# Chilipepper Rockfish STAR Panel Report

**National Marine Fisheries Service**
**SWFSC Santa Cruz**

**June 25-29, 2007**

**Reviewers:**

David Sampson, Scientific and Statistical Committee Representative, Panel Chair
Patrick Cordue, Center for Independent Experts
Norman Hall, Center for Independent Experts
Kevin Piner, NOAA Fisheries Service, Southwest Fisheries Science Center

**Advisors:**

Gerry Richter, Groundfish Advisory Subpanel Representative
John DeVore, Groundfish Management Team Representative

**STAT Team:**

John C. Field, NOAA Fisheries Service, Southwest Fisheries Science Center, Fisheries Ecology Division.

## Overview

The STAR Panel met June 25-29$^{th}$ in Santa Cruz, CA and reviewed the draft stock assessment for chilipepper rockfish off California, which was conducted using the SS2 software and based on the following data sources: annual landings from 1892 for four fisheries (trawl, hook and line, set net, and recreational); five biomass/abundance indices (the triennial bottom trawl survey, the NWFSC shelf/slope combination bottom trawl survey, catch rates from trawl logbooks, catch rates from the northern California CPFV observer database, and the coast-wide SWFSC juvenile rockfish abundance survey); age-composition data from three fisheries (trawl, hook and line, and set net) and the NWFSC combination survey; and length-composition data from four fisheries (trawl, hook and line, set net, and recreational) and two surveys (the triennial bottom trawl survey and the NWFSC combination survey). This stock had not been assessed since 1998.

The draft assessment document distributed prior to the STAR Panel meeting described a preliminary assessment model that included conditional age-at-length compositions rather than age-compositions. Problems with tuning this model resulted in the STAT bringing to the Panel a revised assessment model that had length-compositions and age-compositions and no conditional age-at-length compositions. The STAR Panel accepted that the conditional age-at-length approach should not be pursued during the meeting. To compensate for using some sampled fish in both age- and length-compositions, it was agreed that revised models should down-weight those length-compositions for which there were also age-compositions. A review of the age-composition data uncovered some apparently biased age-composition sampling, as evidenced by large discrepancies between the length-compositions of aged versus un-aged fish. This resulted in additional data filtering to identify and remove suspect age-composition data.

Because chilipepper rockfish are known to be semi-pelagic, there were concerns that the two available bottom trawl surveys would not provide reliable biomass indices. Of all the available indices, the CPFV index, based on observed angler catch rates at defined fishing sites, seemed the most likely to provide a reliable abundance index. Hence the STAR asked the STAT to investigate the consequences of focusing on the CPFV index (based on observer data on recreational CPUE) as the primary tuning index. The model, however, generally predicted flat trends for the CPFV index, which seemed inconsistent with evidence in other data sources that indicated an exceptionally strong 1984 year-class. This inconsistency led to exploration of alternative selectivity configurations that would allow a closer fit to the CPFV index. The accepted base model provides a reasonable fit to the trends apparent in the survey by means of a composite age- and length-based selection curve for the CPFV survey, but there is no direct evidence of a mechanism for such selection.

The preliminary base model (distributed prior to the STAR Panel) was configured to allow time-variation in the growth coefficient K, with changes in K occurring at three-year intervals. It was agreed that time-variation in growth was a sensible feature to explore given the inter-annual variation in mean size-at-age apparent in this stock. However, rather than imposing an arbitrary three-year blocking pattern, time-varying growth in the final base model was incorporated using blocking derived from low-frequency changes in the Pacific Decadal Oscillation (PDO) index.

After exploring the accepted base model along several dimensions of uncertainty, including the rate of natural mortality and the level of historical catch, it was agreed that the major axis of uncertainty should be the steepness parameter (h), which provided reasonable contrast in the level of stock depletion. Low and high values of h for a decision table were derived from a normal prior probability distribution based on a meta-analysis of rockfish steepness parameters. The final decision table was developed after the STAR Panel meeting, based on consultation with the GMT and GAP advisors regarding appropriate harvest levels to include in the projections.

The STAR Panel commends Dr John Field, the STAT, for his hard work and cheerful willingness to address issues arising during the course of the STAR review. Despite encountering technical difficulties before and during the STAR review, Dr Field persisted and was able to find suitable solutions and develop an acceptable base model and alternative runs that adequately captured the uncertainty of the model. The next full assessment of the chilipepper rockfish stock should re-investigate using conditional age-at-length data, rather than non-independent length- and age-composition data, and should further explore environmentally driven changes in growth. There should also be fuller investigation of the effects of uncertainty in the catch history.

