RESEARCH RECOMMENDATIONS FROM RECENT STOCK ASSESSMENTS
AND STAR PANEL REPORTS

Bocaccio Research Recommendations (2015 Assessment)

- Stock structure and stock boundaries for Bocaccio rockfish on the West Coast remains an important issue to consider with respect to both future assessments and future management actions.
- Since large scale area closures and other management actions were initiated in 2001, the spatial distributions of fishing effort (fishing mortality) have changed over both large and small spatial scales. This confounds the interpretation of survey indices for surveys that do not sample in the Cowcod Conservation Areas (CCAs), although the decision to begin sampling for the NWFSC hook and line survey within the CCAs should begin to address this issue with time. .
- Recently updated reproductive biology data (maturity and fecundity) show some differences in length and weight specific fecundity in Bocaccio from those used in the past assessments. Regional differences (southern and northern California, as well as southern Oregon), and multiple brood spawning, are poorly understood.
- As Bocaccio is one of the most abundant and important piscivorous rockfish species, and its interactions with other predator and prey species are poorly known, information regarding diet and movement patterns associated with habitat and prey abundance are key in order to further understand its roles in the ecosystem of the California waters. Northward migratory behaviors of juvenile and young adults are indicated by length frequency data, but such behaviors are also poorly understood. Studies on these behaviors and their associations with oceanographic or other ecological factors can help future assessments in defining stock structure as well as explaining high variability in stock recruitments.

STAR Panel Recommendations (Bocaccio 2015)

- An objective procedure for evaluating the stock boundaries is needed for all rockfish (and potentially other west coast assessments). Such a procedure would more directly point to directions for future research or collaboration across national/international political boundaries.
- Explore better ways to model productivity for stocks like bocaccio that exhibit large episodic recruitment patterns. Lognormal distributions are not a good way to model the recruitment variability for such stocks.
- The strength of recent recruitments is a major uncertainty for bocaccio. Technical methods for capturing and propagating this uncertainty are needed in stock synthesis (especially for axes of uncertainty), perhaps by an improved procedure to fix particular recent recruitment deviations.

- The relationship between stock size and spawning output is critical for interpretation of the CalCOFI index, which is perhaps the most useful index in the bocaccio assessment. Research is needed to better quantify spawning output. This research could include evaluation of environmental correlations of spawning output, and studies of both the prevalence, and the potential demographic and environmental drivers of multiple broods (multiple spawning events by an individual fish within a given spawning season).
- The Panel recommends continued processing of historical CalCOFI samples from northern transects in the early 1950s through the late 1960s. These data would add to the index used in the assessment model, and improve understanding of spatial patterns in population dynamics.
- A data workshop prior to STAR panel review, perhaps for all rockfish stocks due for assessment, should be scheduled to examine assessment information across a broad range of species. The workshop could document protocols used to compile data sets for stock assessment, establish agreed procedures for standardization of abundance indices, and develop alternative data series that capture uncertainty–particularly for historical catch and discards.
- Several estimated selectivity patterns in the bocaccio assessment are very unusual. The NWFSC trawl survey has a curiously flat selection pattern at young ages, and triennial survey has a strongly peaked selectivity at young ages. Research into alternative ways to model the selection pattern of these surveys is needed. Possible approaches include 1) use of age-specific natural mortality, 2) splitting the surveys into separate indices for juveniles (age 0 and/or1) and older fish.
- Available information indicates that the CCAs are a center of abundance for bocaccio. Surveying inside the CCA during the NMFSC hook and line surveys should be continued, though several years of data will be required before the information can be used to inform the assessment. Consideration should also be given to extending the NWFSC trawl survey into the CCAs. A simple analysis of potential catch rates of cowcod, and the impact of survey take on stock rebuilding, would allow the benefits of surveying inside CCA to be compared to potential costs.
- Age data from the NWFSC hook and line survey would increase the utility of the survey for assessment of bocaccio by better defining the selectivity pattern for large fish.

China Rockfish Research Recommendations (2015 Assessment)

- The number of hours fished in Washington should be recorded for each dockside sample (vessel) so that future CPUE can be measured as angler hours rather than just number of anglers per trip. This will allow for a more accurate calculation of effort.
- The number of hours fished in Oregon should be recorded for each dockside sample (vessel), instead of the start and end times of the entire trip. This will allow for a more accurate calculation of effort.
- Compare the habitat-based methods used to subset data for the onboard observer indices to Stephens-MacCall and other filtering methods.
- Explore the sensitivity of Stephens-MacCall when the target species is "rare" or not common encountered in the data samples.
- A standardized fishery independent survey sampling nearshore rockfish in all three states would provide a more reliable index of abundance than the indices developed from catch

rates in recreational and commercial fisheries. However, information value of such surveys would depend on the consistency in methods over time and space and would require many years of sampling before an informative index could be obtained.

