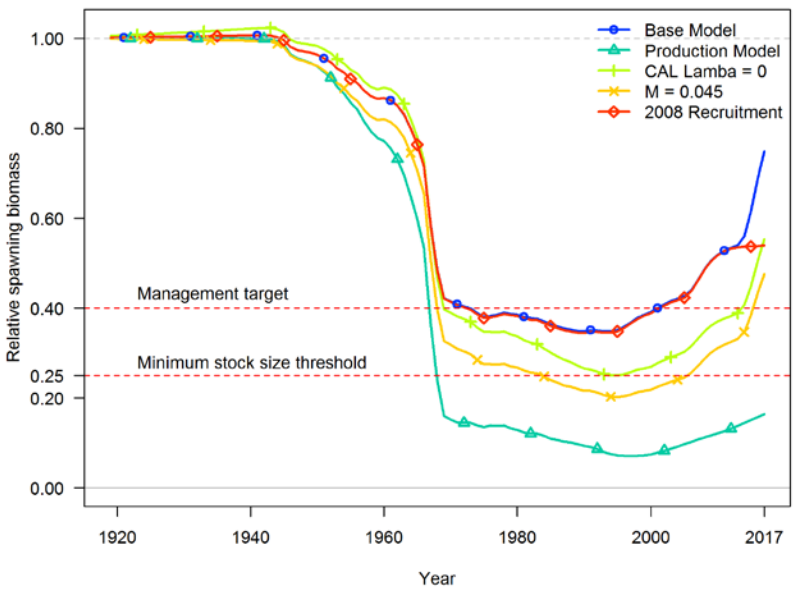

Figure 5. Relative spawning output (depletion) produced in response to Request 1.8 (shown in the light green line with crosses) and several others (Requests 6, 7, 9, and 10). The "base model" shown here is not the final base model.

Request 1.9: *Provide a model run with M that corresponds to a maximum age of 120 yrs.*

Context: The draft base model brought to the STAR assumed a fixed value for *M* (0.054 year^{-1}) that corresponded to a maximum age of 100 years.

Rationale: To understand the effect of *M* associated with a plausible value for maximum age.

Response: The lower fixed *M* (0.045 year^{-1}) lead to model estimates of a more depleted population (the golden line with Xs in Figure 5) and lower scale. Such results indicate the strong influence of the composition data (which provides information similar to a series of catch-curves). The STAR panel concluded that *M* is influential but provided no advice on how it should be determined.

Request 1.10: *Provide a model run with the 2008 recruitment deviation set to zero and provide the corresponding detailed likelihood components.*

Rationale: Identify the source of the extraordinarily large 2008 year-class estimate.

Response: The length composition data from the NWFSC shelf-slope appear to be the source of the information underlying the very strong recruitment in 2008 (Fig.6). The STAT team identified the NWFSC shelf-slope survey as the most trusted data on the population.

Draft base model **Draft base model with Rec-Dev[2008] at zero**

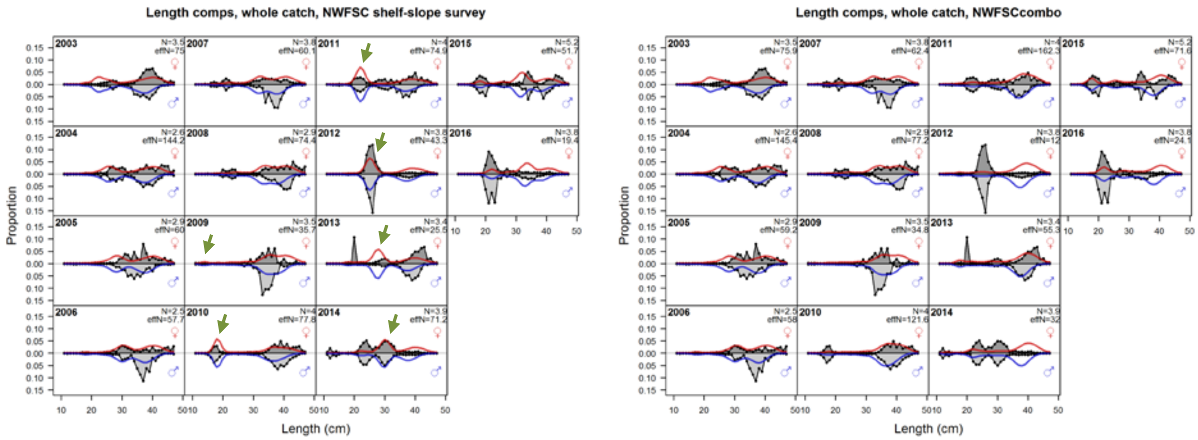

Figure 6. Fits to the length-compositions from the NWFSC shelf-slope survey for the draft base model (left) and that model with the recruitment deviation for 2008 fixed at zero (right). The green arrows (left panel only) identify the 2008 year-class.

Request 2.1: *Drop the Fishery CPUE index (set fleet number to a negative number). Add the data changes / corrections the STAT recommended the previous data. Pending review, this will be part of the new base model.*

Context: The fishery CPUE index was catch in mt per hour from the U.S. domestic fishery combined for the INPFC Vancouver and Columbia areas based on data from Gunderson (1977), which was an early "stock assessment" for Pacific ocean perch.