## Analyses requested by the STAR Panel

The initial presentation by the STAT and accompanying supplemental material distributed on the first day of the STAR indicated that the preliminary model was very sensitive to "tuning" adjustments to the variance weightings of the likelihood components, there being large differences between the results of tuned versus un-tuned models. The STAR Panel considered this sensitivity to tuning to be an indication of tension between inconsistent data sets and proposed that the STAT explore the problem using a simplified model, with additional complexity being introduced later in a stepwise manner. The STAR Panel endorsed the view of the STAT that the juvenile core survey index should be removed from the SS2 analysis as the data are extremely noisy and the information content appears inadequate given the limited spatial coverage of the core survey. The STAR Panel also agreed with the STAT's view that the CalCOFI index was not suitable for chilipepper rockfish because the survey misses much of the spatial range of the stock.

**Round 1 requests**

A. Compare the length-composition of the aged fish with non-aged fish for each fishery and each year.

B. Fix the code for the recreational CPUE to be number-based rather than biomass-based.

C. Reset the lambdas on LFs to 0.1 if age data exist, and to 1 if there are no associated age data for the same samples. Run with:

- No CalCOFI or core juvenile;
- No time varying K – fix at the values of all growth parameters of the earlier conditional runs;
- Trawl CPUE indices;
- Rec CPUE;
- Triennial Survey;
- Combined survey;
- Coast-wide juvenile index;
- Fix h at something reasonable;
- Fix M for females and estimate offset for males;
- Fix CV of length at age at 0.06 [based on external analysis done by the STAT];
- Profile over M including likelihood components;
- Estimate selectivity parameters;
- Estimate SSB0;
- Estimate depletion.

D. Save the results from the un-tuned model

E. Tune the trial reference model – see fit for everything. Plots and tables of diagnostics and results.

F. Profile over M for the tuned model looking at individual likelihood components – identify inconsistencies among data sources.

G. Plot or tabulate spatial distribution of samples in recreational data from observers over time.

*Round 1 responses*

A. The length-compositions of the aged and non-aged chilipepper rockfish were plotted by the STAT for approximately 50% of the samples from each fishing gear. These plots indicated that, for a number of years, fish selected for ageing appeared to be larger than the non-aged fish from the same year, which not only affected the length-compositions but also affected the sex ratio because of sexual dimorphism in growth. While the proportions-at-age of fish of a given length are unlikely to be affected, the mean value of length-at-age is likely to be positively biased if the biased data are included in the SS2 analysis.

B. The SS2 control switch for the CPFV survey (the recreational fishery CPUE index) was corrected to indicate that the data represented numbers of fish rather than biomass.

C. The SS2 model specified for this request was set up and run with h fixed at 0.57 (based on a meta-analysis by Martin Dorn, personal communication) and M for females was fixed at 0.16. These values were consistent with the maximum likelihood estimate based on profiling over M and h. The length-composition data were down-weighted as requested as an ad hoc correction for non-independence of the data. A more appropriate method to address the problem is discussed in Appendix 1.

D. Results of the un-tuned model were saved as requested.

E. Predictions of the abundance/biomass indices from the tuned model failed to reflect the large 1984 cohort apparent in the model-estimated recruitment and catch at age data. The predicted values for the CPFV survey in particular showed no decline despite a clear downward trend in the observed values for this index.

F. The profile plot over M revealed the tension between the data sets, especially between the trawl fishery and the recreational CPFV survey. Higher estimates of spawning stock biomass were associated with higher values of M.

G. To address concern that possible unbalanced sampling in the CPFV observer data could invalidate the GLM time series as an index of abundance, the STAT generated tables and plots of both the number of trips in which samples were taken and the number of chilipepper rockfish caught by depth categories and year. A small number of samples, each with a large number of fish, collected from depths greater than 80 fathom were recorded in the years prior to 1994. To ensure consistency in depth ranges covered by the survey through time it was agreed that fish from depths greater than 80 fathom would be excluded from the GLM analysis.

Other analyses presented by the STAT in response to issues raised by the Panel.

- The time series of estimates of the CPFV recreational CPUE index produced using the Stephens/MacCall filter was very similar to that produced by the GLM using depth and block data, which suggests that the filter was working properly to identify trips likely to catch chilipepper. The CVs of the results from the filter were less than those from the GLM. Results produced by the GLM using only year effects were highly correlated with those produced by the original GLM but lay below them. A very similar result was produced when the deepest depth bins were dropped from the analysis.