- A coastwide evaluation of genetic structure of China rockfish is a research priority. Genetic samples should be collected at sites spaced regularly along the coast throughout the range of the species to estimate genetic differences at multiple spatial scales (i.e., isolation by distance).

- Difficulties were encountered when attempting to reconstruct historical recreational catches at smaller spatial scales, and in distinguishing between landings from the private and charter vessels. Improved methods are needed to allocate reconstructed recreational catches to sub-state regions within each fishing mode.

- There was insufficient time during the STAR Panel review to fully review the abundance indices used in the China rockfish assessments. Consideration should be given to scheduling a data workshop prior to STAR Panel review for review of assessment input data and standardization procedures for indices, potentially for all species scheduled for assessment. The nearshore data workshop, held earlier this year, was a step in this direction, but that meeting did not deal with the modeling part of index development.

- The Marine Recreational Fisheries Statistics Survey (MRFSS) index in Oregon was excluded from the assessment model because it was learned that multiple intercept interviews were done for a single trip. Evaluate whether database manipulations or some other approach can resolve this issue and allow these data to be used in the assessment.

- Many of the indices used in the China rockfish assessment model used the Stephens-MacCall (2004) approach to subset the CPUE data. Research is need to evaluate the performance of the method when there are changes in management restrictions and in relative abundance of different species. Examination of the characteristics of trips retained/removed should be a routine part of index standardization, such as an evaluation of whether there are time trends in the proportion of discarded trips.

- Fishery-dependent CPUE indices are likely to be the only trend information for many nearshore species for the foreseeable future. Indices from a multi-species hook-and-line fishery may be influenced by regulatory changes, such as bag limits, and by interactions with other species (e.g., black rockfish) due to hook competition. It may be possible to address many of these concerns if a multi-species approach is used to develop the indices, allowing potential interactions and common forcing to be evaluated.

- Consider the development of a fishery-independent survey for nearshore stocks. As the current base model structure has no direct fishery-independent measure of stock trends, 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.

- Basic life history research may help to resolve assessment uncertainties regarding appropriate values for natural mortality and steepness.

- Examine length composition data of discarded fish from recreational onboard observer programs in California and Oregon. Consider modeling discarded catch using selectivity and retention functions in Stock Synthesis rather than combining retained and discarded catch and assuming they have identical size compositions. Another option would be to model discarded recreational catch as a separate fleet, similar to the way commercial discards were treated in the southern model.

- Ageing data were influential in the China rockfish stock assessments. Collection and ageing of China rockfish otoliths should continue. Samples from younger fish not typically selected by the fishery are needed to better define the growth curve.
- Consider evaluating depletion estimators of abundance using within season CPUE indices. This approach would require information on total removals on a reef-by-reef basis.
- The extensive use of habitat information in index development is a strength of the China rockfish assessment. Consideration should be given to how to further incorporate habitat data into the assessment of nearshore species. The most immediate need seems to be to increase the resolution of habitat maps for waters off Oregon and Washington, and standardization of habitat data format among states.
- Although all the current models for China rockfish estimated implausibly large recruitment deviations when allowed to do so, particularly early in the modeled time period, further exploration of available options in stock synthesis could produce acceptable re-sults. In addition, this work may provide guidance on any additional options that could be added to stock synthesis to better handle this situation. For example, assuming different levels autocorrelation in the stock-recruit relationship for data-moderate stocks may help curb the tendency to estimate extreme recruitment with sparse datasets.
- Research is needed on data-weighting methods in stock assessments. In particular, a standard approach for conditional age-at-length data is needed. The Center for the Advancement of Population Assessment Methodology (CAPAM) data weighting workshop, scheduled for later this year, should make important progress on this research need.

