Report on

Blue and Black rockfish STAR Panel

May 21-25, 2007

Pacific States Marine Fisheries Commission

205 SE Spokane Street

Suite 100

Portland

Oregon

P.L. Cordue

Fisheries consultant

New Zealand

***For* University of Miami Independent System for Peer Review**

8 June 2007

Executive summary

A STAR Panel met from May 21-25, 2007 in Portland, Oregon, to review assessments of blue rockfish, black rockfish (Washington), and black rockfish (Oregon and California). The Panel consisted of three reviewers (two CIE) and an SSC representative as the chair. This is the report of one reviewer and it should be read in conjunction with the other review report and the STAR Panel reports.

The documents for the three assessments were in various stages of completion when they were distributed to meeting participants some ten days in advance of the meeting. In each case the STATs supplied altered documentation to the Panel at the start of the meeting and/or alerted the Panel to errors in the existing documentation.

The draft blue rockfish assessment was done using a non-equilibrium production model (ASPIC). The STAT had prepared data sets for use with SS2 but had been unable to achieve a working SS2 model. It appeared that ASPIC was used simply as a means of producing an assessment given the approaching deadline. The STAT argued that it was an appropriate tool given the available data but the arguments were not convincing. The Panel was sympathetic to the plight of the STAT but could not accept the hurried assessment as a suitable starting point for the week's activities. The STAT were not prepared to attempt SS2 runs during the meeting. The Panel was not prepared to accept an ASPIC based assessment *unless* it contained supporting SS2 runs which demonstrated that ASPIC was satisfactorily modeling the population dynamics. Given all the preparatory work that the STAT had done there was an expectation (on behalf of the Panel) that a satisfactory SS2 assessment could be achieved in time for the October mop-up meeting (and the Panel recommended that it be sent to the "mop up").

Two black rockfish assessments were presented at the meeting, one for a Washington stock, and one for a single stock off Oregon and California. Ideally, these assessments would have been done by the same STAT – to ensure a high level of consistency in the assumptions for each stock. As it was, it appeared that the two different STATs had had minimal consultation – there was some consistency in the division of catches (between states) for some periods – but in general they had made different assumptions with regard to biological parameters and data interpretation.

Both black rockfish assessments used SS2 to integrate a variety of data sources including length frequencies, mean length at age, and age frequencies. The Washington assessment incorporated biomass estimates from a tagging study (used as relative abundance) and a CPUE index from the same tagging study. The southern black rockfish assessment used CPUE indices from the recreational fisheries. Results from a tagging study were briefly presented but were not incorporated into the assessment.

The Washington black rockfish assessment came with the most complete documentation. However, the draft assessment contained errors and inappropriate data weightings. The STAT made various changes (and corrections) to their base run at the suggestion of the

Panel. The main impediment to achieving an assessment satisfactory to the Panel and the STAT was the STAT's firm belief, on the basis of the tagging study results, that the stock was not overfished (or equivalently, that the *q* on the tag abundance indices should be relatively low). The estimated *q* in the draft assessment was consistent with the STAT's belief but that from the final assessment was not. Sensitivity runs were made for low and high values of natural mortality (relative to the base) and for an alternative catch history (which used a proportion of the large rockfish catches in the 1940s).

The southern black rockfish assessment was poorly documented prior to the meeting but the STAT came with a working SS2 model, a well constructed collection of data sets, and a willingness to explore a wide range of models. The draft assessment model was for a single stock within four (explicitly defined) areas – with three fisheries defined in each area (commercial trawl, hook and line, and recreational). An extensively researched catch history went back to 1915.

The STAT and the STAR Panel worked together to try to produce a technically sound assessment. At the end of the week, after exploring many alternative runs, including single-area and two-area models, the STAT did not find a viable base model. The STAT's main concern was that the runs with satisfactory technical diagnostics had estimated exploitation rates that were inconsistent with the tagging study results (which suggested very low exploitation rates in the recreational fishery off Newport, Oregon). Given a base model was not available, the Panel recommended that the assessment go to the mop-up meeting (with an expectation that a satisfactory assessment can be achieved given the extra time).

The Washington black rockfish assessment was the only one completed during the meeting. The assessment was much improved by the STAR Panel process, but it is far from the best assessment that could be done for the stock. The uncertainty associated with the tag abundance *q* was not explored due to a lack of time and the assessment would be greatly improved if an informed prior was developed for the *q*.

The two failed assessments were a product of inadequate resources applied to the job in hand. In the case of blue rockfish the STAT were not given adequate training and technical support (or did not seek it in a timely manner). In the case of southern black rockfish too much was taken on by the STAT; their diligence in reworking the catch history, and using a spatially explicit model, made the task too big given the time they had budgeted.

Background

The blue and black rockfish STAR Panel met at the Pacific States Marine Fisheries Commission offices in Portland, Oregon, from May 21-25, 2007. This was the second of five 2007 STAR Panels in the biennial meeting schedule.

Three assessments were scheduled for presentation at the meeting, two black rockfish assessments and a blue rockfish assessment. The black rockfish assessments were undertaken by two different STATs with the stocks split on state boundaries (Washington, and Oregon and California combined).