- To explore inconsistency between GLMM and area-swept estimates of survey biomass, which had been highlighted during the initial presentation and review of data sources, the STAT contacted Dr Tom Helser (NWFSC), who had provided the STAT with the survey biomass indices. Dr Helser sent a more detailed description of the GLMM analysis, accompanied by results and diagnostic plots. The discrepancy between the scales of the results of the GLMM analysis and the swept-area approach appears to reside in the presence of occasional large catches of chilipepper rockfish (i.e., the patchiness of the distribution) and the use in the GLMM of a log-normal distribution. Diagnostic plots from the GLMM indicated that the model provides good predictions of the data.

**Round 2 requests**

Based on the reference run that was established on Monday evening (Round 1):

H. Test for block-year interaction in GLM for recreational observer CPFV data. If a strong interaction is detected, report back to this issue and complete points I to M, but do not undertake the additional runs at points N to P.

I. Plot length-compositions of aged versus non-aged fish in remaining samples to determine those samples which are relatively unbiased. Weed out obviously biased samples from the SS2 input including those samples that had infeasible numbers of large males.

J. Investigate samples that had extraordinarily large proportions of males.

K. Link RecFIN length-compositions to the recreational fishery and CPFV observer length-composition to the CPFV CPUE survey to assist in elucidating the respective selectivity curves.

L. Remove whole of deep trips >80.

M. Use Helser's GLMM rather than area swept index.

N. Estimate an appropriate selectivity pattern for triennial survey.

O. Systematically set lambda for recreational observer CPFV index to 1, 5, 10, … till a reasonable fit to this index is attained and investigate changes in likelihood for all other components.

P. Profile over R0 as was done for M, plotting against B0.

*Round 2 responses*

H. There was concern that presence of a strong block-year interaction would require a different analysis to derive a suitable blending of non-parallel abundance trends. It was not possible, however, to detect block-year interactions in the GLM for the recreational observer CPFV data because the data were too sparse. The value of AIC produced did not indicate an improved fit that justified the block-year interactions.

I. The length-compositions of aged and non-aged samples were plotted for samples not previously examined to subjectively identify samples for which there may have been biased selection of fish for ageing. The STAT advised that the data were raw and unexpanded. (Later in the week the Panel concluded that it would have been more appropriate to do the data screening based on length-compositions expanded to account for differing sampling rates.)

In subsequent SS2 runs for this round the STAT had “turned off” the length-compositions of the biased samples. However, as it was the age sample selection that was biased, it would have been more appropriate to retain the length-compositions and “turn off” the age-compositions. The Panel suggested that the likelihoods calculated in SS2 for the length samples and associated age samples may be inappropriate as the two samples were not independent. Patrick Cordue advised that he would derive the likelihood function and that he would include the necessary equations in an appendix.

J. After filtering to remove outliers, the length-composition for one sample still contained a number of unfeasibly large males. This length-composition should also be “turned off” in the SS2 analysis.

K. In the preliminary model the CPFV index was biomass-based and was linked with the recreational fishery, and the CPFV length-composition data were used to represent the recreational fishery. There were no RecFIN length-composition data to represent the full range of recreational fishing modes. Linking RecFIN length-compositions to the recreational fishery and CPFV length-composition to the CPFV survey produced little change in the biomass trajectory. (This request was done after completing changes specified in request M). The STAT advised that the RecFIN length-compositions were expanded length-compositions as produced by RecFIN.

L. The STAT reported that removing the data for trips >80 fathoms, including associated length data, had little effect on the biomass trajectory.

M. The use of the GLMM results rather than the swept area indices for the triennial and NWFSC combination survey resulted in slightly greater depletion than in the previous run. (This request was done before completing the changes specified in request K).

N. The STAT encountered difficulties fitting the selectivity curve for the triennial survey. The resulting logistic curve was essentially a horizontal line, apparently so the model could accommodate small fish. Also, the Hessian for this run could not be inverted.

O. The Panel wanted to understand why the model was not providing a reasonable fit to the CPFV recreational observer index, which should have been a more reliable index than other available indices. Elevated lambdas on the CPFV index resulted in lower biomass trajectories and apparently greater depletion, with a better fit to the CPFV and triennial indices but poorer fit to the trawl CPUE index. However, even with lambda = 25 the predicted CPFV index failed to reflect the strong 1984 year class, which was evident in other data sources and seemed to be reflected in the observed CPFV index value for 1992. Further exploration by the STAT found a configuration that produced a slight signal of the strong year class in the predicted CPFV index: lambda = 5 for the CPFV survey, selectivity for the CPFV survey was dome-shaped for both length and age, and growth was time-varying (with a 3-year blocking pattern). The resulting predicted length-compositions for the CPFV survey reflected the bimodality present in the observed length data. The predicted length-compositions using length-based selectivity alone did not appear to fit as well and failed to produce similar bimodality.