STAR Panel Recommendations (China Rockfish 2015)

- A coastwide evaluation of genetic structure of China rockfish is a research priority. Genetic samples should be collected at sites spaced regularly along the coast throughout the range of the species to estimate genetic differences at multiple spatial scales (i.e., isolation by distance).
- Difficulties were encountered when attempting to reconstruct historical recreational catches at smaller spatial scales, and in distinguishing between landings from the private and charter vessels. Improved methods are needed to allocate reconstructed recreational catches to sub-state regions within each fishing mode.
- There was insufficient time during the STAR Panel review to fully review the abundance indices used in the China rockfish assessments. Consideration should be given to scheduling a data workshop prior to STAR Panel review for review of assessment input data and standardization procedures for indices, potentially for all species scheduled for assessment. The nearshore data workshop, held earlier this year, was a step in this direction, but that meeting did not deal with the modeling part of index development.
- The Marine Recreational Fisheries Statistics Survey (MRFSS) index in Oregon was excluded because it was learned that multiple intercept interviews were done for a single trip. Evaluate whether database manipulations or some other approach can resolve this issue and allow these data to be used in the assessment.
- Many of the indices used in the China rockfish assessment model used the Stephens and MacCall (2004) approach to subset the CPUE data. Research is need to evaluate the performance of the method when there are changes in management restrictions and in

relative abundance of different species. Examination of the characteristics of trips retained/removed should be a routine part of index standardization, such as an evaluation of whether there are time trends in the proportion of discarded trips.

- Fishery-dependent CPUE indices are likely to be the only trend information for many nearshore species for the foreseeable future. Indices from a multi-species hook and line fishery may be influenced by regulatory changes, such as bag limits, and by interactions with other species (e.g. black rockfish) due to hook competition. It may be possible to address many of these concerns if a multi-species approach is used to develop the indices, allowing potential interactions and common forcing to be evaluated.
- Consider the development of a fishery-independent survey for nearshore stocks. As the current base model structure has no direct fishery-independent measure of stock trends, 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.
- Basic life history research may help to resolve assessment uncertainties regarding appropriate values for natural mortality and steepness.
- Ageing data were influential in the China rockfish stock assessments. Collection and ageing of China rockfish otoliths should continue. Samples from younger fish not typically selected by the fishery are needed to better define the growth curve.
- Consider evaluating depletion estimators of abundance using within season CPUE indices. This approach would require information on total removals on a reef-by-reef basis.
- The extensive use of habitat information in index development is a strength of the China rockfish assessment. Consideration should be given to how to further incorporate habitat data into the assessment of nearshore species. The most immediate need seems to be to increase the resolution of habitat maps for waters off Oregon and Washington, and standardization of habitat data format among states.
- Although all the current models for China rockfish estimated implausibly large recruitment deviations when allowed to do so, particularly early in the modeled time period, further exploration of available options in stock synthesis could produce acceptable results. In addition, this work may provide guidance on any additional options that could be added to stock synthesis to better handle this situation. For example, assuming different levels autocorrelation in the stock-recruit relationship for data-moderate stocks may help curb the tendency to estimate extreme recruitment with sparse datasets.
- The China rockfish models made a number of simplifying assumptions, such as asymptotic fishery selectivity, and no deviations from the stock-recruit curve. It would be worthwhile to conduct a simulation-estimation exercise to evaluate potential errors associated with the assumptions commonly made for data-moderate assessments that use length and age data.
- Research is needed on data-weighting methods in stock assessments. In particular, a standard approach for conditional age-at-length data is needed. The Center for the Advancement of Population Assessment Methodology (CAPAM) data weighting workshop, scheduled for later this year, should make important progress on this research need.

Black Rockfish Research Recommendations (2015 Assessment)

- Further investigation into the movement and behavior of older (> age 10) females to reconcile their absence in fisheries data. If the females are currently inaccessible to fishing gear, can we find where they are?
- Appropriate natural mortality values for females and males. This will help resolve the extent to which dome-shaped age-based selectivity may be occurring for each.
- All states needed improved historical catch reconstructions. The trawl fishery catches in particular need particular attention. Given the huge historical removals of that fleet in each state, the assessment is very sensitive to the assumed functional form of selectivity. A synoptic catch reconstruction is recommended, where states work together to resolve cross-state catch issues as well as standardize the approach to catch recommendations.
- Identifying stanzas or periods of uncertainty in the historical catch series will aid in the exploration of catch uncertainty in future assessment sensitivity runs.
- The ODFW tagging study off Newport should be continued and expanded to other areas. To provide better prior information on the spatial distribution of the black rockfish stock, further work should be conducted to map the extent of black rockfish habitat and the densities of black rockfish residing there.
- An independent nearshore survey should be supported in all states to avoid the reliance on fishery-based CPUE indices.
- Stock structure for black rockfish is a complicated topic that needs further analysis. How this is determined (e.g., exploitation history, genetics, life history variability, biogeography, etc.) and what this means for management units needs to be further refined. This is a general issue for all nearshore stocks that likely have significant and small scale stock structure among and within states, but limited data collections to support small-scale management.