Rationale: These [the Fishery CPUE data] are non-reproducible / undocumented data.

Response: The STAT conducted the requested run and it made virtually no difference to the results. In this same run the STAT also corrected the 2016 sample size for the NWFSC shelf-slope CAAL data (relative to previous models), which also made virtually no change in the model results.

Request 2.2: *Starting from the response to Request 2.1 above, compare model runs with the Triennial split and not split – compare likelihoods by component.*

Context: Whether or not to split the Triennial survey has been a controversial issue of long-standing. The practice of splitting this index into separate early (1980-1992) and late (1995-2004) indices started with the 2007 assessment for canary rockfish, due to the realization during the STAR of an abrupt change in 1995 in the seasonal timing of the survey operations. During the early period the surveys started and ended about one month later than in the late period.

Rationale: To see if there is evidence in the Triennial data for a change in survey q (catchability).

Response: Splitting the Triennial index produced a small improvement in fit (about four negative log-likelihood units). Most of the improvement came in the fit to the Triennial index. However because an extra_SD parameter was estimated for the survey the likelihood comparison was not strictly fair. However the q change was as expected (see table below) with the early series having a much higher q than the late. STAR panel did not use this run to suggest splitting the series or not. That decision is left to the STAT.

	One Time-series	Split Time-series
Q POP	0.898	0.845
Q Triennial Early	0.161	0.267
Extra SD Triennial Early	0.391	0.345
Q Triennial Late	-	0.129
Extra SD Triennial Late	-	0.095
Q AFSC Slope	0.084	0.076
Q NWFSC Sloope	0.059	0.053
Q NWFSC Combo	0.074	0.067
Extra SD NWFSC Combo	0.027	0.026

Request 2.3: *Starting with the response to Request 2.1 above, run the model with only the Triennial survey and fishery composition data. Fix growth. Attempt to estimate steepness using the prior if there is time (lower priority).*

Rationale: To see what is causing the lack of fit to the early Triennial survey. Also to see if steepness is estimable if data conflicts are reduced.

Response: The STAT conducted the requested run, which resulted in a lower spawning output and lower depletion, as indicated in the table below. The estimates of steepness went to approximately 0.3. Given earlier profiles from the draft base model it was apparent that the Triennial index is driving steepness to lower levels. It was unclear if this result means that steepness should be estimated or if the Triennial should be removed from the model. The decline in the Triennial index is in some conflict with other data sources given the model's structure.

Models considered:

The Base model from Request 2.1.

A. Drop all indices and associated composition data except the Triennial survey.

B. With the model from A, estimate steepness (h) using a prior:

1. h ~ Uniform [0.2, 1.0]
2. h ~ Normal [-.5, 0.2]
3. h ~ Beta Prior [0.7606, 0.146]

	Base	A	B.1	B.2	B.3
Likelihood:					
Total	1862.5	140.6	131.1	131.6	134.7
Survey	-12.75	-1.01	-3.96	-3.94	-3.83
Discard	-32.07	-40.43	-39.27	-39.30	-39.43

	Base	A	B.1	B.2	B.3
Length	147.14	77.90	76.41	76.45	76.59
Age	1746.31	94.04	91.89	91.91	92.07
Recruitment	12.83	9.07	5.94	5.91	5.89
Parameter Priors	1.00	1.00	0.13	0.56	3.43
Parameters:					
log(R0)	9.371	8.942	9.228	9.217	9.178
Steepness	0.5000	0.5000	0.3082	0.3138	0.3341
Derived quantities:					
SB Virgin	6620.4	4257.3	5666.4	5604.5	5390.8
SB 2017	5040.5	1090.8	507.8	520.3	567.5
Depletion 2017	0.7614	0.2562	0.0896	0.0928	0.1053
Total Yield - SPR 50	1774.1	1157.7	0.0	0.0	14.1

Request 2.4: *Starting with the response to Request 2.1 above, allow dome-shaped age selection for the NWFSC combo survey.*

Rationale: Explore the consequences of violations of the CAAL assumption that ages are random at length in the survey.

Response: The STAT conducted the requested model run. The resulting age-based selection curves were estimated to be asymptotic (fully selected) for all ages, indicating no strong age-based selection. It is unclear if this would eliminate potential age-based availability issues as there is little information in the model to inform the population age distribution other than the NWFSC shelf-slope CAAL composition data.

Request 2.5: *Starting with the response to Request 2.1 above, consider more time-variant selection curves [in the trawl fishery].*

Rationale: To explore if rigid selection functions are contributing to the lack of fit to the Triennial survey.

Response: The STAT team explored several time-varying options for selectivity in the primary fishing fleet by adding time-blocks and additional parameters (length at the peak, the final selection, and the retention width). However there was little improvement in fit to the data (approximately four units). Combinations of the varying options offered slightly more improvement. The fishery selectivity patterns tended to shift towards smaller fish at the end of the modeled period, apparently to account for the higher numbers of small fish observed by the fleet. The STAT suggested a new block of selectivity (for the parameter controlling selection for the last length bin) from 2000-2016 for the fishing fleet, as this was the time period when the stock was declared overfished. Additionally, the retention curve for the post-ITQ period was allowed to shift to retain smaller fish relative to the earlier periods. These two changes improved the model fit by about nine units. However, these changes had negligible effects on model estimates of spawning output or depletion (Fig.7).