P. The STAT had insufficient time to satisfy this request.

**Round 3 requests**

Q. Modify the SS2 input specification to turn off the age-composition data where samples were biased (as determined from comparison of aged and non-aged LF data) and turn length-composition data back on. For the sample with an infeasible number of large males, turn off both age and length-compositions.

R. Using lambda for CPFV survey data set to 1, run SS2 to provide a reference for subsequent runs

S. Investigate alternative parameterisation for sex-specific selection curves for the CPFV survey using either age OR length selection (but not both) and hence determine a suitable selection pattern to use. Save runs.

T. Using the final selection curve from Request S, produce a simple profile analysis based on R0 to explore the tension among different indices and data sets.

*Round 3 responses*

Q. The changes were completed to remove the effect of biased sampling for age but retain the associated length data.

R. The run was completed as requested. Turning off the biased age-composition data did not have a major impact on the predictions of biomass, nor did it help the fit to the CPFV survey data.

S. The rationale for this request was to find a selection curve for the CPFV survey that would fit the CPFV index and length-composition data without the complexity of the composite age- and length-based curve that the STAT had used in response O. The STAT replaced the CPFV length-based selection curve with an age-based curve, which went asymptotic when fitted. The resulting fit appeared slightly better than that obtained with length-based selectivity. Unfortunately, the STAT had not noted the request that the selectivity curve should be sex-specific, and this had not been implemented.

T. Although the STAT produced profiles on R0 as requested, the runs were for CPFV selectivity that was age-based but not sex-specific. The profiles did not provide the information that the Panel had sought.

**Round 4 requests**

U. Complete Request S. That is, search for alternative parameterisation for sex-specific selection curves for the CPFV survey using either age OR length selection (but not both) and hence determine a suitable selection pattern to use. Save runs.

V. Using the final selection curve from Request U, produce a simple profile analysis based on R0 to explore the tension among different indices and data sets.

W. Explore alternative blocking for time-varying growth based on external environmental variables.

*Round 4 responses*

U. In its exploration, the STAT had been unable to fit an age-based or length-based, sex-specific selection curve that provided as good a model fit as that obtained by the age- and length-based selection curve (not sex-specific). The Panel noted that in tuning a model where sampled fish contribute to both length- and age-compositions the effective sample size for aged sub-samples should be linked to that of the associated length samples. The Panel also advised that when age data are derived from a sub-sample of the length data the likelihood function describing the length- and age-at-length composition data has additional components that are “constant” (i.e., independent of estimated parameters) but could influence model fits if the sub-samples were biased and the bias was estimated (Appendix 1).

V. The profile over R0 was completed as requested. The relative impact on the overall likelihood of the different model components at different values of R0 could not be compared easily using the profile plots because the plots did not account for the effect of lambda, which was reduced to 0.1 for some components. Using sex-specific selection for the CPFV survey did not appear to warrant further investigation.

W. The Panel was concerned that there was no basis for arbitrarily blocking changes in growth at three-year intervals, especially given the STAT's view, expressed during its initial presentation, that changes in growth were driven by changing oceanographic conditions. The STAT presented information on the Pacific Decadal Oscillation (PDO) index, which others have shown to be related to zooplankton production. The Panel agreed with the STAT that the PDO provided an adequate basis for blocking the modelled time period into six blocks. This model, with six time-blocks for growth parameter K, resulted in a large improvement in the log-likelihood, but the value of K for the final time-block was much lower than the values for previous time-blocks. The Panel suggested using an informed prior to ensure that the offset of K for the last block did not fall to an unrealistic level as it had in the runs produced for this request.

**Round 5 requests**

X. Investigate feasibility of driving K with PDO (spend no more than half hour on this task).

Y. Adopt time-varying growth based on the better of using either PDO blocks (with slightly-informative prior on K to avoid infeasible reduction in K for last period) or using environmentally-driven growth (Request X), and using both age and size-selectivity on the CPFV CPUE recreational survey, create tuned base. Demonstrate adequate convergence of tuned run.

Z. Produce profile plots on R0 accounting for lambda.

AA. Using base run, produce standard diagnostics for STAR Panel review.

*Round 5 responses*

X. The Panel and STAT agreed that effects similar to those obtained from time-blocking might be obtained by directly relating K with the PDO index, on which the blocking pattern was based. Use of the PDO index to drive K showed promise but this simpler model structure failed to improve the fit obtained using K-offsets in the six blocking periods. It was agreed that the base model should use the time-blocking approach.

