Stock Assessment Review (STAR) Panel Report for Black Rockfish

NMFS, Northwest Fisheries Science Center
Newport Research Station, Bld. 955
2032 SE OSU Drive
Newport, OR 97365

July 20-24, 2015

STAR Panel Members

Dr. Andrew Cooper, Simon Fraser University, SSC (Chair)
Dr. Neil Klaer, Center for Independent Experts
Mr. Patrick Cordue, Center for Independent Experts
Dr. Paul Spencer, National Marine Fisheries Service Alaska Fisheries Science Center

Stock Assessment Team (STAT) Members:

Dr. Jason Cope, National Marine Fisheries Service Northwest Fisheries Science Center
Dr. Andi Stephens, National Marine Fisheries Service Northwest Fisheries Science Center, GMT
Dr. David Sampson, Oregon State University, SSC
Ms. Megan Stachura, National Marine Fisheries Service Northwest Fisheries Science Center
Dr. Theresa Tsou, Washington Department of Fish and Wildlife, SSC
Mr. Troy Buell, Oregon Department of Fish and Wildlife
Mr. Brett Rodomsky, Oregon Department of Fish and Wildlife
Mr. Patrick Mirick, Oregon Department of Fish and Wildlife
Mr. Phil Weyland, Washington Department of Fish and Wildlife
Ms. Kari Fenske, Washington Department of Fish and Wildlife

STAR Panel Advisors

Mr. John Holloway, GAP
Ms. Heather Reed, Washington Department of Fish and Wildlife, GMT
Mr. John DeVore, Pacific Fishery Management Council

Overview

The STAR panel reviewed a new full assessment of black rockfish (*Sebastes melanops*) off the west coast of the United States during a five-day meeting in Newport, OR. Following the consensus recommendations from a preliminary stock assessment workshop in April 2015 (PFMC 2015), the stock assessment team (STAT) prepared separate geographic assessments that were spatially stratified with boundaries at the CA/OR border (42°00' N latitude) and OR/WA border (46°16' N latitude). The last assessment for black rockfish was conducted in 2007 for areas north and south of Cape Falcon (45°46´ North latitude), excluding Canada and Mexico, using Stock Synthesis 2.

The STAR Panel recommends that the current assessments for black rockfish in Washington and California constitute the best available scientific information on the current status of the stock in those areas and that those assessments provide a suitable basis for management decisions. The STAR Panel was unable to endorse the assessment for black rockfish in Oregon as best available science because a base model could not be developed prior to the conclusion of the meeting.

Summary of Data and Assessment Models

Data

Without definitive research to provide justification for stock boundaries, biogeographic information suggests that Cape Mendocino at 40°10' N latitude and the Columbia River plume at 46°16' N latitude would provide justifiable boundaries within the western U.S. area. The Columbia River at the WA/OR border was used for the Washington stock region. For simplification of statistical aggregation, the CA/OR border 42°00' N latitude was used instead of Cape Mendocino. There is a possibility of interaction of the Washington stock (adults, juveniles and larvae) and black rockfish to the north, particularly the population in British Columbia. Each of these points suggests that selection of stock boundaries remains a considerable uncertainty for the assessment of black rockfish, and that further research to improve the justification for stock boundaries is warranted. There appears to be incomplete catch history reconstruction and no formal stock assessment for black rockfish in Canadian waters at this time.

Catch for most rockfish species is uncertain, particularly for the historical period where unspecified rockfish catch needs to be separated by species using assumptions about species ratios. Further work can be done to evaluate catch uncertainty and to provide alternative plausible catch series for sensitivity testing using the assessment model. Formal historical rockfish catch reconstructions have recently been completed in Oregon and California but not for Washington. A first attempt at a black rockfish-specific historical catch reconstruction in Washington was attempted for this assessment. Additional work is still required in each state to better justify most likely catch histories and also to define alternatives that encapsulate major uncertainties for sensitivity testing.

Given the large number of available abundance indices (Table 1) it was noted during the meeting that the Panel was unable to examine each in detail. The Panel was able to endorse standard procedures used and endorsed by the SSC for many of the indices: delta-GLM for individual fishing operations, and Stephens-MacCall filtering of aggregated data by trip or stop followed by a delta-GLM. Given this was common for the other nearshore assessments that shared data and approaches, an improved process would be for a data group to examine and approve input data and methods for standardization prior to stock assessments. A data meeting was carried out for nearshore fishes in March/April 2015 but did not provide endorsement for specific standardization procedures to be used for each abundance index.

Table 1: A summary of the biomass/abundance time series used in each stock assessment.

Region	ID	Fleet	Years	Name	Fishery independent	Filtering	Method	Rank	Endorsed
WA	1	4	1981-2014	Dockside CPUE	No	trip,area, month, Stephens-MacCall	delta-GLM (bin-gamma)	1	SSC
WA	2	5	1986-2013	Tagging CPUE	No	Spring trips, PCA 2	delta-GLM (bin-gamma)	2	SSC
WA	3	5	2000-2013	Tag abundance	No	Spring trips, PCA 2	Petersen	3	No
OR	1	5	2001, 2003-2014	Onboard observer CPFV	No	Positive drifts	delta-GLM (bin-lognormal)		SSC
OR	2	6	2002-2013	Tag abundance	No	None	Mark-recovery		SSC
OR	3	7	1980-2000	MRFSS recreational	No	Stephens-MacCall trip	delta-GLM (bin-gamma)		SSC
OR	4	8	2001-2014	ORBS survey	No	Stephens-MacCall	delta-GLM (bin-gamma)		SSC
OR	5	9	2004-2013	Commercial logbook CPUE	No	Custom criteria	delta-GLM (bin-gamma)		No
CA	1	4	1988-1999	Onboard observer CPFV 88-99	No	Custom filter, Positive drifts	delta-GLM (bin-lognormal)	2	SSC
CA	2	5	2000-2014	Onboard observer CPFV 00-14	No	Custom filter, Positive drifts	delta-GLM (bin-lognormal)	1	SSC
CA	4	6	1980-2003	Dockside-MRFSS CPUE	No	Stephens-MacCall	delta-GLM (bin-gamma)	3	SSC

Length compositions are primarily from the recreational fishery, although some were from the commercial fisheries. Trawl length composition data from landings in Astoria in Oregon were determined to have mostly been taken in Washington, so were removed from the Oregon model and put into the Washington model during the meeting.