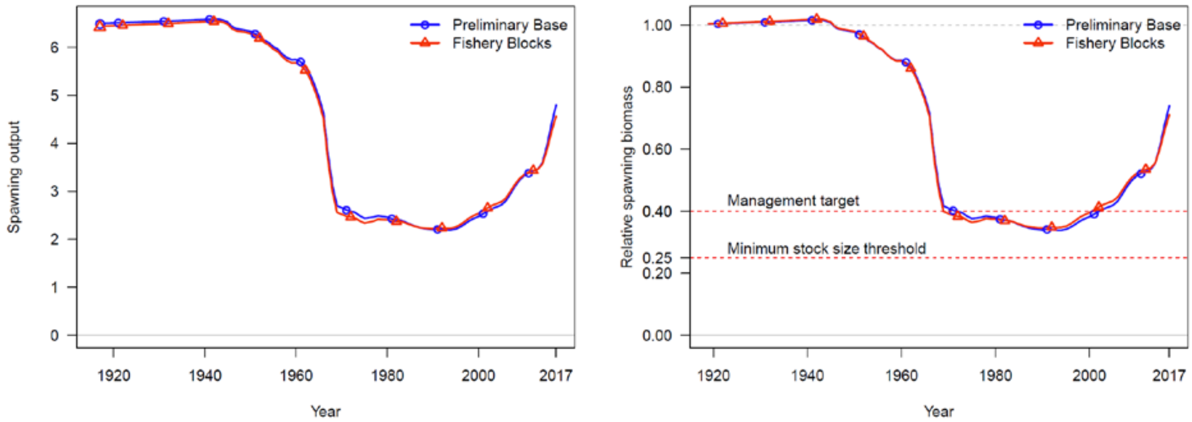

Figure 7. Estimated trajectories for spawning output (left) and relative spawning output (depletion, right) for a preliminary base model (starting from Request 2.1) to which time-blocks have been added for selection and retention by the primary fishing fleet.

Request 2.6: *The STAT should work towards a revised base model and provide a steepness sensitivity analysis for a new preferred base model.*

Rationale: The original steepness assumption (0.5) is somewhat arbitrary.

Response: For the new preferred base model the STAT decided to remove the Triennial survey (the index and the associated composition data) because the data from this survey do not appear consistent with the other data in the model (given the model structure) and they do not consider the Triennial survey to be a good one for slope rockfish. The model results are relatively insensitive to the Triennial survey with respect to the relative spawning output (depletion) but the Triennial survey does influence the absolute scale (Fig.8). The STAT team also decided to use the commercial fishery blocks developed in response to Request 2.5 and the CAAL from the NWFSC shelf-slope survey (as in previous models explored during the STAR).

With regard to sensitivity to steepness, the STAT explored using a steepness value of 0.63 (changed from the value of 0.5 in the draft base model) but decided to use the value from the Thorson analysis (h=0.72) in the preferred base model even though it produces a considerable upturn in biomass at the end of the modelled period.

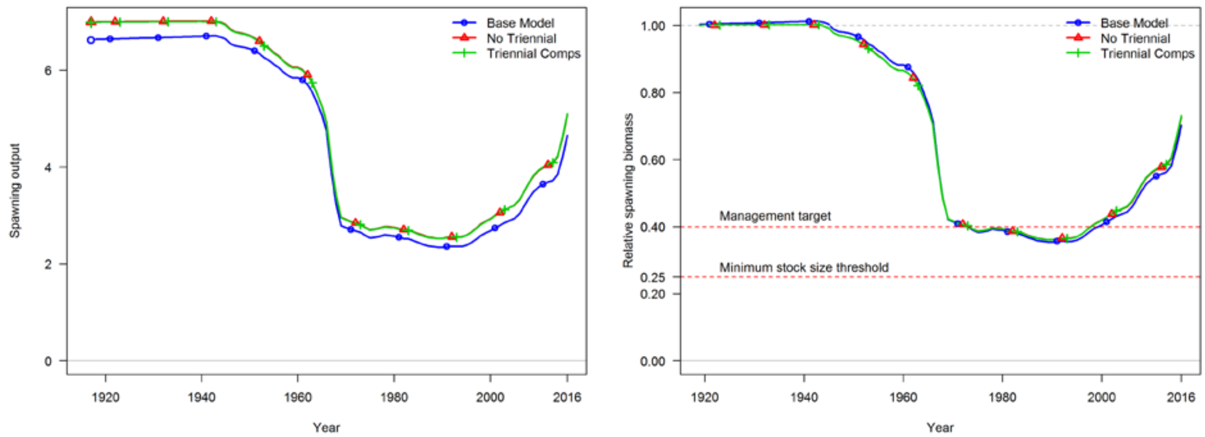

Figure 8. Estimated trajectories for spawning output (left) and relative spawning output (depletion, right) for a preliminary base model (starting from Request 2.1) that includes time-blocking for fishery selection and removes the Triennial survey data (the index, then the compositional data).