STAR Panel Recommendations (Black Rockfish 2015)

- *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

Kelp Greenling Research Recommendations (2015 Assessment)

- Fishery-independent surveys of abundance for nearshore species, including Kelp Greenling, would provide information about population trends that don't rely on data collected directly from the fishery and the natural complexities that those data entail. Surveys that result in a time series of information covering a representative spatial extent of the population would be most advantageous.
- Improved data collection relevant to basic fishery statistics (catch/effort) for recreational shore and estuary-boat fleets, including biological sampling where possible, to monitor changes in these highly dynamic fishing modes.
- The collection of gender-specific information is generally straightforward given the visual ease (color and markings) of identifying Kelp Greenling by gender and the collection of this information should be implemented for Ocean Recreational Boat Samplers (ORBS).
- The double reading of Kelp Greenling otoliths would provide some indication into error and bias for this influential source of information.
- Kelp Greenling stock structure needs to be studied and the results accounted for in future assessments. In particular, ontogenetic and gender-related movement according to offshore depth and spawning seems plausible for Kelp Greenling, and data to support that hypothesis would be beneficial for future assessments.
- Research into the implications and complexities of managing a stock where both genders contribute to spawning potential (e.g., through a Management Strategy Evaluation) would

help guide future assessments and management for species such as Kelp Greenling (males exhibit nest-guarding behavior).

STAR Panel Recommendations (Kelp Greenling 2015)

Specific recommendations for the next assessment:

- Acquire estimates of reader error for kelp greenling through routine double-reading samples of otoliths rather than relying on error estimates for other species, which may not accurately reflect actual read errors for this species.

- Acquire more age-readings, particularly of younger fish caught in the shore based or estuary fisheries, and try to estimate the full set of growth parameters.

- Consider seasons or a shorter time step for the model than a year (e.g., a quarter) or request a modification to Stock Synthesis to allow non-integer growth increments to increase accuracy for the estimated growth rate and error.

- Consider developing an alternative conditional likelihood (condition length-at-age) to make better use of ages to inform on population age structure where a significant proportion of aged fish are within one standard deviation of the asymptotic length. This would be limited to sampling that is random with respect to age or length.

- Composition data weighting should start with the number of samples as the effective sample size rather than the number of fish.

- Sample catches from the shore and estuary fisheries so that they can be estimated and characterized.

- Explore starting the model at the historical point where removals become more reliable (e.g., 1970-1980) and estimate initial conditions reflecting prior exploitation such as free numbers at age and an offset to equilibrium recruitment.

- Consider estimating recruitment deviations for the entire time-period of the assessment model.

- Density estimates from visual surveys and other methods could be used to derive a prior for unexploited biomass (perhaps transformed to R_0). This could help with the very poorly estimated scale in the current assessment.

- Future assessments should consider assessing a single stock for Washington, Oregon and California; this analysis could include explicit spatial areas and or state-based fleets as necessary. This was also a specific recommendation of the 2005 STAR panel. Expanding the spatial scale could make the interpretation of steepness based on life history theory more straightforward.

- The next assessment should be an update, given the status and trends. There is no urgency to update the assessment, unless negative trends appear in biological or catch-rate data.

Specific suggestions for the SSC:

- Given current estimates of growth and natural mortality, kelp greenling is probably much more productive than an $F_{45\%}$ policy would imply. None of the data series show any signal of depletion. The SSC may wish to consider recommending a different target SPR rate for this species.

General recommendations for nearshore species assessments:

- It may be more accurate in some circumstances to consider when historical catch history is so uncertain that the model should be started in a later year with the initial depletion estimated by the model.

- Consider the development of a coastwide fishery-independent survey for nearshore stocks. Any work to commence collection of such a measure for nearshore stocks, or use of existing data to derive such an index would greatly assist with this assessment.

- MRFFS data are difficult to use properly, making quantities derived from it potentially unreliable. Broadly, the MRFSS database needs to be cleaned (e.g., removing derived estimates that are not easily identified as such and that currently appear to be observed 'data'), better documented, and made more accessible for future assessment authors. It is important that the work of cleaning and interpreting these data not have to be repeated for each assessment. No further MRFSS data will ever be collected.