Most recreational length data arises from the MRFSS program that does not record sex. Length and age data from the Washington biological database and Oregon ORBS program are the only source that records sex. Washington did not use MRFSS data at all, while the other two states relied on MRFSS data to varying degrees.

MRFSS recreational length composition data are expanded according to year, boat type, bi-monthly period and state. Length or age data from ORBS and commercial fisheries samples for individual strata were combined by expanding by the estimated numbers of fish in that strata to produce weighted average estimates of length or age composition.

There is little evidence in any of the length composition data of modes that may indicate strong cohorts moving through the population from year to year.

Age data have recently become available for the California assessment from Abrams (2014), Lea et al. (1999), and the NWFSC and the SWFSC, providing in the order of 2,000 commercial and recreational age samples from 1979 to 2013, mostly by sex.

Oregon has in the order of 500 age samples per year from 1974 (n=242), and 1992-2013 mostly from the commercial live fish and recreational fisheries

Washington has the most available age data of the three areas, with about 8,000 samples from commercial fisheries primarily from 1980 to 1995, and in the order of 1,000 samples per year from 1984 to 2014. Nearly all age samples are by sex.

For all three areas, age data were almost exclusively modeled as conditional age-at-length.

Assessment Models

The stock assessments were done using SS3 (ver. 3.24v) to produce estimates from the mode of the joint posterior distribution (MPD estimates), which are identical to maximum likelihood estimates when all the priors are uniform. Full Bayesian estimation, where the estimates are taken from the median of marginal posterior distributions, was not performed. The use of only MPD estimates is standard practice in west coast assessments.

The availability and treatment of the input data was similar across all three stocks. Only fishery-dependent biomass indices were available and these were fitted assuming lognormal observation error. Length compositions were fitted assuming multinomial distributions and age data were almost exclusively used as conditional age at length (allowing the estimation of growth within the models).

The population models were very similar across all three stocks. The models were single-area, single-growth morph, and two-sex, with a maximum age of 40 years (with a plus group). Multiple fisheries were modelled for each stock, generally a trawl fishery, a recreational fishery, and two non-trawl fisheries (live-fish and dead; Washington did not have a live-fish fishery). Generally, logistic length-based selectivities were used (except for live-fish fisheries). However, for some fisheries an additional dome-shaped age-selection was specified because of the relative absence of old females in the data. In previous assessments only length-based logistic selection had been used and the absence of old females was explained by ramping up the female natural mortality rate (Sampson 2007, Wallace et al. 2008).

The previous assessments used a ramped "kill 'em" model to explain the absence of old females. An alternative is to "hide 'em" by using age-based domed selection (or a combination of age- and length-based selection to produce derived dome selection) in the fisheries where the old females were missing. Domed-shaped age selection, for females in particular, may be problematic in that a proportion of the old female spawning biomass is protected from exploitation. That is, in the absence of a survey that has asymptotic selection, domed-shape selection can lead to a substantial "cryptic biomass" (it exists but is not available to fisheries or surveys).

The base models included in the initial document provided to the STAR Panel included dome-shaped age-based selectivity for all fleets, and thus had fixed values of natural mortality (*M*) which were derived externally from estimates of maximum age and growth parameters (using methods in Hamel 2015, Then et al. 2014), a common method of estimating *M* on the west coast. The fixed values of *M* ranged from 0.07–0.10 with male values being slightly higher than those for females. In the 2007 assessments, *M* had been fixed at 0.16 for males (of all ages) and "young" females with a ramp going up to 0.24 for "old" females (Sampson 2007, Wallace et al. 2008). The contrasting *M*s between the assessments are due to the shift from a "kill 'em" hypothesis to a "hide 'em" hypothesis. Low *M*s are inconsistent with the age data unless domed selection is invoked. The STAT was concerned that the use of domed selection would prevent the estimation of *M* within the models (because of the confounding of the domed selection and the right-hand-limb of age frequencies). The Panel and CA/WA STAT considered it to be preferable to estimate *M* within the models as there appeared to be adequate age data available (definitely for Washington and Oregon) and logistic selection for the males prevented a confounding with the domed selection for females.

The models all assumed a Beverton-Holt stock recruitment relationship with a fixed steepness (*h*) of 0.773 (the median of the latest rockfish prior; J. Thorson, pers. comm.). Attempts to estimate *h* within the models had, as expected, been unsuccessful with *h* hitting the bound at 1 or likelihood profiles indicated the data contained no information about steepness.

The female maturity at length in the models was set equal to an estimate of "functional" maturity rather than the more usual "sexual" maturity. Functional maturity excludes fish which are sexually mature but did not spawn, and thus gives a better representation of the amount of reproductive output (i.e. larvae) produced. However, during the meeting there was concern that L_{50} was too high (compared to L_{inf}). It was decided to re-estimate the maturity curve after excluding large females (> 46cm) that were deemed to be "immature". These fish were considered to be mature but not spawning in the year they were sampled.

Recruitment deviations were estimated for all models but the earliest cohorts for which deviations were estimated were not observed in the composition data. The purpose of this approach was to allow the uncertainty in the early cohorts to "flow through" into assessment outputs. However, the Panel suggested that a deviation should only be estimated when there was evidence that the data contained information on the cohort. This avoids the situation where the MPD estimates of the early cohorts were all less than 1 because of the assumed lognormal prior. It was suggested that the variance plot, produced to determine the bias corrections, should also be used to decide which deviations should be estimated.

The base models for CA and WA were tuned to balance the relative weights of the different data sets and also to determine a final value for sigma-R (the log standard deviation for the lognormal prior on recruitment deviations). Panel members were concerned that for California, in particular, the resulting sigma-R (at 0.25) was far too low and was unnecessarily restricting the ability of the model to fit the data. It was suggested that anything below 0.5 was unlikely to be appropriate.