Request 3.1: *Provide a model run with only the Triennial survey (no added variance) and associated composition data. Include the fishery length-compositions. Assume h = 0.4 and alternatively h = 0.5 and 0.72 (re-weight with h= 0.5 before running the other steepness scenarios).*

Rationale: To understand how the Triennial survey might affect the results and provide an alternative view of POP population dynamics. The Triennial survey appears to provide information that conflicts with the other surveys.

Response: The STAT conducted the requested model runs (Fig.9). If only the Triennial data are used the model results are considerably more pessimistic. Thus the STAT's decision to exclude the Triennial was a substantial decision point and contributed to the difference between this assessment and previous assessments.

During post-STAR email discussion the STAT countered that the Triennial had little influence on the model results relative to the 2011 assessment because the Triennial survey in the 2011 assessment was split (early versus late) and had a large estimated extra_SD parameter. The differences between the final base model here and the 2011 assessment (especially the historical trajectory and decline in status) are due primarily to a slightly higher assumed *M* and a large increase in steepness (from 0.40 to 0.72).

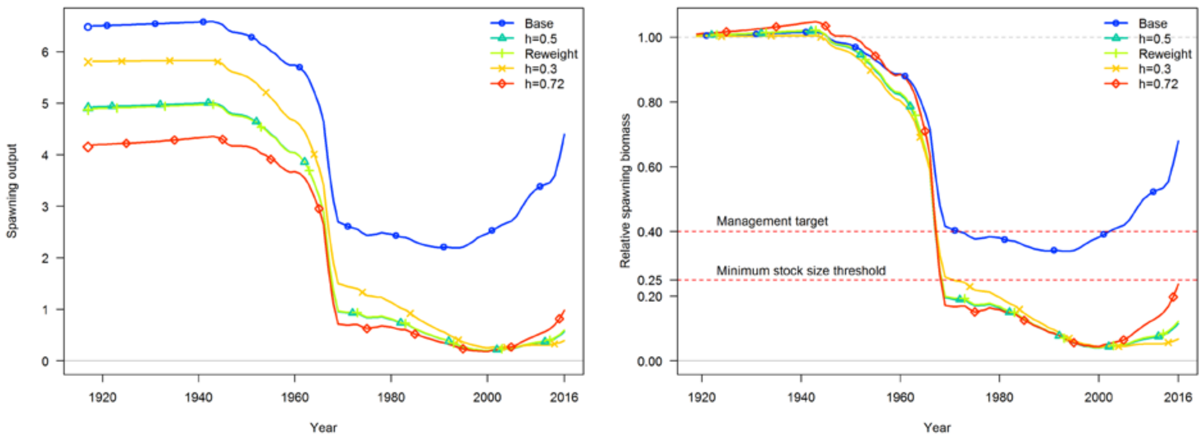

Figure 9. Spawning output (left) and depletion (right) from models using the Triennial survey as the only tuning source (labelled as “h=0.5”) compared to the revised base model (labelled as “Base”) that used other surveys but not the Triennial survey.

Request 3.2: *Provide a plot of fishery selection from the 2011 assessment and compare to the proposed base model with the time block ending at 2010.*

Rationale: To better understand how changes in fishery selectivity relative to the previous assessment may contribute to our current perception of stock productivity, status, and abundance.

Response: The STAT provided the requested comparison (Fig.10), which indicated that the fishery selection and retention as estimated in the new assessment are not markedly different from the previous assessment.

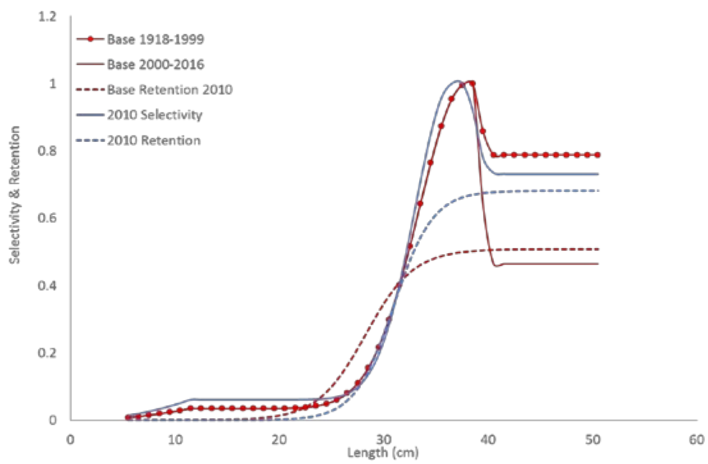

Figure 10. Fishery selection and retention curves from this assessment (in red) and the previous assessment (in blue).

Request 3.3: *Explore a potential base model that has the following specifications:*

- Remove the Triennial survey and associated composition data.
- Change fishery selection and retention time-blocks as proposed by the STAT in their response to Request 2.5.

- Assume the mean of the steepness meta-analysis (h = 0.72).

Retune the model and provide full diagnostics. Provide likelihood profiles on fixed female M, the new fixed h, and ln(R0).

Rationale: These specifications converge on a consensus base model and this request is needed as a final check on its suitability.