The STAR Panel had four members. My three colleagues were: Dr Owen Hamel, the SSC representative and Panel Chair, Dr Tom Helser, and Dr Neil Klaer, my fellow CIE reviewer. This report should be read in conjunction with the STAR Panel reports and Dr Klaer's CIE report.

Review Activities

Pre-meeting

Meeting documents and materials were received in electronic form in advance of the meeting (*see* Appendix 2). I familiarized myself with the background material and current assessments prior to the meeting. Paper copies of the assessment documents were also made available at the meeting, which was helpful.

Meeting

The meeting was convened at 12.30 pm on Monday, May 21, 2007 and closed Friday afternoon, May 25, 2007. I will only give a brief summary of the meeting's activities, concentrating on blue rockfish. For black rockfish, details of the requests to the STATs and their responses are contained in the STAR Panel reports. I was the designated rapporteur for southern black rockfish, Dr Helser had blue rockfish, and Dr Klaer covered Washington black rockfish.

The first afternoon was devoted to blue rockfish, with the assessment presented primarily by Dr. Meisha Key with assistance from Dr. Alec MacCall. The assessment had been done only for northern California, with southern California excluded because of the loss of habitat in the area (a "kelp index" was presented which showed a large loss of kelp beds). A summary of the data that had been compiled for an SS2 assessment was presented including three alternative catch histories, numerous length frequencies and some age frequencies. No fishery independent abundance indices were available but two CPUE time series had been derived for recreational fisheries.

There was a brief explanation of the difficulties experienced trying to run SS2 with the input data. Almost all runs apparently failed because the model experienced “crash penalties” (associated with too little model biomass to satisfy the specified catches). No SS2 runs were presented and the STAT stated that they were not prepared to do any SS2 runs for the Panel. Instead they offered an assessment based on ASPIC runs. They justified the use of ASPIC but it was clear that had they got SS2 running they would not have been using ASPIC. In my opinion, the justification was a rationalization rather than a reasonable argument for why ASPIC was preferable to a model that could integrate all of the available data sources and properly capture an appropriate level of uncertainty. The meeting closed at 5.30 pm with a view to further discussing the blue rockfish assessment the following morning.

In the morning the inevitable conclusion was reached after brief discussion. The Panel could not accept the hurried ASPIC assessment (or whatever refinement of that assessment may have been possible during the week). It was not that the Panel could not accept *any* ASPIC assessment but that we could not accept the one that was offered without at least some supporting SS2 runs showing that ASPIC was adequately modelling the population dynamics. The Panel suggested that an acceptable SS2 assessment was possible and recommended that a new assessment be brought to the mop-up Panel in October. There was extra discussion on what the Panel thought was required in a new assessment.

The remainder of Tuesday and almost all of the rest of the meeting were devoted to the two black rockfish assessments. The southern assessment was presented by Dr David Sampson. The fish off California and Oregon (as far north as Cape Falcon) were assumed to be a unit stock. The draft assessment for Washington black rockfish was presented by a three-member STAT (Dr Farron Wallace, Dr Yuk wing Cheng, and Dr Tien-Shui Tsou). Dr Wallace presented the stock assessment modelling.

Southern black rockfish was presented first. The Panel made several suggestions for modifications to the base model and for exploratory runs. Then Washington black rockfish was presented, and likewise the Panel made requests for modifications and investigations. The meeting proceeded through to Friday morning with, generally, alternating presentations of results from the two STATs and associated supplementary requests from the Panel.

On Friday morning the Panel and the Washington STAT agreed that an acceptable assessment had been obtained, consisting of a base model and three sensitivity runs. On Friday afternoon, the Southern STAT concluded that he had no acceptable candidate for a base model. The STAT had worked tirelessly with the Panel, but we had run out of time.

Post-meeting

As rapporteur for southern black rockfish, I completed a draft STAR Panel report as soon as possible after the meeting. A partial draft was circulated by email to the other members of the STAR Panel on Friday evening. A complete draft was circulated after my

return to New Zealand (after a week in Chile on other fisheries business). Likewise, Dr Klaer circulated partial and complete drafts for Washington black rockfish. The black rockfish reports were revised via an email collaboration amongst the Panel (at the date of this report the final drafts are still at least a week away from finalization due to Panel members other commitments).

The blue rockfish report was essentially completed before the meeting closed, but it still needed to be circulated to other meeting participants and formally agreed by the Panel. Extensive and critical comments were received from the STAT on the report. At the request of the Chair, I briefly commented on the STAT's comments – standing by our findings and recommendations in the STAR Panel report.

Review findings

The findings for each assessment are discussed below under a variety of sub-headings. The "standard' headings are appropriate only for Washington black rockfish (for which an assessment was accepted). For this stock, I summarize the draft assessment and the changes that were made to reach the final assessment. I then summarize the main uncertainties and finally summarize the merits and deficiencies of the accepted assessment.

For southern black rockfish, a draft assessment was presented and many alternative models were explored but an acceptable base model could not be found. For this stock I summarize the draft assessment, the alternatives explored, and the reason why an acceptable assessment was not achieved.