For the base models included in the initial document provided to the STAR Panel, effective sample sizes for conditional age-at-length data were not tuned for any of the stocks. For Washington, extra variance was added to the CPUE indices and length frequencies were tuned using Francis (2011). A similar approach was tried for California but the addition of variance to the biomass indices resulted in them being given increasingly high CVs (so high that they were essentially without information). Therefore, in the California base model, the only reweighting was for the length frequencies. For the Oregon model, extra variance was added to indices (unintentionally as it happened), a mistake that was identified and corrected during the STAR Panel review. In the preliminary base model the length frequencies were tuned using the harmonic mean, but were subsequently tuned using the Francis weighting approach. None of the models attempted to tune the input Ns for the conditional age at length compositions.

Standardized procedures for relative weighting within and across different data sources (particularly length and age composition, age at length composition and abundance indices) are currently an area of active research. The approach used in the assessments was consistent with currently accepted methods. There is yet to be full consensus on how to weight data sets and there are particular problems with down-weighting conditional age-at-length data. Because the data are spread across many cells, down-weighting may result in sample sizes within cells that are less than 1. In SS3 this creates a particular problem because all scaled sample sizes less than 1 are set equal to 1 (thus the proportional numbers at age for given length can become severely distorted). A data weighting workshop planned for later this year might provide guidance if work to resolve such issues is currently being done and will be presented.

The Panel requested that reference models be produced for each of the three stocks which included cumulative changes to the base models:

- Use the functional maturity estimate from the modified dataset
- Set sigma-R to 0.5 and do not tune it
- Use the variance plot to determine which recruitment deviations to estimate
- Estimate *M* for females with a small fixed positive offset for males; with the domed age selection for females.

A sensitivity run for each of the reference models, excluding domed-selection and using a ramp for female *M* was also requested. For this run, in California and Washington, the *M* for higher ages was estimated to be lower than the *M* for younger ages. This suggested that a ramp was not needed and a new run was requested, excluding domed selection and simply estimating separate *M*s (constant at age) for males and females.

Other sensitivity runs were also requested by the Panel and also initiated by the STAT Teams. The final outcome of the explorations was a joint decision by the Panel and the Washington / California STAT to move to new base models. A base model was not achieved for the Oregon assessment due to run-time error and errors in the SS3 input files. The main differences between the Washington and California base models brought to the STAR Panel and the final base models were:

- The exclusion of domed-shaped selection except for the live-fish fishery, which was length-based.

- The estimation of sex-specific *M* within the model
- Only estimating recruitment deviations for which the model appeared to contain information
- A fixed value of sigma-R = 0.5 with no tuning of sigma-R.

Treatment of uncertainty

Model uncertainty was addressed by conducting sensitivity analyses, likelihood profiles, and retrospective analyses. A major uncertainty across the three models is a lack of female fish for ages older than ~10 years (as indicated by age-specific sex ratios from catch data), which could be explained either by higher female mortality rates or lower female selectivity (or some combination of these processes). For the Washington and California models, comprehensive sensitivity analyses were conducted prior to the STAR Panel Review that evaluated the removal of data and (independently) the various assumptions regarding the 'extra' process error variance on abundance indices, growth, maturity ogives, fecundity, selectivity, and natural mortality. A total of 30 and 29 sensitivity runs were conducted for the Washington and California models, respectively (not including likelihood profile and restrospective analyses), with additional ones brought to the STAR panel. The California and Washington models were sensitive to the treatment of maturity, selectivity and natural mortality, as ramping of *M* up for females at older ages decreased the stock scale and increased stock productivity. Recruitment deviations were relatively insensitive to alternative configurations of selectivity and natural mortality.

Likelihood profiles for the California and Washington models indicated that the initial recruitment (ln R_0) was well informed, whereas steepness showed a relatively flat profile in the range of 0.8 to 1.0. The likelihood profile of *M* for the pre-STAR Panel Washington base case indicated a well-informed estimate of approximately 0.15, an increase from the fixed sex-specific values of < 0.1 in the base case. For the pre-STAR Panel California model, the likelihood profile also indicates an increase from the fixed base case values, but with a relatively flat likelihood profile from approximately 0.11 to 0.17.

Several sensitivity analyses and likelihood profiles were provided for the Oregon model in the initial assessment report and for a different preliminary base model brought to the STAR Panel Meeting. These models and the uncertainty analyses were rejected by the STAT and the STAR because of run-time errors that were occurring prior to and during the initial portion of the STAR Panel meeting and model configuration errors that were discovered during the STAR Panel meeting.

The final model accepted for the California and Washington models deviated from the assessment models developed for the assessment report, primarily in the modeling of selectivity and sex-specific natural mortality rates. Because of the time spent evaluating various alternatives for selectivity and natural mortality processes, evaluation of uncertainty for the final accepted base model was not as extensive as the pre-STAR model, though several sensitivity runs were provided during the STAR panel. Specifically, likelihood profiles for natural mortality rates and the unfished recruitment level (R_0) would have illuminated the model sensitivity to these important parameters. The amount of age composition data for the California model is less than the other two models, which may have resulted in higher uncertainty on the natural mortality rate.

Some limited evaluations of uncertainty for the California and Washington models were conducted when attempting to identify model runs for the decision table. For example, the models appear to be relatively insensitive to plausible alternatives for the catch history (i.e., ± 50% for the trawl time series). Additionally, the models appear to be relatively insensitive to the steepness, although only steepness values of 0.6 and 0.9 and were considered. A more complete evaluation of uncertainty of the final model should be evaluated and is expected for the final assessment document. In particular, the accepted base model explains the absence of older female fish with sex-specific natural mortality and logistic fishery selectivity (except for the live- fish fishery). It is plausible that that some degree of reduced fishery selectivity exists for older ages, but the sensitivity to various degrees of this reduced selectivity has not been evaluated.

Requests by the STAR Panel and Responses by the STAT

Requests for the Black RF STAT (CA&WA models)

1. For the base WA dockside CPUE index, include an area:year interaction and compare trends across areas graphically (i.e., year effect plots with and without area interactions). Also provide the total removals by year for each area. If possible, statistically test for differences in trend.

Rationale: These runs will partially test the validity of the statewide index.