Response: The STAT produced the model runs based on the h = 0.72 assumption as requested (e.g., Fig.11). Supporting analyses included a profile on M, steepness, and ln(R0), as well as jitter runs (50 runs, showing no evidence the base model had not converged to the maximum likelihood parameter estimates) and a five-year retrospective analysis (showing no untoward behavior).

The fits to the majority of the indices were flat indicating the dominance of the compositional data in the assessment. The STAT concluded that this model configuration is the final base model and the STAR Panel concurred.

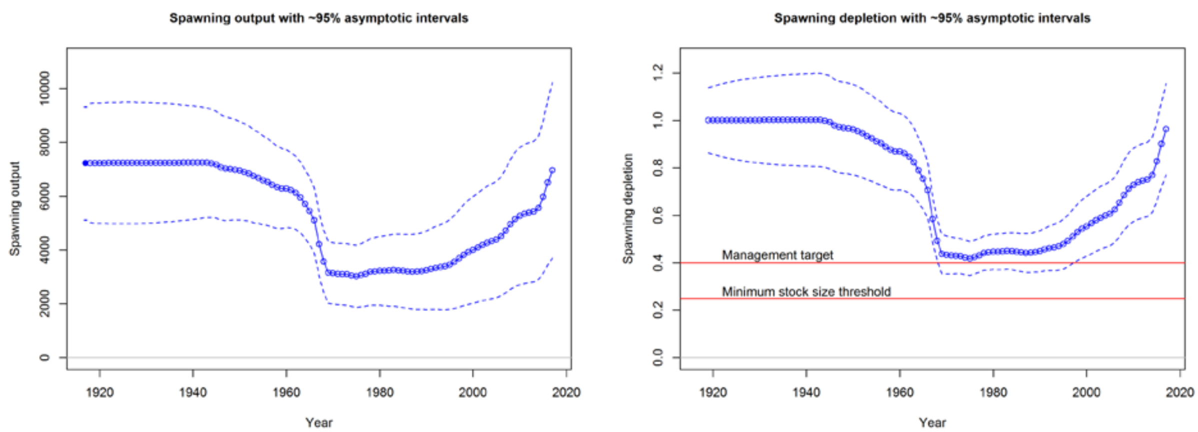

Figure 11. Estimated trajectories for spawning output (left) and depletion (right) for the proposed base model as specified in Request 3.3.

Request 3.4: *Explore quantiles on steepness from Thorson's meta-analysis for the major axis of uncertainty in the decision table.*

Rationale: This may inform the decision table.

Response: The STAT provided runs with steepness at the 12.5 percentile of the prior (h_{low} = 0.509) and at the 87.5 percentile (h_{high} = 0.894).

Request 3.5: *Explore quantiles on M for the major axis of uncertainty in the decision table.*

Rationale: This may inform the decision table.

Response: The STAT provided runs with M at the 12.5 percentile of the prior (M_{low} = 0.0325) and at the 87.5 percentile (M_{high} = 0.089).

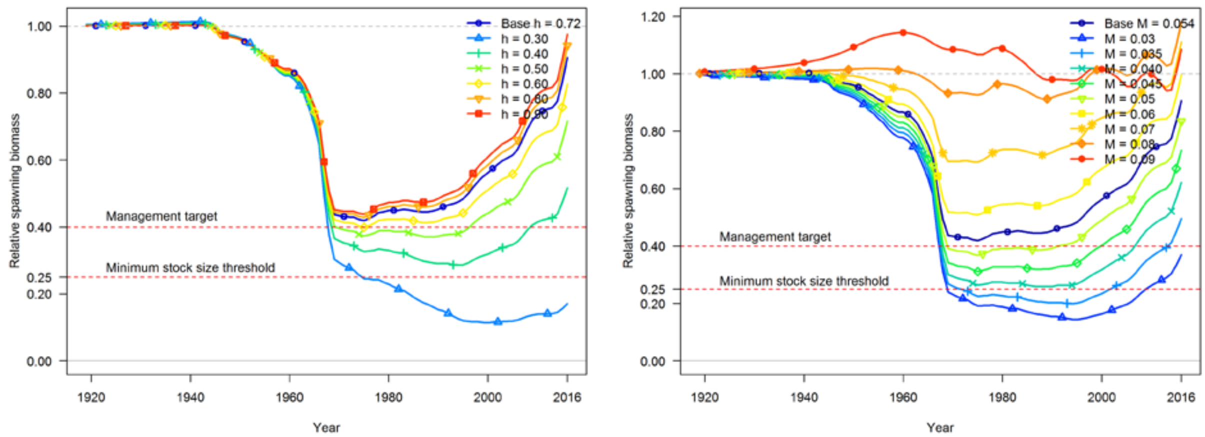

Figure 12. Relative spawning output trajectories under different assumed values for steepness (*h*, left panel) and natural mortality (*M*, right panel).

It was agreed that steepness (*h*) would provide a better axis of uncertainty than *M*. However, the steepness prior was not wide enough to capture the upper range of uncertainty in the base model's estimates of 2017 spawning output. The upper 87.5 percentile for spawning output given the estimated spawning output (6966.0 million eggs) and its associated standard deviation (1665.39) is 8881.2 million eggs. The revised plan agreed at the STAR was for the STAT to produce a decision table that uses values for natural mortality consistent with the 12.5th and 87.5th percentiles of the asymptotic variance of spawning output from the base model.

