

# **Black Rockfish (Northern)**

## **STAR Panel Report**

Pacific States Marine Fisheries Commission

Portland, Oregon

May 21-25, 2007

### Reviewers:

Owen Hamel (Chair), Northwest Fisheries Science Center and SSC  
representative

Thomas Helser, Northwest Fisheries Science Center

Patrick Cordue, Center for Independent Experts

Neil Klaer (Rapporteur), Center for Independent Experts

### Advisors:

Kelly Ames, GMT representative

Kenyon Hensel, GAP representative

### STAT Team Members Present:

Farron Wallace, Washington Department of Fish and Wildlife

Yuk Wing Cheng, Washington Department of Fish and Wildlife

Tien-Shui Tsou, Washington Department of Fish and Wildlife

## Overview

A draft assessment of black rockfish (*Sebastes melanops*) off the Washington coast was reviewed by the STAR Panel. This assessment used a recent version of the SS2 model. A Petersen tag and recapture study that was explicitly modeled within the previous assessment was included this time as providing a relative abundance index. During the review a number of alternative model configurations were explored that incorporated changes including using the correct CV\_growth\_pattern in the control file to allow correct interpretation of CV on length at age, alternative catch histories, freeing growth parameters, using a steepness value of 0.6, adding adjustments to the CV on tag abundance, removal of early tagging length composition data, freeing up peak parameter for selectivity for all fisheries, using a base value  $M$  male 0.16 and ramp to 0.2 for old females, setting  $\lambda$  values to 1 (except length compositions), adding 1983/84 trawl mean size at age data and re-weighting  $\sigma_r$ , length and age compositions.

Biological features unusual to this stock were discussed, including the lack of old females in population samples compared to numbers of males. It may be that females provide sustenance to the young and therefore have a “harder” life than males, and are therefore killed off more quickly than males. Alternatively, there may be a sex-specific selectivity difference with old females becoming less available to the fishery. In short, modeling methods to deal with these alternatives methods for dealing with older females may be termed “kill them or hide them” methods.

Modeling selectivity separately by sex is managed in SS2 using offset values, so the previous method using a change to a higher  $M$  for older females (kill them) is the only option that has been explored at present. Sex-specific selectivity (hide them) should be pursued as an option in future. The STAT also pointed out that black rockfish may have unusual breeding habits where about 10% of the older females don't appear to spawn in any year.

Input data are available from three main fisheries – commercial trawl, commercial non-trawl and recreational sport fishery. Known catches commence in 1963 for trawl, 1970 for non-trawl and 1975 for the sport fishery. It is known that the species was caught back to at least the 1940s, so historical catches were reconstructed by assuming a linear increase from 1940 to the 1964-65 average for trawl and to 1974 for the sport fishery. The non-trawl fishery was assumed to commence with a linear increase from 1950 to 1969. Particularly in early years, black rockfish were not identified at the species level in catches, and were recorded as part of a combined catch of all rockfish. Ratios from periods where the black rockfish fraction of the catch was known have been applied to unknown periods for each fishery. Some of this procedure was presented by the STAT, but a complete detailing of all of the assumptions made to generate the historical catch series is required.

Size and age composition samples commencing in 1976 are available for each of the fisheries. The Panel noted that the size samples often include the same fish as in the age samples, so there is not complete independence of these series. Results from SS2 model

presented by the STAT show a large 1999 year class that is now 8 years old, forming a central portion of the fishery. By the mid-1990s length compositions and age compositions from the sports fishery show a definite truncation of older age classes indicating an impact of fishing.

Abundance indices are available from a tagging program that commenced in 1981 as Petersen tag and recapture estimates and a CPUE series from the tagging effort is also available.

Statistical methodologies for deriving the Petersen estimates from tag-recapture, sex-specific length-weight and age-length relationships, aging error, age-weight conversion errors, age-length-maturity relationships, total mortality and natural mortality were presented. The Panel noted that there was a residual pattern in the fitted relationship used to estimate tag loss for spaghetti tags, suggesting a non-linear relationship. Also, in fitting fecundity, the model has a positive intercept, so is not strictly proportional to weight. The Panel suggested that effort used in the M estimation should be from all sources of mortality, and not just the sport fishery, and that there is also an element of double use of the data if these estimates are used in the assessment. However, total and natural mortality estimates from these procedures were not used in the assessment, and were provided for information and comparison with estimates from other sources.