Response: Most of the samples are from areas 2 and 4 (as well as most of the catch). While reasonably flat overall, the indices from areas 2 and 4 do show some difference in trend over time, with area 2 generally declining and area 4 increasing. This suggests an area:year interaction that can be addressed for the index and also the possibility of population sub-structuring in WA that has implications across all input data sets. A clustering procedure was applied across indices that grouped areas 1/2 and 3/4, also suggesting a N/S structuring. The clustering procedure is a useful way to examine comparable indices that could be applied more widely as a standard diagnostic.

2. Examine the output of the delta-GLM to explain why adding the terms BagLimits and DepthRestrict for the WA dockside models had no effect (the change in AIC = 0) on the AIC values for the gamma and binomial error distributions.

Rationale: See request.

Response: Output factor values for bag limits and depth restriction for this model are completely flat. Standard software has been applied that appears to be producing AIC values that are difficult to understand given the addition of parameters to the model, even though they are apparently not being fitted. Depth restrictions and year are confounded, so such a structure is questionable.

3. Provide a general overview of all indices and their attributes (see Neil's summary table in the Requests subfolder).

Rationale: This would be a useful overview of indices used in the base case models.

Response: This was done (Table 1) and the Panel agreed that it provides a useful overview of all indices. Such a table should be considered as part of the standard assessment document template.

4. For all CA CPUE indices, rather than stepwise selection, perform model selection based on an *a priori* set of models. Comment on the timeframe to do this. If this can be done within 36 hrs., then please do this. Explain how the "best" models were chosen in the final document.

Rationale: Best practices recommendation.

Response: None.

5. Provide recreational mean lengths and length comp. sample sizes by area and year in the WA model.

Rationale: Indices need to be informed by consistent annual sampling strategies.

Response: These were provided as a plot and show consistent sampling for areas 2 and 4 with most of the sampling and catches. They also show little change in mean length through time in those areas.

6. For the estimation of functional maturity, provide fits to data and if there are large length classes that are under the maximum maturity, remove them, refit, and compare the functional maturity curves.

Rationale: To validate the pre-STAR functional maturity ogive.

Response: There were outliers with low maturity for large fish (potentially due to "warm blob" being an unrepresentative year with high atresia) that appeared to have a disproportionate influence in forcing asymptotic maturity to the right that was alleviated by removing those data. The asymptote was still missing the data at the top of the curve suggesting that the fit would be improved if fitted to log lengths. However, newly provided proportions were fitted well by the existing model, so all that was required was the removal of low maturity fish >45cm for an agreed functional maturity for the base case.

7. Provide male age frequencies and *Z* estimates for a fishery with logistic selectivity (i.e., do a catch curve analysis of marginal male ages for a fishery with logistic selectivity).

Rationale: This may help evaluate plausibly high estimates of *M* in the model.

Response: Available *Z* estimates were generally less than about 0.3 and in five years near 0.17 , providing an additional guide for an upper bound for *M*.

Requests 8-10 are cumulative changes proposed to create a new reference case. Requests 11-14 would use the new reference case.

8. Fit the logistic functional maturity curve using log lengths to fit the maturity samples after removing the zeroes >45 cm for new reference case. If log lengths do not work, fit using a LOESS.

Rationale: To provide a better fit to the proportion positive mature.

Response: It turned out that the proportion mature at length had been incorrectly calculated and the logistic fit was actually much better than had been presented. The removal of the zeroes shifted the maturity curve to the left by about 1cm and showed a good fit to the proportions mature. The new curve was accepted for use in the reference runs.

9. Set sigma-R to 0.5 for CA and WA models. Use the bias adjustment curve to decide when to estimate early and late recruitment deviations for the new reference case.

Rationale: If sigma-R is tuned down below about 0.5 it is likely to unnecessarily restrict the recruitment deviations and the overall fit to the data. Estimating recruitment deviations too early may give the model too much freedom to adjust in a period when there is no information.

Response: The variance plots indicated that there was little or no information for many of the early recruitment deviations and these were not estimated in subsequent runs.

10. For CA and WA models, estimate female *M* with a small male offset from growth and/or longevity-based estimates of *M* including dome-shaped female selectivities and the Hamel prior on *M*. Tune as before except for sigma-r.

Rationale: Good practice to estimate *M* if there are data that are informative. Logistic selectivity for males aids in estimating *M* despite dome-shaped selectivity for females.

Response: The run produced estimates of *M* that were far higher than the fixed values that had come from the growth-based estimation procedure (0.13-0.16 compared to 0.07-0.12). However, the estimates were not inconsistent with the Hamel prior for rockfish. When male *M* was freely estimated it came in at about the same value as when the small fixed offset was used (despite the confounding issue). For WA, there was a poor fit to some composition data. In particular, the female length composition for trawl was poorly fit because the model could not produce enough large females due to the low L_{inf} of the growth curve. The STAT was asked to investigate this problem either by down-weighting the conditional age at length data (which was constraining the growth curve) and/or using an informed prior on L_{inf} for the females.

11. Provide a tuned *M* ramp version of the model with logistic selectivities (except for the CA live-fish fishery) as a sensitivity.

Rationale: Provides an alternative to explain missing old females.

Response: For WA, the estimated *M* for the older females was lower than the *M* for the younger females. This result suggested than a ramp was not required and that constant *M*s at age could be adequate. The STAT was requested to produce another "kill 'em" run for WA where separate *M*s were estimated for males and females (with no ramp). The initial examination of the diagnostics was very encouraging with good fits to the main index and the composition data.

12. As a sensitivity, add variance to the following CA index points if needed: 1998 onboard CPUE value, 1997 dockside CPUE value if these points are still outliers.

Rationale: To investigate whether the tension created in the model by these outliers is important to the model results.

Response: The model results were not substantially different when the outliers were down-weighted.

13. Potential development of a cryptic female spawning output estimate from an average age selectivity weighted by catch to female numbers at age and functional maturity. Step 1: Female numbers at age matrix by year; step 2: calculate catch by fleet by year; step 3: create an average derived age-based selectivity across fleets and years using a weighted average catch by fleet by year; step 4: multiply the numbers at age matrix by the average selectivity curves to derive exploitable numbers at age; step 5: generate spawning output estimates from exploitable numbers at age; step 6: subtract the results of step 5 from the spawning output from the assessment to determine cryptic spawning output.