- For CPUE abundance indices, further evaluation of the Stephens-MacCall method would be valuable to ensure that the filtering method is robust. For example, the Stephens-MacCall method for filtering logbook records produced anomalous inclusion of the same species as both a positive and negative indicator in similar datasets. Pooling among similar series (e.g., charter boats and private boats from the same areas) to develop the filtering criteria could make this more stable. More generally, 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.

- Definition and measurement of suitable habitat for nearshore species such as kelp greenling, especially when combined with density estimates, would assist assessments, particularly as an independent indicator of plausible relative scale of modeled virgin biomass by area/region/state.

General recommendations for all assessments:

- Whenever age-readings are done, some portion of the effort should routinely include double-reads for estimating ageing error.

Widow Rockfish Research Recommendations (2015 Assessment)

- Historical landings and discards: The historical landings and discards are uncertain for Widow Rockfish and improvements would increase the certainty that fishing removals are applied appropriately. Because landings are assumed to be known exactly in the assessment model, uncertainty in the predictions does not include uncertainty in the landings. A thorough look at historical landings, species compositions, and discarding practices would potentially account for and possibly reduce the uncertainty. More importantly, though, a measure of uncertainty on the estimated historical landings would allow for reasonable sensitivities to be investigated.

- Natural mortality: Uncertainty in natural mortality translates into uncertain estimates of status and sustainable fishing levels for Widow Rockfish. The collection of additional age data, re-reading of older age samples, reading old age samples that are unread, and improved understanding of the life-history of Widow Rockfish may reduce that uncertainty.

- Maturity and fecundity: There are few studies on the maturity of Widow Rockfish and even less recent information. There have been no studies that reported results of a histological analysis. Further research on the maturity and fecundity of Widow Rockfish, the potential differences between areas, the possibility of changes over time would greatly improve the assessment of these species.
- Age data and error: There is a considerable amount of error in the age data and potential for bias. Investigating the ageing error and bias would help to understand the influences that the age data have on this assessment.
- Basin-wide understanding of stock structure, biology, connectivity, and distribution: This is a stock assessment for Widow Rockfish off of the west coast of the U.S. and does not consider data from British Columbia or Alaska. Further investigating and comparing the data and predictions from British Columbia and Alaska to determine if there are similarities with the U.S. West Coast observations would help to define the connectivity between Widow Rockfish north of the U.S.-Canada border.

STAR Panel Recommendations (Widow Rockfish 2015)

- The next iteration of this assessment should be an update assessment.

- Minor anomalies in the weight-length data from the PacFIN Biological Data System (BDS) should be excluded or reconciled.

- A reanalysis of the foreign at-sea index that best overlaps the period of largest stock decline could be conducted before the next assessment. In particular, an analysis should consider effort measures that include search as well as towing time, given the schooling nature of this species. Other fishery indices are unlikely to have an appreciable impact on the results and may not be worth reanalyzing.

- Widow rockfish should be considered in any future discussions about trans-national stocks. Although a joint assessment with Canada may be difficult to arrange, it should be explored. It is possible that lack of information from Canada affects estimates of productivity and, in particular, steepness. Until such time as a joint assessment can be conducted, evaluation of relative catches and trend information on abundance in Canadian waters would also be helpful. Potential exchange also clouds the clear interpretation of what represents steepness for this stock.

- Updated maturity data representing the current stock distribution should be collected and analyzed, preferably using histological methods.

- Since there was so little information in the data on steepness, the informative prior might be strong enough to allow for estimation in future assessments. This should be explored.

- Based on the variability estimated for the juvenile index, it should be removed from future analyses unless it can be improved and validated. Specifically, the estimated variance is greater than the RMSE of the recruitments, so it will add more noise than signal at the end of the time-series when there are no other data to inform recruitment. This decreases the predictive ability of the model.

- Although recreational removals are low in relation to other removals for this stock, these should at least be reported in a table for comparison in future documents.

- It may improve the model if the H&L and NET fisheries are combined with other fleets, as these represent very little removals and noisy data. Removals of these data did not appreciably change the results for this assessment and their selectivity showed similar patterns to other fleets. Removing these as separate fleets would likely to make the modelling simpler with no loss of signal.

- Select one or more fleets (as run-time allows) and create conditional age at length data in order to inform growth and selectivity from more than just the most recent years where survey data are available.

General recommendations for all assessments:

- 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.