The draft blue rockfish assessment was not extensively reviewed as there was no prospect of reaching an acceptable assessment during the meeting. I give a brief summary of the draft assessment and explain why the Panel quickly dismissed the blue rockfish draft assessment.

Blue rockfish

The assessment summary below is an edited version of that found in the STAR Panel report.

Assessment summary

A single northern Californian stock was assumed (north of Point Conception). There were recorded historical catches from southern California, but little catch in recent years. These were assumed to be from another stock (which has diminished in size as habitat has disappeared).

Fishery data for this assessment consisted of commercial landings from the CALCOM database, 1969-2006, and recreational landings from MRFSS from 1980-2003. For 2004-2006, recreational catches were taken from the California Recreational Fisheries Survey (CRFS), also available on the RecFIN website. The STAT provided three scenarios; 1) estimated catches, 2) fisherman supplied catch information for some years, and 3) intermediate catches between the two (being the average; this was the STAT's base case).

No fishery-independent abundance indices for blue rockfish were available. Instead, several fishery-dependent abundance indices were derived based on delta GLMs applied to catch and effort data from both the CPFV RecFIN and the CDFG CPFV data bases. Two pre-recruitment indices were presented based on the SWFSC mid-water trawl young-of-year data; one index from 1986-2004 based on the "core area" and the second from 2001-2006 based on the pooling of the NWFSC and SWFSC mid-water trawl data. In both cases, pre-recruit indices were developed by applying a GLM. Limited age composition data were available from the recreational fishery from 1979-1984 (based on 2300 otoliths). A small amount of recent age data was available from the groundfish ecology cruise in 2003-2004 (101 otoliths). Length composition data were available from recreational CPFV and private boats, 1980-2006.

The STAT presented model results based on the ASPIC surplus production model. The only data sources provided to the base model were fishery catches from 1969-2006 (intermediate catch series) and the RecFIN CPFV CPUE time series from 1980-2006. The Schaefer logistic production model, which consists of four parameters to be estimated, B1/K, K, MSY and CPUE *q*, was preferred over Pella-Thompson parameterization. The ratio B1/K was fixed at 0.77 (or greater) to achieve a model result that did not drive biomass to unrealistically low levels. Sensitivity runs explored different fixed levels of B1/K, the alternative catch series, and an alternative CPUE time series.

Reasons for dismissal

The STAT had compiled data sources for an assessment using SS2. However, they were unable to get any "sensible" results from SS2 model runs because of "crash penalties" (which occur when the model biomass is lower than that required to allow the specified catch to be taken). The STAT assumed that this was because the maximum likelihood estimate was close to the minimum level of biomass (necessary for catches to be taken). In my opinion, it is more likely that the STAT had one or more errors in their input files and/or had inappropriate starting values (in the minimization). The Panel could have explored the reasons for the SS2 failure but because the STAT was unwilling to do any SS2 runs during the meeting, there was little or no point in the STAR Panel attempting to correct their input files.

The STAT turned to ASPIC because they could not get SS2 to work. ASPIC was not their first choice. They presented reasons why an ASPIC assessment was acceptable and perhaps even preferable to an SS2 assessment but this was, in my opinion, just a *rationalization* of their actions. The assessment they were asking the Panel to consider

fell so far short of the standard expected from modern stock assessments that we had no choice but to reject it.

I found the ASPIC assessment unacceptable for three main reasons:

- ASPIC could not use much of the data that was available.
- It was not clear that ASPIC could properly mimic appropriate population dynamics for this stock.
- Although uncertainty could be captured with multiple runs it was clear that uncertainty on individual runs was grossly under-estimated.

An ASPIC based assessment could have been acceptable, but it would have required a demonstration that similar results would be obtained from a statistical model which integrated the available data sources. It would have needed to be a finely crafted assessment with an extensive set of ASPIC runs supported by *some* SS2 runs.

Black rockfish (Washington)

Assessment summary

A single stock was assumed for black rockfish distributed between Cape Falcon, Oregon, and Cape Flattery, Washington. The draft assessment used data from fishery and fishery independent sources. The catch history included three fisheries: commercial trawl (starting 1940), commercial non-trawl (starting 1950), and a recreational fishery (starting 1974). Length and age frequencies were available from 1976 in each of the fisheries. Tagging studies had been done since 1981 and two abundance indices were derived from the data: a CPUE index from 1981 to 2006 (with gaps) and Petersen tag and recapture estimates from 2000 to 2006. A CPUE time series had been derived for the recreational fishery but was not used because of concerns about the impact of regulation changes.

Input sample sizes for length and age frequencies had been calculated using a method recommended at the NMFS 2006 Stock Assessment Data and Modeling workshop (using the number of samples and the number of fish per sample). However, to improve the fit to the abundance indices and age frequencies, and also to compensate for the double use of some length data (in the age frequencies), all of the emphasis factors (lambdas) on the length frequencies were set to 0.01. To allow for the lack of old females in the observed age frequencies, the natural mortality (*M*) for females was assumed to "ramp up" from ages 10 to 15, going from 0.14 to 0.22. A constant *M* of 0.14 was assumed for males. The value of 0.14 was the value within the model that maximized the fit to the tag abundance indices. The value of 0.22 was "based on comparison to models external to synthesis". Steepness had been set at 0.7.