Rationale: This may help determine the best way forward (i.e., estimating *M* vs. dome-shaped selectivity) for a base case.

Response: An R-script was developed to perform the calculations from the standard SS3 output. This was applied to the Washington and California reference models and it was seen that cryptic spawning output was at very high levels (Figure 1). This diagnostic was influential in the decision by the STAT to use the "kill 'em" assumption in the base model.

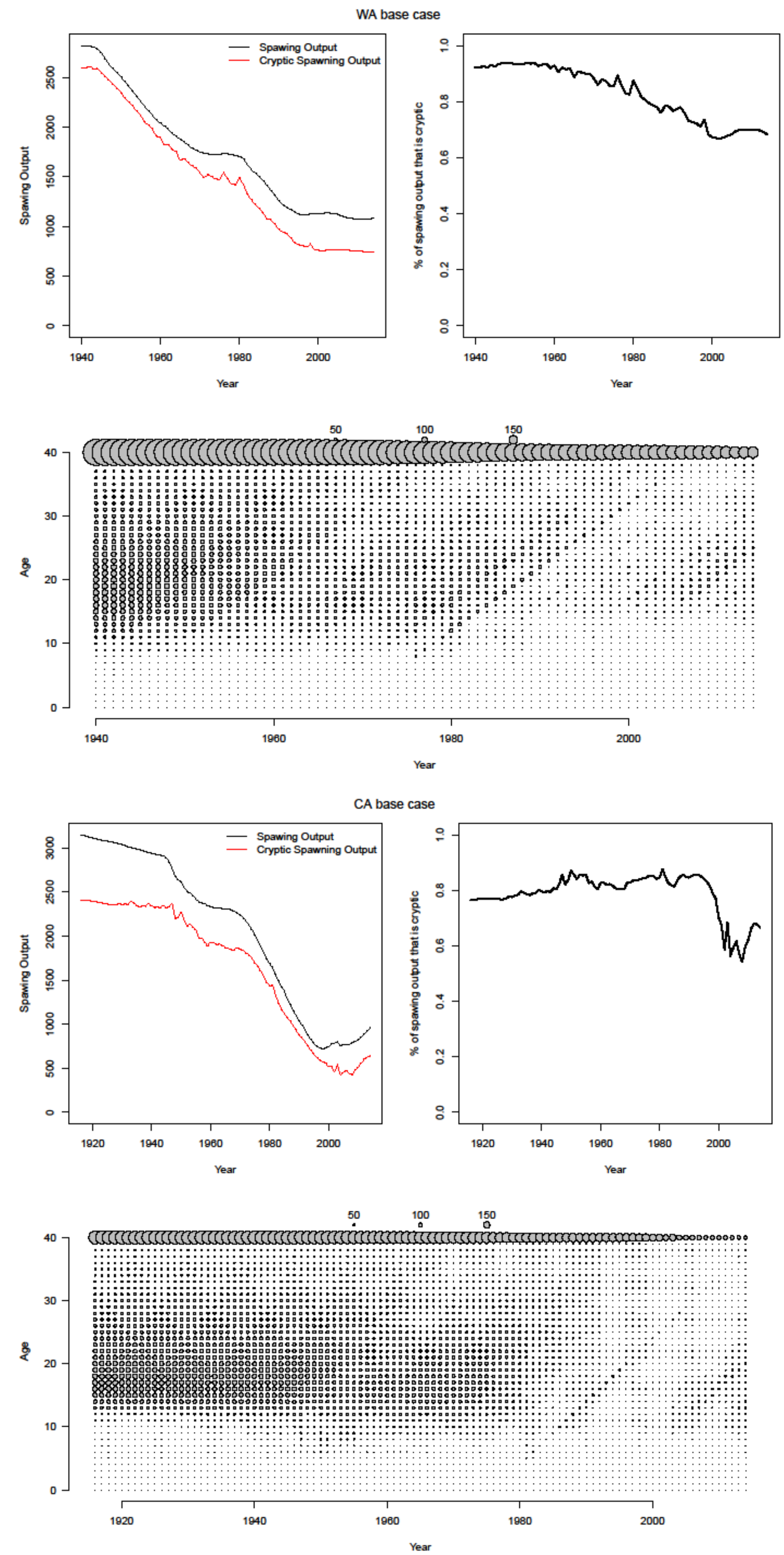


Figure 1: The spawning output trajectories, including the percentage of hidden or cryptic output, for the Washington (WA) and California (CA) base models in the draft SAFE document.

14. Provide the year effects for the binomial error distribution and positives for the two components and the combined for the WA dockside index.

Rationale: Best practice request but a lower priority.

Response: It proved difficult to do this with the available software, but some results were briefly produced that showed that the binomial component of the combined index had no influence on the overall trend.

Day 3 Requests for the Black RF STAT (OR model)

The STAR Panel did not receive a presentation on the OR model until the afternoon of the 3rd day of the Panel meeting. All changes to the agenda regarding the presentation of the OR model were approved by the STAT lead for the Oregon model because it gave him additional time to debug his model.

Requests 2-8 are cumulative changes proposed to create a new reference case.

1. Determine the amount of composition data collected in Astoria used in the OR model.

Rationale: To examine the potential level of data contamination.

Response: 86% of the OR trawl composition data came from Astoria.

2. Remove the Astoria trawl composition data from the OR model and migrate the data to the WA model.

Rationale: The trawl catches from Astoria were correctly in the Washington model, but the composition data had been left in the Oregon model.

Response: Removal of this data from the Oregon model left little data with which to estimate selection for the trawl fleet, but the trawl catches in Oregon are minimal and the estimated trawl selection curves were not implausible. The addition of the Astoria data had little effect on the Washington model results.

3. Avoid double use of the composition data from the Small-fish sampling.

Rationale: Best practice.

Response: This was done.

4. Correct the input N for the ORBS marginal age frequencies and, if necessary, scale the input Ns to be no more than 50.

Rationale: To prevent undue influence of these comps.