Final production of the decision table occurred after the STAR meeting concluded. During its construction the STAT found that the values of *M* that were consistent with the 12.5th and 87.5th percentiles (0.045 yr^{-1} and 0.06 yr^{-1}) did not provide a wide range of contrast relative to the base model value of *M* (0.054 yr^{-1}). The STAT produced a second decision table with low and high *M* values (0.038 yr^{-1} and 0.077 yr^{-1}) based on the uncertainty in the *M* prior to reflect the error in the estimates used in the meta-analysis which created the prior distribution.

Description of the Base Model and Alternative Models used to Bracket Uncertainty

The base model

The final base model for Pacific ocean perch was structured as having one area, one season, two sexes, and covered the period 1918-2016, with catches and recruitment deviations beginning in 1918 from an unfished, non-equilibrium age-distribution; the main period for recruitment deviations is 1940-2014. Steepness (*h*) was fixed at the mean of the beta prior (0.72); natural mortality (*M*) for both sexes was fixed at 0.054 $year^{-1}$ (the median of a prior based on a maximum age of 100 years). The model used an internal structure for ages that ranged from zero to an accumulator age of 60 and an internal structure for lengths that ranged from 5 to 50 cm in 1-cm increments. Parameters defining length-at-age were sex-specific and fully estimated except that females and males at age-3 were forced to have the same mean length and variability in length.

The final base model was informed by survey biomass indices and length-compositional data from the POP survey (1979 and 1985), the AFSC Slope survey (1996-2001, excluding 1998), the NWFSC slope survey (1999-2002), and the NWFSC shelf-slope survey (2003-2016). There were marginal age-compositions for some years for the POP survey and the NWFSC slope survey and conditional age-at-length compositional data for all years of the NWFSC shelf-slope survey. There were also length-compositional data from the commercial fishing fleet (1966-2016) and the at-sea hake fleet (2003-2016), with discard lengths for the commercial fleet (1986 and 2004-2015) to inform the length-based retention curve for this fleet. Observations of discard fractions were available for the periods 1985-1987, 1992, and 2002-2015.

Selection curves for the four surveys were fully length-based with no male offsets and forced to be asymptotic. Selection curves for the fishing fleets were also fully length-based, double-normal curves with no male-offsets. Selectivity for the at-sea hake fleet was forced to be asymptotic but selectivity for the main commercial fishing fleet was dome-shaped. The foreign fleet (for which there was no compositional data) was mirrored with the main commercial fishing fleet. The commercial fishing fleet had two blocks for its selection curve and four blocks for its retention curve.

In the final base model the length- and conditional age-at-length compositional data were iteratively weighted using the Francis approach and the model had an estimated extra_SD parameter for the NWFSC slope-shelf survey index (but not for the other biomass indices). The model was also tuned to have suitable bias adjustments for recruitment variability. There was no adjustment to the σ_R parameter value relative to the draft model brought to the STAR.

Following the STAR the STAT conducted additional jitter runs to confirm convergence of the final base model. The STAT did not find a better fitting model than the one reviewed on the final day of the STAR.

Bracketing uncertainty

The STAR and STAT agreed that the decision table should use natural mortality that would produce the 2017 spawning output consistent with the 12.5th and 87.5th percentiles of the asymptotic variance of spawning output from the base model (i.e., base model SB_{2017} +/- 1.15 times its estimated standard deviation). The Panel notes, however, that this approach ignores other potentially important aspects of uncertainty. For example, using a run fit only to the Triennial survey index (plus the survey and fleet compositional data, as in Request 3.1) would provide a pessimistic alternative that captures a major source of structural uncertainty.

After the STAR meeting concluded the STAT produced a second decision table that provided slightly wider contrast between the low and high states of nature, using low and high *M* values (0.038 yr^{-1} and 0.077 yr^{-1}) based on the uncertainty in the *M* prior.

Technical Merits of the Assessments

- The draft assessment document for POP was very well constructed and thorough in its description of the draft base model brought to the STAR and the underlying data that informed the assessment.
- The STAT anticipated many of the questions raised by the large differences in perception of the stock's dynamics between previous assessments and the new assessment. The

presentations at the beginning of the STAR were very well constructed and informative, which greatly facilitated the review process.

- The STAT was extremely responsive to STAR Panel requests and demonstrated considerable skill revising the draft base model in response to Panel requests, producing presentations to illustrate the relevant results, and working with the Panel to develop acceptable base models that addressed the major concerns during the course of the review.