Y. A value of 0.5 was used as the standard deviation for a slightly informative prior on K for the configuration with six PDO-based time-blocks for changes in K. The convergence-test runs that used "jittered" starting parameter values revealed convergence problems, with many runs clearly not converging. Four runs apparently converged but produced different solutions at different values of R0: two with high overall likelihood values and two with low. The different solutions appeared to be associated with changes in the values of the K-offset for the last two time-blocks. Evidently the likelihood surface is quite irregular.

Z. This request was not completed because of the convergence problems uncovered in response Y.

AA. This request was not completed because of the convergence problems uncovered in response Y.

**Round 6 requests**

AB. Explore convergence and results of time-varying K with (a) last two blocks combined into a single large block and (b) changing the standard deviation for the prior on the deviations on K from 0.5 to 0.35.

AC. Use 0.5 on the K-dev prior. Run with five-block rather than 6-block model. Examine results.

AD. Turn off all priors. Run with five-block rather than 6-block model. Examine results

AE. Use run from Request AD. Clean up initial values. Make qs analytical. Clean up phasing. Do jitters and alternative phasing to confirm model convergence. If not converged, report back ASAP. If converged, produce a full set of diagnostic results and profile plots on R0 accounting for lambda. If these are satisfactory, this will be the base model.

*Round 6 responses*

AB. The two requested runs explored alternative methods for constraining the growth coefficient K in the final time block. The Panel was concerned that the unconstrained estimate for the final K value was extremely small and would have a strong influence on forecasts. The run with the standard deviation for the prior probability reduced to 0.35 still produced a low value for the final K. The run that merged the last two blocks in combination with a standard deviation of 0.35 for the prior probability resulted in an intermediate value of K.

AC. The Panel sought confirmation that having the longer final block in the five-block model would provide sufficient constraint for the final K value and that the prior probability on the K-offsets could be eliminated. The use of a standard deviation value of 0.5 for the prior probability on the K-offsets did not have much effect on the results.

AD. Likelihood summaries examined in connection with earlier responses indicated that some parameters other than the offsets to K were also being constrained by prior probabilities. Because there were no appreciable differences between runs with prior probabilities and runs in which they had been eliminated, the Panel concluded that it would be appropriate to remove all prior probabilities and thus simplify the model configuration.

AE. The Panel requested a general clean-up and simplification of the SS2 control file in hope that this would improve convergence of the model. However, convergence test runs with jittered initial parameter values indicated there still were convergence problems. Roughness that was evident in the profile plots was probably a reflection of lack of full convergence but the effects did not appear too severe.

**Round 7 requests**

AF. Set process error added to CPFV survey indices to 0. Re-run. Confirm that this is appropriate to use as a base model through jitters and alternative phasing to confirm model convergence.

AG. With settings resulting from Request AF, increase emphasis to 20 on both CPFV survey indices and length frequencies to estimate age-based, sex-specific selectivity. Assess whether this gives sensible selection patterns. If so, using the resulting parameter space and selectivity pattern (possibly fixing selectivity parameters to the resulting values), de-emphasise, re-fit, and re-tune to produce plausible alternative results (removing process error if necessary after tuning). Note – no more than ~45 minutes to be spent on this task. Produce a plot of the biomass trajectory of this compared with the result from Request AF as a sensitivity analysis. Compare the depletion estimates.

AH. With settings resulting from Request AF, explore the following dimensions of uncertainty using low and high values for (a) historical catch prior to 1978 (half and double), (b) M, and (c) h. Retain SS2 results from each run. Produce comparative plots of the biomass trajectories of these compared with the result from Request AF. Produce a table showing comparison of likelihood contributions from different components. Produce a table of comparative depletion estimates.

*Round 7 responses*

AF. Removing the variance adjustment on the CPFV survey index had the desired effect of producing a better fit to the CPFV survey. After reviewing diagnostic plots the Panel recommended acceptance of this model configuration as the base model.

AG. These sensitivity runs re-explored using an alternative configuration for the CPFV survey selection curve. Previous explorations had increased the lambda on the CPFV survey index but not on the CPFV length-composition data. The new runs produced a very good fit to the CPFV index even when lambda was decreased from 20 to 10, but the CPFV selectivity curve had been configured as age- and length-based and sex-specific. Convergence tests with jittered initial parameter values still produced fits that appeared not fully converged.