Response: The correction was made and the samples sizes were all larger than 50. Rather than scaling down the numbers so that the maximum was 50 they were all set to 50.

5. Use the newly fitted logistic functional maturity curve.

Rationale: This was the preferred maturity curve.

Response: Done.

6. Don't change sigma-R from 0.5. Use the bias adjustment curve to decide when to estimate early and late recruitment deviations for the new reference case.

Rationale: As for the Washington and California models.

Response: The original Oregon model presented in the draft assessment document had sigma-R fixed at 0.5 and subsequent models never attempted using any other value.

7. Estimate female *M* with a small male offset from growth-based estimates of *M* or use an estimated offset including dome-shaped female selectivities (asymptotic for males in the trawl fishery) and the new Hamel prior on *M*. Tune length comps. with Francis weighting; don't tune sigma-R or conditional age-at-length.

Rationale: To conform with the Washington and California reference models.

Response: Done.

8. Provide a sensitivity to the reference run with logistic selectivities, no *M* ramp, and estimate male and female *M*s separately.

Rationale: To conform with the new Washington and California base models. (Kill 'em, don't hide 'em).

Response: Done. The reference model and the sensitivity were produced and results were shown to the meeting. However, exploration of the Synthesis Results.sso file on the afternoon of Day 4 of the STAR uncovered an error in how the bounds were specified for the Tag-Q parameter, which meant that the results were only useful for diagnosing the errors rather than informing on the suitability of potential base models. There was insufficient time to develop a base model.

9. Normalize time series of indices and plot together.

Rationale: To compare trends.

Response: Done. All indices showed similar trends except for the Nearshore Commercial logbook index.

10. Provide the year:area effect in indices.

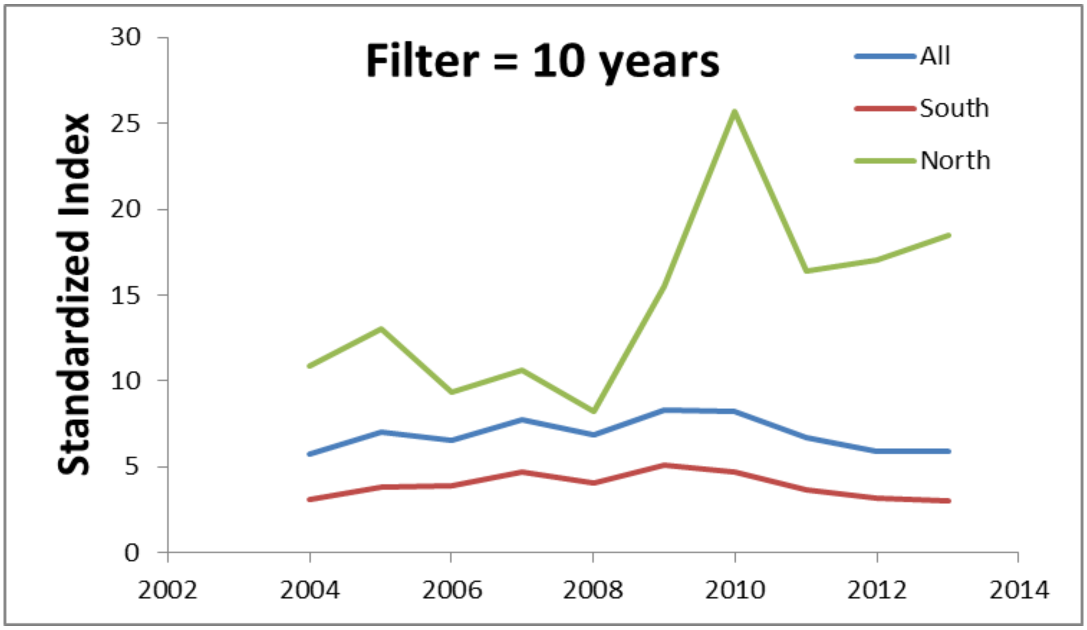
Rationale: To determine whether there are different CPUE trends in different areas (e.g., N&S or by port).

Partial Response: The Nearshore Commercial logbook index is being driven by trends in the south. This raised a new request to investigate the data filtering on vessels retained in the index.

11. Apply a different filtering of the number of vessels retained in the Nearshore Commercial logbook index (vessels that were active (at least 1 trip) in at least three years) to retain much more data.

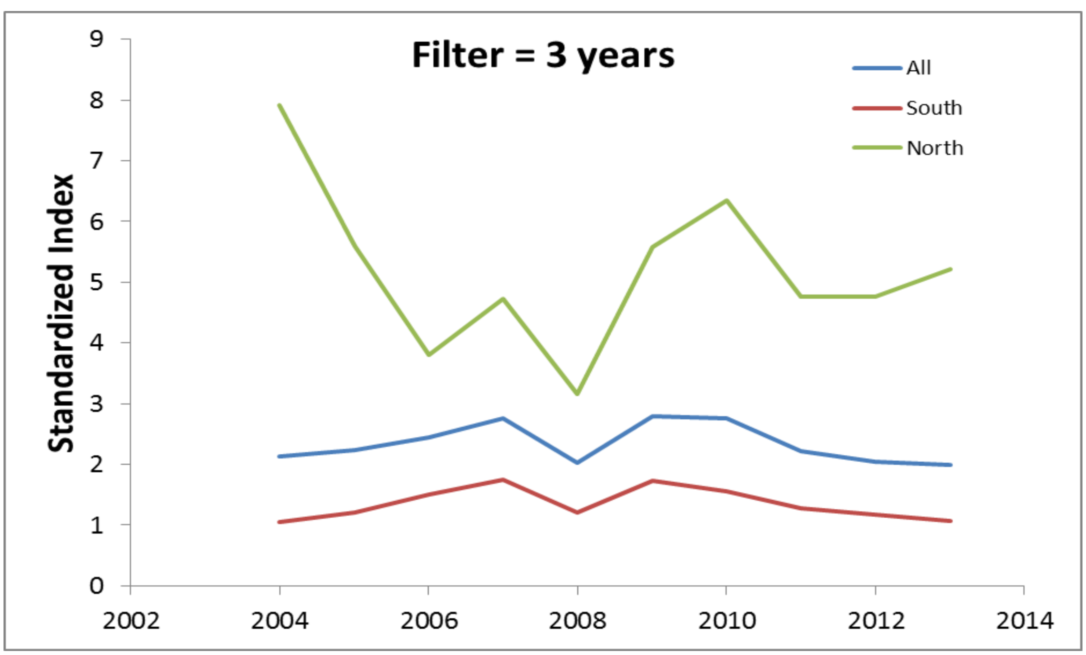
Rationale: To determine if the new filtering shows a dramatically different index trend.