Technical Deficiencies of the Assessment

- Overall, there were no serious technical deficiencies with this new assessment for Pacific ocean perch assessment, which made good use of the available data and standard modeling approaches. That said, some STAR Panelists were concerned by the model's poor fit to the indices – conflicting indices were removed from the final base model – and the fact that uncertainty in key productivity parameters (*M* and *h*) is not incorporated in the model's estimates of uncertainty.
- The process for developing the model started with fixed weights (sample sizes) for the composition data and estimating extra standard deviation (extra_SD) parameters for the Triennial and NWFSC shelf-slope survey indices. Such an approach may tend to condition decisions regarding model structure on the compositional data, thereby leading to a final model whose structure is most consistent with the compositional data. In the model for POP this tendency may have been exaggerated because the compositional data were connected to the population dynamics via relatively rigid observation processes (e.g., limited time-variation in selectivity and relatively inflexible functional forms). An alternative model development process that prioritized fitting the indices of abundance ahead of fitting the compositional data, coupled with a more flexible linking between the observation process and the compositional data, might have led to a different outcome. This is a limitation of any statistical model-selection process, which can only explore a limited range of potential model structures in a finite period of time.
- The use of conditional age-at-length (CAAL) data was very influential but lacked any model estimation of an age-based process linking the data to the underlying the population numbers at age (e.g., age-based selectivity or age-based movement). Such a modelling approach implicitly assumes that any potential spatial structure in the population's age- or size-composition is due strictly to size-based movement, which is inconsistent with the general assumption of ontogenetic movement to depth (an age-based movement process). If an age-based process links the population with the CAAL data, the data will be biased.
- The sample composition data for the surveys were expanded to the total based on design-based estimators, and therefore may not be representative of the composition from a VAST model-based estimate of relative abundance.

Areas of Disagreement Regarding STAR Panel Recommendations

Among STAR Panel members (including GAP, GMT, and PFMC representatives):

None.

Between the STAR Panel and the STAT Team:

None.

Management, Data, or Fishery Issues Raised by the GMT or GAP Representatives During the STAR Panel Meeting

None.

Unresolved Problems and Major Uncertainties

The *Technical Deficiencies* section above describes several modeling issues that warrant further consideration but these issues fall in the general topic of "how can one build a statistical assessment model that is complicated enough to model a stock's dynamics without over-stretching or exhausting the information content of the available data". This is not a nut that will be easily cracked. Below are some other, slightly more tractable issues that the final base model left unresolved.

- Pacific ocean perch off the U.S. West Coast may be a fraction of a much large population extending into Canada or even Alaska. Modelling only a part of the total population might contribute to the lack of correspondence between the survey indices and other data sources, as seen in the ln(R0) profiles and ASPM diagnostics as well as some of the observation variability. While this comment is not intended to reflect badly on the STAT's capabilities, it is important to recognize that stock structure could potentially be a major source of uncertainty regarding the assessment results.
- The indices of abundance used in the final base model provide almost no information on population scale, as demonstrated in the ln(R0) profiles examined during the review. The Triennial survey was the only index that provided signal with respect to population scale (e.g., see Figure A.1 in the Appendix, which was taken from the STAR draft stock assessment document). However, this survey was removed in the final base model due to concerns about the quality of the survey and conflicts with other data. There are large amounts of composition data in the model, with both age- and length-compositions being included for some fleets. The compositional data and catch are providing the majority of the information on the estimated and derived quantities.
- Use of CAAL composition data provides information on parameters beyond those of the length-at-age relationship. The CAAL data are robust to length-based processes (Piner et al. 2016), however they are also influenced by age-based processes (Lee et al. in press). No age-based observation model processes were used in the assessment model as a link to the data, meaning that the CAAL data were assumed to be unbiased with respect to the population. The CAAL data were shown to be very influential on the estimated dynamics beyond growth estimates. More theoretical work in this area is needed to understand how to best the use this type of information and what potential systems or observation model processes could invalidate the assumption of randomness at length.

Prioritized Recommendations for Future Research and Data Collection

Specific recommendations for the next assessment of Pacific ocean perch (POP)

1. Further investigation of POP stock structure is recommended. One approach would be to look for correlations of U.S. West Coast recruitment deviations and survey biomass estimates with corresponding results from POP assessments in Canada and the Gulf of Alaska.
2. The next iteration of this assessment could be an update assessment.

General recommendations for all assessments

1. Comprehensively evaluate the appropriateness of using the Triennial survey in assessments for other rockfish species and whether the survey should be split into early and late segments. The lingcod assessment reviewed during this STAR split the Triennial survey into separate early and late surveys, whereas the draft POP assessment brought to the STAR had a single Triennial survey.
2. Explore the assumption that conditional age-at-length data are random samples of the age-composition.
3. A standard approach for combining conditional age-at-length sample data into annual CAAL compositions should be developed and reviewed. If age data are not selected in proportion to the available lengths, simple aggregation of the ages by length-bin may provide biased views of the overall age-composition and year-class strength.
4. Further explore the VAST approach for constructing relative abundance indices. The upcoming workshop at the Center for the Advancement of Population Assessment Methodology (CAPAM) will address this issue.

Acknowledgements

The STAR Panel thanks Stacey Miller for providing excellent logistical support both in advance of the STAR meeting and through the week of the review. The Panel also acknowledges the helpful assistance provided by Dr. Jim Hastie for answering questions regarding past assessments and Council management actions and Dr. Jim Thorson for providing an overview of the VAST software and its recommended use when applied to the analysis of the NWFSC trawl survey data.

References

Hamel, O., Hamel, O.S., and Ono, K. 2011. Stock Assessment of Pacific Ocean Perch in Waters off of the U.S. West Coast in 2011. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, OR 97220.