- Additional work is required to further develop an objective procedure for evaluating the chosen stock boundaries across all rockfish (and potentially all other) assessments may be beneficial, and also more directly point to required directions for future research or assessment collaboration across national/international political boundaries. Further investigation is required for whether the stock boundaries assumed in the assessment are appropriate for management as well as scientific assessment.

- There is a need for more detailed examination of input data independent of the stock assessment, particularly in relation to sample size and representativeness. An examination of data sources by year and sub-area in particular may suggest appropriate methods for post-stratification of composition data. Potential stratification that should be considered should include season, latitude, depth, and boat type.

- Reports should include a section on how the recommendations from any data workshop and previous CIE reviewers were addressed. This would be an extension of the section for addressing previous STAR Panel recommendations (Appendix B in the 2014 Terms of Reference).

- Additional work is required in developing catch histories. An evaluation of the plausible range of proportions of species in the aggregated catch on the reconstructed catch time series is recommended. It would be most useful not only to provide single best estimates, but to define ranges suitable for use in bracketing uncertainties and sensitivity analyses.

- The state of Washington still needs a formal catch reconstruction to standardize approaches across assessments and ensure the best available estimates are being used.

- An objective procedure for identifying 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. The standard delta-GLM procedure should allow for different factors to be considered in the binomial and sub-models. A standard set of diagnostics should be provided to review panels for each abundance index including: 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. 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). It may be far more efficient to produce and review this output for groups of species together, rather than try to include it in each species-specific review.

- Reporting the extreme catch encounter probability in the ECE models would be an interesting diagnostic and additional piece of information for understanding how frequent exceptional catches are estimated to be.

- Expand and weight conditional age at length data to accurately represent both the sampling process and the numbers of fish predicted in each strata. This achieves logical consistency among data sets,

choices of the number of length bins, and imparts the greatest amount of orthogonal information possible to the assessment model. Where length sampling is random, marginal length data should be associated with conditional age at length data (by year and fleet) such that the recruitment information contained in the ages is not lost relative to what would be included if using marginal age compositions.

- An objective procedure using maximum age for a natural mortality prior needs to be developed and fully explored. Specifically, unless age samples were collected prior to significant exploitation, the oldest fish may be missing from the observed data, ageing error will tend to bias the maximum age to higher values and there may be sampling bias (e.g. domed selectivity). A percentile based method (e.g. use the 90th percentile) is likely to be more robust than using the single maximum age. Other information (e.g., GSI, growth, exploitation rates etc.) should be included, where possible in the derivation of the *M* prior.

- Where there is significant uncertainty in a very recent recruitment estimate (informed by very little data), it may be helpful to perform a likelihood profile over the strength of that year-class (running the model adjusting that deviation in the par file and using a "noest" option) to see where the primary signal was coming from.

- Aggregated residual plots (weighted and combined across all fleets would help to understand whether the model is fitting the available data adequately, even where patterns in residuals might show trade-off among fleets. This could reduce the over interpretation of residual patterns within the fit to a single fleet.

- Where there are marginal age- and length-compositional data being used from the same fish, iterative reweighting should be done first, then at the final stage an additional multiplier of 0.5 should be applied and no additional iteration performed. This retains the goals of both logically consistent reweighting and down weighting the doubly-used data.

- MCMC results are a useful tool to measure uncertainty and diagnose problems in the assessment as well as provide an alternative to MLE-based results, which can differ appreciably in terms of point estimates and uncertainty (Stewart et al. 2013). MCMC should be routinely used and reported where possible.

- Producing at least one model run with the full time-period of estimated recruitment deviations would be a very helpful diagnostic and could be a plausible base case depending on the model behavior.

- Triggering a future full assessment could be based on monitoring the most reliable indices, such as surveys with confidence bounds. In comparing observed and predicted values, values outside confidence limits may suggest a higher priority for more immediate assessment.

In addition to current R4SS and SS3 functions, the following additional features and standardized procedures should be developed:

- Procedures for examining sources of information on recruitment events is required. This could include profile over recruitment events or partition likelihood components.

- A method to examine observed and expected sex ratio by age and through time would resolve questions about the consistency of sex ratios being produced for the modeled population.

- Developing residual plots that are weighted across data sources would allow comparisons to be made that might help to identify common patterns.

- Removal of the re-scaling to 1 problem after weighting is applied to composition data

- Development of standard procedures for the selection of the most appropriate weighting system that should be applied to input data (additional sd for indices, harmonic mean/Francis/other for length and marginal age comps, harmonic mean/Francis/other for conditional age-at-length data).