Response: The original vessel filter, in conjunction with other "missing/bad data" filters, only retained 18% of the total records and catch. When the less restrictive vessel filter was applied 90% of the catch was retained. The overall trend was not substantially changed as it was driven by the trend in the south. The northern trend, which was very different from the southern trend, was dramatically altered when all of the extra data were included (Figure 2). The relative weighting of the two trends



within the CPUE model (16% north, 84% south) was different from the weighting implied by the proportions of habitat in the north and the south (30% north and 70% south). The habitat weighting, with the new vessel filter, should be preferred when constructing the overall CPUE time series but it may not make much difference to the stock assessment results.

Figure 2: The Oregon commercial logbook CPUE indices for a vessel filter requiring participation in every year (10 years, top) and a filter requiring only 3 years participation (bottom). The indices are shown for the north, south, and the overall trend.



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

A base model was not achieved for Oregon. The description below applies to both the Washington and California models.

The models for CA and WA were single stock, area, and growth-morph, with two-sexes and ages 1–40 years (with a plus group). Multiple fisheries were modelled including a trawl fishery, live-fish and dead-fish hook and line fisheries, and a recreational fishery. Discards were estimated to be low and were added to the catches (and not explicitly modelled). The biomass indices and composition data as described above were used to produce MPD estimates.

The key points of the base model structure were:

- Logistic length-based selection except for the live-fish fishery, which was length-based and dome-shaped.
- The estimation of *M* within the model (constant at age and separate for males and females).
- Only estimating recruitment deviations for which the model appeared to contain information.
- A fixed value of sigma-R = 0.5 with no tuning of sigma-R.
- Tuning of the effective sample sizes for length frequencies using Francis (2011)

The base model was bracketed by low-*M* and a high-*M* runs where 0.03 was subtracted and added to the MPD estimate for each stock.

Technical Merits of the Assessment

The main merits of the Washington and California assessments are:

- The use of SS3 which is a well-tested package that allows the integration of multiple data sources in a well-understood estimation procedure.
- Multiple sources of indices and composition data are used.
- There are substantial age data in the Washington assessment which supports the estimation of *M* within the model.
- The use of "functional maturity" rather than "sexual maturity" was used to determine spawning output.
- The relative absence of older females was explored through two alternative hypotheses (hide 'em or kill 'em). This was an improvement on past assessments that had automatically used a kill 'em model (with the unrealistic and over-parameterized ramp on female M).

Technical Deficiencies of the Assessment

The following deficiencies apply to all three stock assessments:

- No alternative catch histories were supplied despite large uncertainty in historical catches.
- Improved data preparation is necessary. It is noted that standard West Coast practice was adopted in this regard. However, best practice requires a thorough and careful analysis of available composition data, including conditional age-at-length data, to determine whether post-stratification and scaling are required. The objective is to use all available composition data that can be formed into consistent time series. This may require that some data are excluded from a time series when sampling was inadequate (e.g., spatially and temporally).
- No fishery-independent biomass indices were available.

- With only one exception, the diagnostics supplied for CPUE indices were inadequate (see recommendations).
- The SS3 input files are unnecessarily complicated (e.g. see CASAL input formatting for a less complicated approach) and require further automated QA/QC safeguards to detect errors.
- The base models brought to the STAR Panel were modified extensively during the meeting. This substantially reduced the time available to explore sensitivities to the new base models.

The California assessment is marginal in the amount of age data that is supporting the estimation of *M* (it may be better to borrow an *M* estimate from one of the other stocks given they have vastly more age data).

Several problems were detected for the Oregon assessment during the meeting:

- A standard deviation of 1 was inadvertently added to the standard deviations of the biomass indices (because the STAT mistakenly thought that the variance adjustments in the control file for the indices were multiplicative).
- The composition data from the trawl catches landed in Astoria were used (the catches had been transferred to Washington but not the composition data).
- The informed prior on the tag *q* was incorrectly specified (the bounds and lognormal prior were appropriate for *q* but wrong given that log(*q*) was the parameter being estimated).

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:
The STAT lead for the Oregon model (Dr. David Sampson) takes great exception to the model configuration requested by the STAR Panel as a “reference model” for all three states. During the STAR panel Dr. Sampson presented formal arguments based on ecological theory to the STAR Panel registering his opposition to the reference model configuration. Dr. Sampson will be providing a thorough discussion of his arguments as Agenda Item H.3, Attachment 7, September 2015

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

The GMT representative stressed that the proxy SPR harvest rates used to determine F_{MSY} should be re-evaluated given the evolution of the steepness prior used in assessments.

Unresolved Problems and Major Uncertainties

The bullet points below apply to the Washington and California assessments. There are unresolved problems for the Oregon assessment but hopefully that assessment can benefit from borrowing some of the model structure that worked successfully for Washington and California. The Panel suggests that the model structure for Oregon should be similar to Washington and California unless data peculiar to Oregon dictate otherwise.

The unresolved problems are:

- The complexity of SS3 input files makes it difficult to detect errors that may still reside in the input files. None were specifically suggested or thought to occur in the Washington and California models, but this remains unknown.

- Standard practices for data preparation need further improvement. The CPUE indices may contain spatial trends that require re-weighting using habitat based weights. The composition data may require post stratification and scaling and the removal of data in years when sampling was inadequate.