When the fits to the length frequencies were presented it was noted that the predicted length frequencies were extremely "spiky" and multi-modal. It was concluded that the CV of length at age must be very small. It was eventually discovered that the input CVs

had been coded as standard deviations (hence a CV of length at age of 0.08 became a standard deviation of 0.08 cm). The Panel also wanted to have some consistency between the two black rockfish assessments in terms of assumed biological parameters. Discussion was held with the STATs and it was agreed to use a common value of 0.6 for steepness, to have consistent ramp-up ages for female *M* (those used for Washington) and for each STAT to use the same values. There was agreement to have base model values of *M*: 0.14, ramp to 0.2; with sensitivities: 0.1, ramp to 0.16; and 0.18, ramp to 0.24. A consistent approach was also adopted with regard to the weighting of data sets. Length and age frequencies were to be iteratively re-weighted, except that lambda values were set to 0.1 for length frequencies that were double used.

There were other minor changes with some experimentation with the estimation of growth parameters and attempts to better fit the tag abundance indices with higher emphasis levels. It was found that the model simply could not achieve a good fit to the relatively flat but noisy abundance indices. Indeed, the best fit generally managed to miss almost all of the indices 95% confidence intervals (the predicted trajectory threading its way through the indices but missing the intervals).

The STAT was generally very responsive to the Panel's requests, and maintained good humor during the meeting, but there were two curious incidents. After the second set of requests the STAT presented results with a "new" base model. They had changed from the agreed base values of *M* to somewhat higher values (0.16, ramp to 0.22 rather than 0.14, ramp to 0.20). The main reason seemed to be that the model estimated exploitation rates were then more consistent with the (low) estimated exploitation rates from the tagging study. Also, after the fifth set of requests, the STAT proposed yet another new base model, this time with *M* at 0.18 ramping up to 0.22. They argued that the difference between male and female *M* at old ages was indicated to be about 0.4 from a study external to the model and that the level of 0.18 was the value obtained from a catch curve analysis done for some 1980 data. The STAR Panel pointed out that the catch curve was estimating *Z* rather than *M* and so a value less than 0.18 was appropriate. The STAT eventually agreed to use 0.16, ramp to 0.20, with sensitivities: 0.12, ramp to 0.16; and 0.19, ramp to 0.23. The STAR Panel was more comfortable with a base of 0.14, ramp to 0.20 as previously agreed.

The STAT was clearly unhappy with the estimated tag *q* from the successive base models that were examined during the meeting. In the original draft assessment the estimated tag *q* had been at about the level they were comfortable with (the STAT believed the value should be about 0.3 based on the proportion of total habitat covered by the survey). In later model runs the estimated *q* was typically at about 0.7-0.8. Technically, the problem was that the tag abundance indices were used as relative abundance with an *uninformed* prior, and so none of the information they contained on absolute abundance was available to the model. However, comparison of the estimated tag *q* with the STAT's belief was problematic because the biases associated with the tagging estimators had not been quantified in any way. I believe that the construction of an informed prior for the tag *q* is a priority if a robust assessment is to be obtained.

The final assessment consisted of a base model with *M*: 0.16, ramp to 0.20, steepness = 0.6, growth estimated within the model, iterative re-weighting of age and length frequencies, with lambdas set equal to 1 except for double-used length frequencies where lambda = 0.1. The error on the CV at length code was fixed as were some incorrectly specified CVs on the tag abundance indices. Sensitivities were done for lower and higher *M* and for an alternative catch history which included a large pulse of commercial trawl catches in the 1940s (the base catch history had almost no catch in this period).

The final assessment was technically much better than the original draft assessment but it could have been substantially improved had much more time been available.

Primary sources of uncertainty

Major uncertainties:

- The value of the tag *q* remains very uncertain (and hence so does absolute biomass).
- Natural mortality and steepness remain uncertain.
- Sex-specific natural mortality was assumed in all model runs; the alternative of sex-specific selectivity was not explored.
- The historical catch history is not well determined.

Strengths and weaknesses of current approach

The final assessment is at the lower end of what I would consider to be acceptable in a modern stock assessment. There is a base model and three sensitivity runs, none of which were taken forward to MCMC runs to obtain posterior distributions and an appropriate measure of within-run uncertainty (approximate confidence intervals will no doubt be obtained for each run using boot-strapping or likelihood profiles but these are poor substitutes).

The data associated with the tagging study was used to provide two relative abundance time series, which were assumed to be independent. There is some level of dependency between the tag CPUE and the Petersen tag estimates but there was no time to explore this at the meeting. Also, the main strength of the tag abundance indices is not in any trend they may show over time but in terms of the information they provide on *absolute* abundance. Using the tag estimates as a relative time series *without* an informed prior on the *q* dissipated the main value of the data.

Merits:

- SS2 was used and as such brings the advantages of a standard and well tested package.
- A fishery independent source of abundance data was used (tagging study).