During discussions the STAT indicated that the CVs for the triennial and combination surveys had been reduced externally rather than with a variance adjustment factor in the SS2 control file. Because the model provided good fits to several survey data points that had very large input CVs, the standard variance adjustment approach would have produced negative CVs for other data points with small input CVs. The Panel notes that further consideration is needed to develop an appropriate approach for handling survey variance adjustments that could potentially become negative.

AH. The runs were completed as requested. The resulting profile plots were somewhat jagged, suggesting that the model had failed to converge fully at many values of the reference variable. Following examination of the profile plots the Panel concluded that, of the variables considered, h was likely to provide the most useful axis of uncertainty. The Panel recommended assuming a normal distribution for h with a mean value of 0.573 and standard deviation of 0.183 to determine the bracketing values.

**Round 8 requests**

AI. Complete Request AG to estimate age-based, sex-specific selectivity. Run and produce comparison of results.

AJ. For developing a decision table, run the base model with h = 0.34 and 0.81 [mean values of the lower and upper 25% of the prior probability distribution for h] to obtain results likely to be representative of the lower 25% and upper 25% of values, respectively. Use the alternative phasing supplied by the STAR Panel. Jitter and ensure convergence for each value of h.

*Round 8 responses*

AI. The response to AG had used a sex-specific, age- and length-based selection curve for the CPFV survey. Results demonstrated that, although needing further refinement, an age-based, sex-specific selectivity curve could be developed to replace the age- and length-based, sex-specific selectivity curve.

AJ. While there were still convergence issues that required jittering of input parameter values for each analysis, the jittered runs for each level of steepness produced reasonably similar results. Depletion for the base case was 0.7, while those from the lower and higher values of h were 0.46 and 0.78, respectively. The Panel accepted that use of these values of h produced the required lower and upper runs to bracket uncertainty around the base-run results.

## Final base model description

The agreed base model configuration for chilipepper rockfish had the following characteristics:

- Single-area model with two sexes.
- Stock initially at equilibrium with zero harvest. First harvests in 1892.
- Age- and length-compositions included but with down-weighting of length-compositions that had associated age data.
- No conditional age-at-length composition data.
- Fixed natural mortality coefficients (0.16 for females and 0.202 for males).
- Steepness parameter (h) fixed at 0.57.

- Assumed value of 1.0 for sigma-R.
- Recruitment deviations estimated for 1965-2006.
- Length-based selection with no sex-offset for all fisheries and surveys, except for the CPFV abundance index, for which selection is length- and age-based, with no sex-offset.
- Sex-specific growth coefficients (K) allowed to vary during 1970-2006 according to a five-block pattern based on changes in the PDO index.
- Other growth parameters estimated outside the model.
- No prior probabilities on any parameters.

## Comments on the technical merits and/or deficiencies of the assessment

### Technical Merits

- The STAT excluded some data sets from the assessment model based on pre-evaluations of potential input data sets. This is a more sensible approach than mixing good data with bad.
- The STAT made proficient use of SS2 and accompanying software, which greatly facilitated the Panel's review.
- Use of time-varying growth coupled with changes in the PDO was a useful innovation and the Panel encourages further work on the approach.

### Technical Deficiencies

- Good length-composition data may have been excluded from the model because the data filtering to detect biased age samples was based on unexpanded length-compositions. (The Panel wrongly instructed the STAT to examine unexpanded length-compositions.)
- The approach applied in this assessment of down-weighting length-composition data when associated age-composition data are included is ad-hoc and has no good theoretical basis. The age data from fish that contribute both age and length data should be handled instead as conditional age-at-length compositions. See Appendix 1.
- The model tuning process that adjusted for inconsistencies between the "input" and "effective" sample sizes for length and age compositions treated the age- and length-compositions as independent even though length/age data for some fish were included in both length- and age-compositions. If dependent age and length frequencies are used (which is not recommended) then a method needs to be developed for their joint tuning.
- The model tuning process that adjusted for inconsistencies between the model fits to surveys (RMSE) and the input CVs took an ad hoc approach with surveys that had very large CVs for some index values. The input CVs were reduced proportionally. This is inconsistent with the normal basis for the tuning process, which involves adding a constant to account for process error.
- The estimated growth curves at the L1 value had kinks that could probably be eliminated by reducing the lower bound of the smallest length bin.

- Results from the convergence tests with randomly jittered starting parameter values indicated that the likelihood surface is very irregular, which implies that the final model runs may not have fully converged.  However, the biomass trajectories and other critical results do not appear to be sensitive to any lack of convergence.
- All final runs used a composite length-age selection curve for the main tuning index (the CPFV survey), but currently there is no obvious rationale for such complex selection.  Using an age-based, sex-specific selection curve showed promise as an alternative configuration.  It was able to provide a good fit to the CPFV length-composition data and the decline in the CPFV index when given high lambda values on these likelihood components.  The estimated depletion from this alternative was not inconsistent with the range used in the decision table analysis.