The major uncertainties are:

- The level of cryptic biomass is unknown. The base model has assumed that there is none but this is unlikely to be absolutely true, although there has been considerable fishing at most depths and habitat types coastwide that has not apparently located a concentration of old female fish. It is unlikely that the alternative hide ‘em model represents reality either, but some level of domedness in selection is to be expected in some of the fisheries (especially trawl where large fish may be unavailable due to habitat preference, or able to escape).
- Historical catch history is very uncertain. Sensitivity to this was explored only for plus/minus 50% on the trawl catches. The results were not sensitive in that case but could be sensitive to different trends in the historical catch.
- Natural mortality may be poorly determined, especially for California.
- The stock recruitment relationship is unknown.

Recommendations for Future Research and Data Collection

Continued research on:

- Stock structure, including whether national/international boundaries are appropriate.
- Catch histories, including uncertainty and alternative catch streams to be used in sensitivity analysis.
- Definition and measurement of black rockfish habitat
- Preparation of composition data: post-stratification and scaling supported by a detailed analysis of the data

A specific data workshop, perhaps for all species prioritized for assessment, could examine information across a broad range of species due for assessment, and would also assist with the development of more specific documentation of protocols used to compile best available data sets for stock assessment, continue acceptance of agreed procedures for standardization of abundance indices, and also begin work on procedures for the development of alternative data series that capture uncertainty – particularly for historical catch and discards.

The outline for stock assessments (Appendix B in the 2014 Terms of Reference) includes a section for addressing previous STAR Panel recommendations. If a data workshop precedes the stock assessment, as here for black rockfish, the outline should also include a section on how the recommendations from the data workshop were addressed. A similar process should be outlined to address recommendations from previous CIE reviewers.

Consider the development of a coastwide fishery-independent survey for nearshore stocks. As the current base model structure has no direct fishery-independent measure of recent rebuilding of the adult portion of the stock, any work to commence collection of such a measure for nearshore rockfish, or use of existing data to derive such an index would greatly assist with this assessment.

For abundance indices a multi-species simulation study to test whether the Stephens-MacCall filtering may lead to a bias in abundance estimates given differences in abundance trends among species should be considered. It is the understanding of the panel that some simulation testing has been done; these results should be made generally available. A comparison of alternative filtering procedures should also be considered.

CPUE standardization protocols need improvement:

- An objective procedure for sub-model error structure (usually gamma or lognormal here) is required for delta-GLM procedures. Consistency is required for the model selection process – preferably using *a priori* candidate models rather than a stepwise selection. For Washington and California: removal of the restriction on having the same explanatory variables for the binomial and positive catch rate models. There is no reason why the presence/absence of the species should be explained by the same variables which explain the magnitude of the positive catch rates.
- The AIC tables do not appear to report the correct AIC values in some cases. In particular, models with 1 or 2 additional variables, which were likely confounded with other variables, often had the exact same AIC value as the less complex model. While this is technically possible, it is highly unlikely and its consistency is worrying.
- Better diagnostics for each CPUE analysis: plots of the binomial and positive catch rate year effects in addition to the combined year effects; plots of all estimated effects; production of year:area interactions and a comparison of the trends by area. In cases where the trend in CPUE index differs across areas, the aggregate CPUE index is affected by the method used to weight the CPUE from the areas. Evaluation of the effect of alternative weighting methods on the aggregate standardized CPUE index should be evaluated in these cases.
- The effects of the standardization on the "nominal" or unstandardized indices should also be shown and explained (i.e., which variables have caused a shift in the trend).

SS3

- The input interface is not user-friendly and requires considerable knowledge of formatting requirements and the meaning of some settings in relation to how the model is configured or parameterized.
- The addition of extra standard deviation for biomass indices should be correctly implemented. Standard deviations do not add arithmetically. To apply a constant process error to a time series of biomass indices requires that the variances be added (i.e., square the standard deviations, add them together, and take the square root).
- The fact that some priors are set in normal space and others are set in log space creates confusion on inputting these priors.
- Many of the problems could be solved by creating an "expert system" front end which creates the input files exactly how SS3 needs them (i.e., no erroneous white space or unprintable characters) transforms parameters from arithmetic space as needed, and checks for obvious user errors (e.g., a row of 1s being added to the standard deviations).
- Improved debugging tools are necessary to help track down sources of errors messages such as "-1.INDs"

R4SS

- The plots showing the fits to the indices could also include useful information such as the estimated q and whether extra SD was added to the input SD (users look at the plots but they may not look at the report file). A plot to assist with comparison of all abundance indices where all indices are plotted with available biomass determined by selectivity (such a plot was developed for China rockfish during a recent STAR Panel).
- Calculation of unexploitable spawning output: the procedure developed here should be considered as a standard diagnostic for all assessments.

- The units of spawning output should be shown in the r4ss plots, as confusion was caused when spawning output is defined as egg production that may not necessarily be proportional to spawning stock biomass.
- Assistance with appropriate levels for jittering: Jason Cope has used a procedure for setting appropriate ranges for jittering that might be considered as an objective method for standard practice

Acknowledgements

The STAR Panel thanks the public attendees, Maggie Somers, STAT, GMT and GAP representatives.

References

Abrams, J.L. 2014. The effect of local fishing pressure on the size and age structure of fishes associated with rocky habitats along California's north coast. Ph.D dissertation, Humboldt State University, Humboldt.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1124–1138.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life-history correlates. ICES Journal of Marine Science. doi:10.1093/icesjms/fsu131.

Lea, R. N., R. D. McAllister, and D. A. VenTresca. 1999. Biological aspects of nearshore rockfishes of the genus *Sebastes* from central California. California Department of Fish and Game. Fish Bulletin 177:1–109.

Sampson, D.B. 2007. The Status of Black Rockfish off Oregon and California in 2007. Pacific Fishery Management Council, Portland, OR.

Then, A.Y., J.M. Hoenig, N.G. Hall, and D.A. Hewitt. 2014. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72:82-92. doi: 10.1093/icesjms/fsu136.

Wallace, F.R., Y.W. Cheng, and T.S. Tsou, 2008. Status of the Black Rockfish Resource north of Cape Falcon, Oregon to the U.S.-Canadian border in 2006. In Pacific Fishery Management Council. 2009. Appendix: Status of the Pacific Coast Groundfish Fishery Through 2007 and Recommended Biological Catches for 2009: Stock Assessment and Fishery Evaluation. Pacific Fishery Management Council, Portland, Oregon.