Deficiencies:

- The tagging study data were not used appropriately: tag CPUE and tag abundance indices were assumed to be independent; the tag abundance indices had an uninformed prior on the *q*.
- Uncertainty in the tag *q* was not explored. An informed prior on *q* would greatly improve the assessment.
- Multiple use of some length data (in length frequencies, age frequencies, mean length at age) – assumed independent with *ad hoc* down-weighting.
- Sex specific selectivity was not explored as an alternative to sex-specific natural mortality.
- The full uncertainty of the catch history was not captured (only two alternatives were used with no attempt to bracket the true uncertainty).
- No MCMC runs were done.

Black rockfish (Oregon and California)

The assessment summary is taken from the southern black rockfish STAR Panel report (for which I was rapporteur).

Assessment summary

The draft black rockfish stock assessment for Oregon and California assumed a unit stock in the waters off Oregon and California, covering the PMFC statistical areas from 1B (in the south) to 2C (in the north). A recent version of Stock Synthesis 2 (SS2) was used to integrate data from many sources.

The draft assessment included an extensively re-worked catch history which extended back to 1915 (in contrast to the previous assessment which assumed zero catch before 1946). Explicit spatial structure was included in the stock assessment model with four areas defined by aggregating adjacent PMFC areas (Areas A & B are off Oregon, and areas C & D off California). Three different fisheries were defined in each area: recreational, commercial hook and line, and commercial trawl. CPUE indices were calculated from recreational fisheries data: RecFIN CPUE in all areas, and “secondary” indices from ORBS and CPFV in areas A, B, and D. Length frequencies were available for all areas, and age frequencies for areas A, B, and D. Other inputs included mean size-at-age data and a few mean weight indices from the commercial fisheries. Tag recapture estimates of exploitation rate for a portion of the population off Newport, Oregon were briefly presented but were not included in the stock assessment model.

Alternatives explored

The STAR Panel worked with the STAT to try to develop an acceptable assessment by progressively making changes to the draft "base" model. In all there were six written sets of requests (see STAR Panel report) and the STAT, who was diligently numbering runs, got to over one hundred runs.

The issues that most concerned the Panel and STAT were:

- The number of areas to use
- Allocation of recruitment across areas
- Relative weighting of data sets
- Validity of CPUE time series
- Natural mortality
- Catch history
- Time blocking of selectivities
- Estimation of growth

The draft assessment had four explicitly defined areas as an attempt to be more realistic. Also, there was a legitimate concern that there could have been different trends in biomass across areas. During the meeting one-area and two-area models were tried as alternatives. A two-area state-based model emerged as the preferred choice (but possibly only because it simplified interpretation of results for management). The number of areas to use is an important consideration. However, it only became an issue because the STAT came to the meeting with a four-area model. If they had just had a single area model the issue would never have arisen (but they would have been asked to check for area-year interactions in the CPUE indices).

Because the model had multiple areas there was also the issue of allocation of recruitment across areas. Most model runs were parameterized so that each area received a constant proportion of the annual recruits. This made it impossible for the model to fit flat trends in most CPUE time series and an increasing trend in the Area-D CPUE indices. In some of the later runs time-blocking of recruitment proportions was introduced and this gave a substantial improvement in fit (for relatively few extra parameters).

The relative weighting of data sets is always an important issue unless all data sets have the same "preference" for virgin biomass, depletion level, and recruitment patterns. The STAT was conscious of the need to iteratively re-weight ("tune") the runs to achieve a level of consistency between the input CVs and the "output CVs" (from the residuals). However, model results often varied dramatically between the un-tuned run and the fully-tuned run. Despite this, the fully-tuned run should be a starting point – not necessarily a finishing point (if there are some "reliable" biomass indices in an assessment it is often appropriate to up-weight such indices to ensure the assessment is consistent with the indices).

The assessment relies on recreational CPUE indices to provide some clues to abundance trends. However, there were several major regulation changes within the period covered by the indices. There was strong evidence in the data that changes in bag limits had

affected the distribution of bag sizes. The STAT and the Panel agreed to split most of the CPUE indices into blocks based on when regulation changes occurred. This was an *ad hoc* response forced on us given time constraints. This is a generic issue across most rockfish species and has been known for many years. Techniques can and should be developed to deal with the issue (*see* Recommendations).

Sensitivity runs using alternative values of natural mortality and an alternative catch history were waiting to be run should a suitable base model ever be found. (Alternative catch history runs were actually done in the hope that they might provide a suitable base model.)

For a number of fisheries there was clear evidence in the data of shifts in selectivity, so time-blocking of selectivities was introduced for some fisheries. Initial bad fits to some length frequencies were ascribed to not estimating enough growth parameters within the model. After some experimentation the STAT settled on length at age 3 years equal for males and females with estimation of sex specific k and L_{max}.

No base model

The STAT made runs using many alternative model configurations but did not produce any model runs which they were willing to put forward to the Panel as a base model. The STAT's primary diagnostic was a "reality check" on the estimated exploitation rates from the Newport tagging study. Model runs generally produced estimated exploitation rates for the Area-A recreational fishery that were perceived, by the STAT, to be too high relative to the Newport tagging study estimates.