Lee, H.H., Piner, K.R., Methot, Jr., R.D., and Maunder, M.N. (2014). Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the Pacific Ocean. *Fish.Res.*158:138-146

Lee. H.H., Thomas, L.R., Piner, K.R. and Maunder, M.N. (in press). Effects of age-based movement on the estimation of growth assuming random-at-age and random-at-length data. *J. Fish. Bio*.

Piner, K., Lee, H.H., and Maunder, M.N. (2016). Evaluation of using conditional age-at-length observations and an equilibrium approximation of the population age structure in fitting the von Bertalanffy growth function. *Fish Res. 180:128-137*

Rogers, J. (2003). Species allocation of *Sebastes* and *Sebastolobus* species caught by foreign countries off Washington, Oregon and California USA in 1965-1976. NOAA Technical Memorandum NMFS-NWFSC-57

Wetzel, C.R., Cronin-Fine, L., and Johnson, K.F. 2017. DRAFT Status of Pacific ocean perch (*Sebastes alutus*) along the US west coast in 2017. Pacific Fishery Management Council.

Pacific Fishery Management Council. 2016. Terms of Reference for the Groundfish and Coastal Pelagic Species Stock Assessment and Review Process for 2017-2018. Available from http://www.pcouncil.org/groundfish/stock-assessments/terms-of-reference/.

Appendix

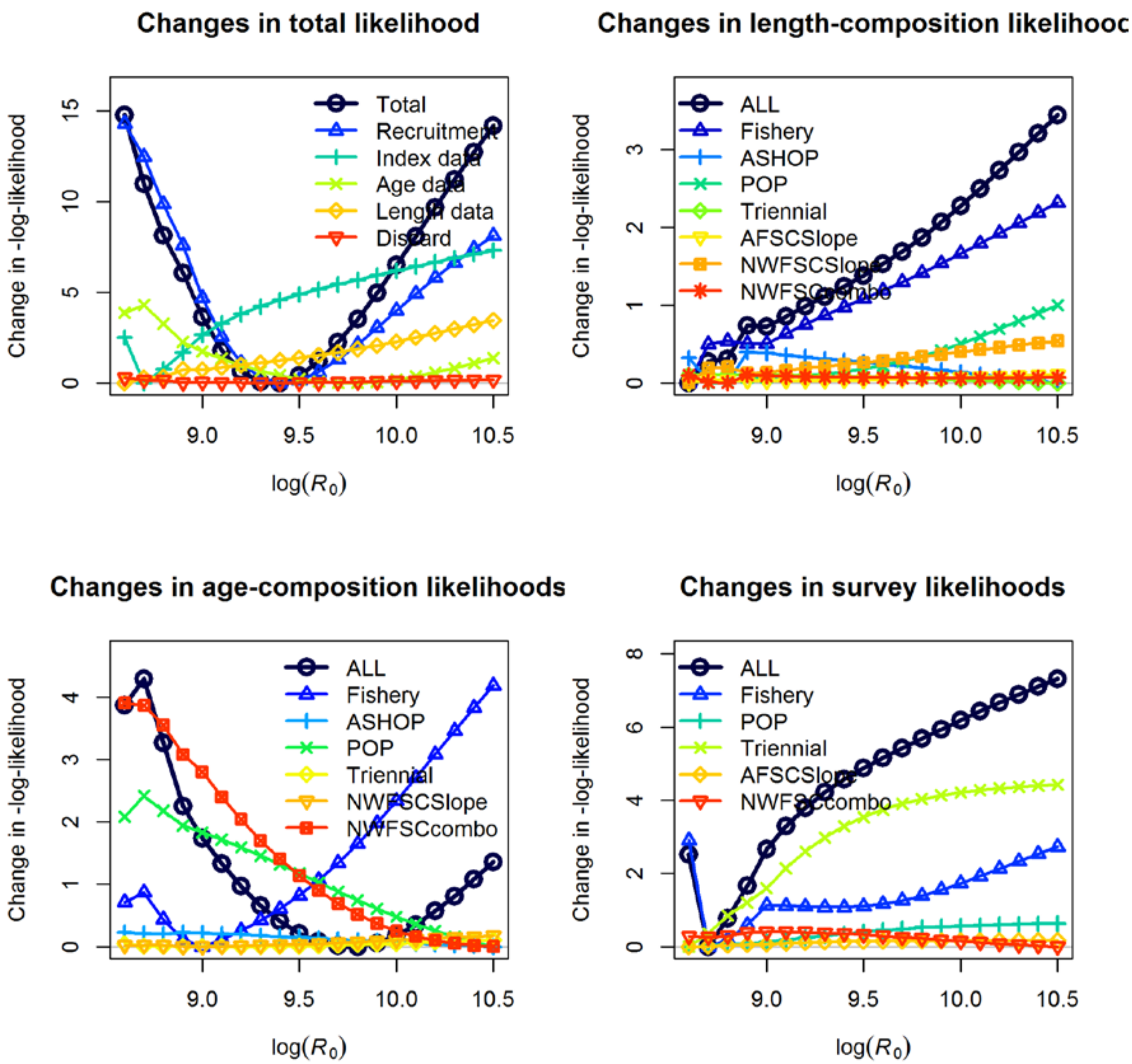

Figure A.1. Likelihood profile for the draft base model brought to the STAR (taken from the STAR draft assessment document).