## Areas of disagreement regarding STAR Panel recommendations

There were no areas of disagreement among the Panellists or between the STAR and the STAT.

## Unresolved problems and major uncertainties

This section focuses on major uncertainties.  Unresolved problem are discussed above under Technical Deficiencies.

- The base model configuration developed during the STAR meeting was based on the assumption that the survey index from the CPFV observer data is a reliable index of abundance.
- The full range of plausible catch histories has not been explored and the final model does not fully capture the influence of catch history on uncertainty in the biomass trajectory.
- The plausible parameter space for the assessment was not fully explored, but it was implausible to do so given the timeframe of the review and current technology.
- Spatial structure has been ignored in the model (e.g., north-south split at Point Conception).

## Concerns raised by GMT and GAP representatives during the meeting

The GAP and GMT representatives expressed concern that the STAT had difficulty gaining access to some of the raw survey data and thus could not fully explore those data.

## Recommendations for future research and data collection

### For the next chilipepper rockfish stock assessment

- Reconstruct the chilipepper rockfish catch history using all available data including catch by gear and by region.  The reconstruction should include an envelope of high and low values to set bounds for exploration of alternative catch histories.  The Panel notes that the SWFSC has made significant progress in retrieving detailed historical landings data, which will facilitate catch reconstructions.  As has been recommended previously by a variety of STAR Panels, the reconstruction of historical rockfish landings needs to be done comprehensively across all rockfish species to ensure efficiency and consistency.

- Read chilipepper rockfish otoliths from the triennial and combination bottom trawl surveys to provide better data on the early stages of growth and possible time-variations in growth.
- Explore use of conditional age-at-length data rather than coupled age- and length-composition data.
- Explore time-varying growth as influenced by environmental changes.
- Explore possible spatial structuring of the data and model.
- The next STAT should have full access to raw data from the NWFSC trawl survey.

**For the longer term**

- Age-validation of chilipepper rockfish should be pursued.
- Develop a fishery-independent time series using fixed sites and volunteer anglers who use standard protocols and are properly supervised.
- Establish a meta-database that provides a comprehensive overview of all relevant data sources and sufficient information to correctly interpret the data.
- Establish an accessible database for rockfish catch histories by species, including envelopes of high and low values for each species to set bounds for exploration of alternative catch histories.
- Relevant raw data, updated in a timely manner, should be readily accessible to assessment authors in on-line databases that are user-friendly.
- Develop comprehensive descriptive analyses of recreational fisheries and fleets to assist in interpretation of recreational CPUE and length-composition data.
- Develop a concise set of documents that provide details of common data sources and methods used for analyzing the data to derive assessment model inputs.

## Acknowledgements

The staff at the SWFSC Santa Cruz laboratory provided excellent hospitality and support for the STAR meeting.  The Panel especially thanks Steve Ralston for making certain we were well provisioned.

## Appendix 1: Modeling of age and length data

By Patrick Cordue

The appropriate use of age and length samples in stock assessments is important in obtaining robust stock assessment results. In a likelihood setting, the key is the application of appropriate likelihoods given the nature of the data – which is dependent upon how it was collected.

Age frequencies and length frequencies for a given fishery or abundance survey may be obtained independently or in combination. The usual likelihood used for both is a multinomial with an "effective sample size" which is smaller than the actual number of fish measured or aged (for length frequencies, the effective sample size is often similar in magnitude to the number of samples taken rather than the number of fish measured).

When a length frequency is sub-sampled for age, it is not immediately clear how the dependence between the length frequency and the age data should be represented. Two approaches have been taken in rockfish assessments. The most common method is to use both the length and age frequency in the assessment but to down-weight the joint contribution of the data to the total likelihood by adjusting emphasis factors on the individual components (e.g., lambda = 0.1 for length samples where a sub-sampled age frequency is also present; or lambda = 0.5 for both the age and length frequency). An alternative, which is theoretically better, when both age and length are used, is to use the age data as conditional age-at-length.

The latter method requires the input of the proportions at age for given length (class). The same approach is used when there are independent age and length samples, but the age sample was obtained from non-random length samples (e.g., to obtain a growth curve). The age frequency is biased, but the conditional age-at-length data are not.