I was not entirely convinced that the STAT's primary diagnostic was appropriate (nor were the other members of the Panel). Without a detailed study and evaluation of the potential biases in the tagging estimators, comparisons between the model estimated exploitation rates and the tag estimates are problematic. However, it was apparent that a carefully crafted assessment could not be put together by the end of the STAR Panel meeting – even if the STAT's diagnostic was ignored. We had simply run out of time.

I believe that an adequate assessment can be produced for the October "mop up" meeting, provided that sufficient time can be spent on the necessary work (*see* Recommendations for the next assessment).

Conclusions and Recommendations

The Washington black rockfish assessment was technically improved by the STAR Panel process. The other two assessments were rejected. The blue rock fish assessment could not have been made acceptable during the week without some supporting SS2 runs (which the STAT was unwilling to attempt). The southern black rockfish assessment, although not accepted, was progressed during the meeting. The STAT is now in a much better position to deliver a good quality assessment to the October STAR Panel.

I support the recommendations given in the STAR Panel reports, many of which are repeated below. The recommendations for blue rockfish and southern black rockfish are primarily aimed at guiding the STATs towards acceptable assessments for the October meeting.

Generic (all rockfish)

The first two recommendations below have been made in past years, possibly more than once, by CIE reviewers and/or STAR Panels. Action on them is long overdue:

- Establish a database for historical rockfish catch histories, "best" guesses and estimates of uncertainty (and processes for updating and revising the database).
- Develop fishery independent time series using fixed sites and volunteer fishers properly supervised using standard protocols.
- Develop a full descriptive analysis of the recreational fisheries and fleets for CPUE interpretation (not limited to "rockfish trips" – interactions with other target species are important).
- Develop standard and validated methods for producing recreational CPUE indices which deal with the peculiarities of the recreational data and regulation changes.
- Develop informed priors for recreational CPUE *q*s (*see* Appendix 1).
- Map rockfish habitat – quantitative estimates of area (which will inform CPUE *q*s and tagging *q*s).

Black rockfish (Washington)

Research recommendations:

- Ageing validation and between-state reader comparisons.
- Development of an informed prior on the tag *q*.
- Investigation of the dependence between the tagging abundance estimates and the tagging CPUE indices.

Black rockfish (Oregon and California)

Recommendations for the next assessment:

Essential elements:

- Informed prior on tag *q* (with the tag estimates input to the model as abundance).
- Catch history uncertainty to be fully captured (alternative catch histories should bracket the uncertainty).
- Descriptive analysis of CPUE – to justify, to some extent, the use of CPUE indices as indices of abundance.
- More GLM diagnostics – separate binomial and positive-catch indices– residual plots – QQ plots – sensitivity on number of co-caught species used (currently top 10 – perhaps 8 or 12).

Desired elements

- Explore alternative stock hypotheses (e.g., split off Area D).
- Continue exploration of using multiple areas.

Research recommendations

- Ageing validation and between-state reader comparisons.

Blue rockfish

Recommendations for the next assessment:

Essential elements:

- Provide explicit documentation and justification for assumed stock hypotheses.
- Fully capture the uncertainty in historical catch. At least three alternative catch histories should be constructed: a "best guess", an upper bound and a lower bound. Alternative assumptions in the timing of small and large catches could also be explored. "Ramping up" of catches from the start of the fishery is far preferable to assumptions of constant historical exploitation.
- Provide a detailed justification for the use of fishery CPUE indices as indices of abundance. Provide a detailed descriptive analysis of the data used, with particular attention to annual changes that affect the fundamental assumptions.
- Provide a more complete suite of GLM diagnostics, including separate binomial and positive-catch indices. Potential year-area interactions should be explored.

Desired elements:

- Document the historical blue rockfish catch in Oregon.
- Catches should also be considered in Southern California. An alternative hypothesis might lead to consideration of a two area model that has depletion in the south (due to fishing and/or loss of kelp). A kelp index could be included in the model.
- Exploration of potential temporal changes in recreational CPUE q should be evaluated, particularly in relation to regulatory measures.

Appendix 1: Development of an informed prior for a CPUE proportionality constant

The development of an informed prior for an abundance-survey proportionality constant (q) is relatively common in New Zealand (e.g., see hoki and orange roughy stock assessments in Sullivan et al., 2006). A prior is often useful to help stabilize stock assessment results and, in a full Bayesian assessment, provides a natural method for incorporating ancillary information into an assessment. Also, comparison of the estimated q with the prior provides a useful diagnostic for point-estimate assessments or full Bayesian assessments (posterior compared with prior). Informed priors for CPUE qs have never been developed in New Zealand, but there is no theoretical reason why they should not be.

For assessments that depend largely on CPUE indices for abundance information an informed prior on a CPUE q could be very useful for ground-truthing assessment results. The equations of a simple model which could be used to develop CPUE q priors are given below. Not all details are covered – this is the presentation of a concept rather than a definitive method.

Let X be a CPUE abundance index in a given year for a given species and area. Assume that it is part of a time series (GLM standardized or not) and that the units of the catch rate have been retained (e.g., numbers per angler hour).

By definition,

$$\mathrm{E}(X) = qN$$

where N is the total number of fish in the vulnerable population (i.e., the fish selected by the associated fishery). Further, assume that the CPUE index is proportional to density:

$$\mathrm{E}(X) = \alpha d$$

where d is the average density across "fishing spots" (i.e., the specific areas which are fished) and α is a proportionality constant. Note the distinction between q and α; they are both unknown proportionality constants, but one relates density to catch rate and the other relates catch rate to population numbers. We need to express q in terms of its components – which we know something about – in order to develop a prior for q, and α is one of those components. The other main component is the area occupied by vulnerable fish.

Let,

A = total area of fishing spots
D = total background area (areas not fished, but which contain vulnerable fish)
b = average background density where $b = \beta d$.

Then,

$$N = dA + bD = d(A + \beta D)$$

and

$$E(X) = \alpha d = \frac{\alpha N}{A + \beta D}$$

Hence,

$$q = \frac{\alpha}{A + \beta D}$$

The denominator in this equation appears tractable. Certainly something is known about the area of the "total habitat" ($A + D$) and the area fished (A). Also, it is not too difficult to obtain suitable experimental data on the relative densities found in the "fishing spots" and the "background" (using the specified fishing method).

The numerator appears to be more difficult. How does catch rate (in a fishing spot) relate to the underlying density? Clearly α is a function of several variables and could be highly species specific. Certainly, the relationship between density and catch rate will vary, even for a given species, by time of day and season and many other factors. However, α relates an average density (over all fishing spots) to an expected catch rate for an associated CPUE index (so daily and seasonal variation are not a particular concern).

One way to explore potential ranges for α is through a simulation study. It might be possible analytically but it would be much easier to simply simulate fishing under a number of different conditions – e.g., density, clusters of lines and hooks, biting probabilities, "effective hook volumes" - and examine the relationship between catch rates and fish densities. Depending on the sub-model used, it may be that information/opinions on values of the sub-model parameters could be available.

There are at least two alternatives which could be supplementary to or used instead of such a simulation study. First, it may be possible to use a depletion experiment design (which need not be destructive - perhaps some/most fish could be retained alive in tanks and later returned to the fishing spot). Second, there may be some comparable species which have reliable assessments which include CPUE indices – and the estimates of their CPUE *q*s could be "borrowed" (this could be possible for *q* if the areas are comparable, otherwise it could be done for α if there is information on the habitat area for the comparable species).

References

Sullivan, K.J. et al. 2006. Report from the fishery assessment plenary May 2006: stock assessments and yield estimates. (Draft document available from the N.Z. Ministry of Fisheries.)

Appendix 2: Bibliography of supplied material

I. Current Draft Stock Assessments

A. Assessment of Blue Rockfish (*Sebastes mystinus*) in California.
Meisha Key, Alec D. MacCall, Debbie Aseltine-Neilson, and Kirk Lynn.

B. Status of the black rockfish resource off the Washington Coast in 2006.
Farron R. Wallace, Yuk Wing Cheng, and Tien-Shui Tsou.

C. Status of black rockfish off Oregon and California. David Sampson.

II. Background Materials

A. 2006 Workshop Summary Reports

1. A Summary Report from the NWFSC Bottom Trawl Survey Workshop held October 31 – November 2, 2006 in Seattle, Washington. NOAA Fisheries, NWFSC, FRAM Division.
2. A Summary Report from the WC Groundfish Data/Modeling Workshop held August 8-10, 2006 in Seattle, Washington. NOAA Fisheries, NWFSC, FRAM Division.
3. Report of the Groundfish Harvest Policy Evaluation Workshop, Southwest Fisheries Science Center, La Jolla, California. December 18-20, 2006. A Workshop Sponsored by the Scientific and Statistical Committee of the Pacific Fishery Management Council
4. Pre-Recruit Survey Workshop. September 13-15, 2006. Southwest Fisheries Science Center, Santa Cruz, California. A Summary Report Prepared by Jim Hastie NOAA Fisheries, Northwest Fisheries Science Center and Stephen Ralston, NOAA Fisheries, Southwest Fisheries Science Center.

B. Previous Black Rockfish Assessments and STAR Panel Reports

1. The Status of Black Rockfish (*Sebastes melanops*) Off Oregon and Northern California in 2003, Stephen Ralston and E. J. Dick. 2003
2. STAR Panel Report for Black Rockfish, April 20-25, 2003.
3. The Status of The Black Rockfish Resources in 1999. Farron R. Wallace, Annette Hoffmann, and Jack V. Tagart. 1999.
4. Black rockfish STAR Panel Meeting Report, June 20-24, 2005.
5. STAR Panel Report 1999. May 24-28, 1999
6. Black Rockfish STAT Executive Summary. 1999. Farron R. Wallace, Annette Hoffman, and Jack V. Tagart

C. SS2 Model Related

1. SS2 Zip File – includes User's Manual, example files, and powerpoint presentations
2. R Software Zip File – Code developed by Ian Stewart to perform model diagnostics and plotting of SS2 output. This is not an official SS2 add-on and is not part of the NOAA toolbox. File contains User's Guide, example files as well as powerpoint presentations.