The distinction between the two situations is the issue of independence between the length frequency and the age-length data. When there is sub-sampling of a length frequency for age, the length data and the age sub-sample are clearly not independent. It follows, in comparison to independent samples, that there must be an additional likelihood component which "links" the two data sets. It is very instructive to derive the likelihood and see why this component is important but also why it does not contribute to the total likelihood when fish are selected at random for the age sub-sampling.

Suppose that $n_l$ fish are sampled at random for length from a population (in a statistical sense). Further, suppose that $n_a$ fish are then sub-sampled at random for age.

Assume that there are $m$ length classes and let $L_i$ denote the number of fish in the $i$th length class for the length sample. Let $X_{ij}$ denote the number of fish in the $i$th length class and $j$th age class in the sub-sample for age. Adopting the notation of lowercase letters for observations of the random variables and bold notation to represent vectors or matrices, it follows from conditional probability theory that,

$$\mathrm{P}(\boldsymbol{L} = \boldsymbol{l}, \boldsymbol{X} = \boldsymbol{x}) = \mathrm{P}(\boldsymbol{L} = \boldsymbol{l})\, \mathrm{P}(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{L} = \boldsymbol{l})$$

The likelihood for $\boldsymbol{L}$ is a multinomial:

$$P(\boldsymbol{L} = \boldsymbol{l}) = \text{Mult}(\boldsymbol{l} \mid n_l, \boldsymbol{p})$$

where $\boldsymbol{p}$ is the vector of proportions at length in the population.

The conditional likelihood is derived by applying a further conditional construction:

$$P(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{L} = \boldsymbol{l}) = P(\boldsymbol{U} = \boldsymbol{u} \mid \boldsymbol{L} = \boldsymbol{l})\, P(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{U} = \boldsymbol{u}, \boldsymbol{L} = \boldsymbol{l})$$

where $U_i$ is the number of fish in the $i$th length class in the age sub-sample.

The conditional likelihood for $\boldsymbol{U}$ is another multinomial:

$$P(\boldsymbol{U} = \boldsymbol{u} \mid \boldsymbol{L} = \boldsymbol{l}) = \text{Mult}(\boldsymbol{u} \mid n_a, \boldsymbol{s})$$

where $s_i = l_i / n_l$ is the proportion of fish in the $i$th length bin in the length sample.

The final component in the joint likelihood is the conditional age-at-length likelihood:

$$P(\mathbf{X} = \mathbf{x} \mid \mathbf{U} = \mathbf{u}, \mathbf{L} = \mathbf{l}) = \prod_{i=1}^{m} \text{Mult}(\mathbf{x_i} \mid u_i, \mathbf{p_i})$$

where $\boldsymbol{p}_i$ is the vector of proportions at age in the population for the $i$th length class.

Hence, the joint likelihood of the length sample sub-sampled for age is the product of the likelihood for the length frequency, the conditional age-at-length, and the "linking" component being the sub-sample for length associated with the age sampling.

If the sub-sample of length is truly at random then the linking component consists entirely of "constants" (in terms of population parameters) and so does not need to be included for estimation purposes. Alternatively, if the sampling is biased, but the bias depends only on the characteristics of the length sample, then the linking component can still be ignored (even across a time series, despite the fact that the "constant" varies).

However, if sub-sampling for length is not random and depends upon population parameters then the linking component is potentially important. To adhere to a strict likelihood approach, it would be necessary to include the population parameters driving the bias in an appropriate parameterization to account for the biased selection process. When a time series of length and age data are used it is important to check for potential bias in the length sub-sampling and to consider if it could be driven by population parameters. If that could be occurring in some years, then the associated age data should perhaps be removed or the annual biases should be estimated using a joint likelihood that includes an appropriately parameterized bias function in the probability vector of the linking likelihood component.

Of course, one does not necessarily need to adhere to a strict likelihood approach. It can be argued that any bias in the sub-sampling for age can be ignored when the age data are used as conditional age-at-length. The argument being that the linking component may potentially provide information about population parameters, if the bias truly is driven by them, but by ignoring the component, potential information is forgone, but existing information in the other data is not compromised.

An important point emphasized by the full joint likelihood is the linkage of the length and age data in terms of their sample sizes. This is perhaps obvious in hindsight, but when “tuning” of age and length data is done during a stock assessment (i.e., an iterative adjustment of effective sample sizes to ensure that input variance assumptions are consistent with residual variances) it is crucial to maintain the consistency of the age and length sample sizes. That is, they must *not* be tuned independently. The relative contributions of each year’s age and length data to the total log likelihood of a full age and length time series will be proportionally maintained if effective sample sizes are scaled by the same multiplier both between and within years.