D. Terms of Reference (TORs) for the West Coast Groundfish Stock Assessment and Review Process for 2007-2008. The Scientific and Statistical Committee (SSC) of the Pacific Fishery Management Council. 2006.

E. GAO Report: Pacific Groundfish: Continued Efforts Needed to Improve Reliability of Stock Assessments. United States General Accounting Office, Report to Congressional Requesters. June 2004.

F. Coastwide Pre-Recruit Indices from SWFSC and PWCC/NWFSC Midwater Trawl Surveys (2001-2006). Stephen Ralston. April 6, 2007.

G. Estimation of black rockfish (*Sebastes melanops*) exploitation rates from mark-recapture data for the recreational fishery off Newport, Oregon. Troy V. Buell, Steven J. Parker, and Robert W. Hannah. 2007. (Provided by David Sampson).

III. Meeting Materials

A. Draft Agenda
B. STAR Panel Meeting Location Information
C. Driving Directions to PSMFC
D. List of STAR Panel Participants

Appendix 3: Statement of work

Consulting Agreement between the University of Miami and Patrick Cordue

Statement of Work

April 21, 2007

General

The Stock Assessment Review (STAR) meeting is a formal, public, multiple-day meeting of stock assessment experts who serve as a peer-review panel for one or more stock assessments. External, independent review of West Coast groundfish stock assessments is an essential part of the STAR panel process that is designed to make timely use of new fishery and survey data, analyze and understand these data as completely as possible, provide opportunity for public comment, and assure the best available science is used to inform management decisions.

The stock assessments will report the status of the black rockfish and blue rockfish resources off the west coast of the United States using age and/or size-structured stock assessment models. Specifically, the information includes a determination of the condition and status of the fishery resources relative to current definitions for overfished status, summaries of available data included in the models, and impacts of various management scenarios on the status of the stocks. The information is provided to the Pacific Fishery Management Council and NOAA's National Marine Fisheries Service to be used as the basis of their management decisions, which are subsequently approved and disseminated by the Secretary of Commerce through NOAA and NMFS.

The consultant will participate in the Stock Assessment and Review (STAR) Panel of the Pacific Fishery Management Council (PFMC) for the review of the black rockfish and blue rockfish stock assessments. The consultant should have expertise in fish population dynamics with experience in the integrated analysis type of modeling approach, using age-and size-structured models, use of MCMC to develop confidence intervals, and use of Generalized Linear Models in stock assessment models.

The Pacific Fishery Management Council's Scientific and Statistical Committee requests that "all review panelists should be experienced stock assessment scientists, i.e., individuals who have done actual stock assessments using current methods. Panelists

should be knowledgeable about the specific modeling approaches being reviewed, which in most cases will be statistical age- and/or length-structured assessment models" (SSC's Terms of Reference for Stock Assessments and STAR Panel Process for 2007-2008)

Documents to be provided to the consultants prior to the STAR Panel meeting include:

- Current drafts of the black rockfish and blue rockfish stock assessments;
- Most recent previous stock assessments and STAR panel reports for black rockfish (this is the first assessment for blue rockfish);
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer);
- The Terms of Reference for the Stock Assessment and STAR Panel Process for 2007-2008;
- Summary reports from the West Coast Groundfish "Off-Year" stock assessment improvement workshops held in 2006;
- Stock Synthesis 2 (SS2) Documentation; and
- Additional supporting documents as available.

Specifics

Consultant's duties should not exceed a maximum total of 14 days: several days prior to the meeting for document review; the 5-day meeting; and several days following the meeting to complete the written report. The report is to be based on the consultant's findings, and no consensus report shall be accepted.

The consultant's tasks consist of the following:

1) Become familiar with the draft stock assessments and background materials.
2) Actively participate in the STAR Panel to be held in Portland, Oregon from May 21-25, 2007. Participants are strongly encouraged to voice all comments during the STAR Panel so the assessment teams can address the comments during the Panel meeting.
3) Comment on the primary sources of uncertainty in the assessment.
4) Comment on the strengths and weaknesses of current approaches.
5) Recommend alternative model configurations or formulations as appropriate during the STAR panel.
6) Complete a final report after the completion of the STAR Panel meeting.
7) No later than June 8, 2007 submit a written report consisting of the findings, analysis, and conclusions (see Annex I for further details), addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Die, via e-mail to ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani, via e-mail to mshivlani@rsmas.miami.edu.

Submission and Acceptance of Reviewer's Report

The CIE shall provide via e-mail the final reports of the consultants in pdf format to Dr. Lisa L. Desfosse for review by NOAA Fisheries and approval by the COTR, Dr. Stephen K. Brown by June 22, 2007. The COTR shall notify the CIE via e-mail regarding acceptance of the report. Following the COTR's approval, the CIE shall provide the COTR with pdf versions of the final report.

ANNEX 1: Contents of Panelist Report

1. The report shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report shall consist of a background, description of review activities, summary of findings (including answers to the questions in this statement of work), and conclusions/recommendations.

3. The report shall also include as separate appendices the bibliography of all materials provided by the Center for Independent Experts and a copy of the statement of work.