The tagging program is carried out off Westport by volunteers, and the effort measure is the number of rod hours. Tagging is mostly done before the commencement of the sports fishery each year. Recaptures are from the wider sports fishery. Although several boats were probably used early in the program, most of the tagging is done from a single vessel. The region tagged is the same each year, but not the exact positions. From 1998 onwards the effort was distribute according to known black rockfish habitat, but before that was across all areas. The Panel noted that Petersen population estimates are from the same tagging effort that produces the tagging CPUE, so there is possibly a problem with independence of these two series.

The STAT thought that the  $q$  value for the tag Petersen index should be about 0.3 or less, as the survey covers about that portion of the available habitat along the central Washington coast, but in models presented,  $q$  was estimated and the index is used as relative index. The CV for the tag index used in the model was 0.6, and the index values were in numbers of fish. Calculated values for the survey CV range from about 0.1 to 0.25. The distribution of the recapture fleet changes through time due to economic factors. The Panel noted that it is questionable whether the assumption of mixing between tag and release holds depending on how far the tagged fish move, and the extent of overlap between tagging and release fishing effort. It is not possible to determine from returns where the fish were caught. The STAT pointed out that 80% of fish move less than about 10 miles. The Panel noted that it would be worthwhile to carry out a study to determine whether there has been any trend in the recapture fleet that may cause a bias in this index.

Selectivity for tag release is different to the sport fishery because fishing is shallower in the water column to avoid barotrauma. The release selectivity is showing as more dome shaped in the stock assessment than the sport fishery. It may be that the sex ratio is affected by this as well.

A CPUE index is also available from the sport fishery. The STAT presented results of standardization of the sport fishery CPUE using a delta lognormal GLM, but did not use this index in the assessment as they regarded it as not reflecting abundance due to the effects of changes in bag limits and a switch to bait fishing in the early to mid-1990's.

The coast wide recruitment survey has not been used as there are only 6 years of data available from this source, which the STAT considers too limited to use at this stage.

At the end of a series of requests and responses a base case model was produced that was acceptable, but with a number of deficiencies. The index for tagging abundance was noisy and the trend almost missed all confidence intervals of the observations. Effective  $q$  for the tagging index was 0.83 and the STAT thought that this was perhaps twice what it should be. The Panel pointed out that the SS2 value of  $q$  is a function of selectivity which is strongly dome shaped for the associated fishery. Without an objective evaluation of an informed prior on  $q$  it is difficult to compare a prior conception of  $q$  based on tagging and the one estimated by SS2. The Panel and STAT agreed that this was the best assessment available at the moment, but there are reservations about the  $q$  for the survey and that this dimension was not explored. The STAT was content to proceed with this base case. They also agreed to use a set of low and high  $M$  values and alternative catch history for sensitivity testing.

## Requests and responses

There was Washington catch landed in Astoria in the 1940s that may have contained large catches of black rockfish based on anecdotal information from Cleaver. The Panel was concerned that the current reconstruction of historical catch does not capture any of this uncertainty and suggested that as a first step, an alternative catch history be developed that accounts for such a potentially large historical catch, and that historical catches may have commenced in about 1915.

For both the trawl and sport fishery to some extent, there is a general underestimation of fish at older ages. The peak parameter for selectivity has been fixed, so these results suggest that it needs to be estimated.

Expected length frequencies show marked spikes, indicating the CV on length at age needs to be increased (The CV\_growth\_pattern had been set to 3 instead of 0).

The model has difficulty fitting length compositions from the tagging fleet prior to about 1990, and the Panel suggested that they might need to be down-weighted or disregarded.

Models presented had recruitment  $\lambda$  set to 0.1 and  $\sigma_r$  set to 0.55 which was the RMSE for one of the model runs after using an initial value of 0.6. The Panel suggested setting all  $\lambda$  values back to 1 and re-weighting  $\sigma_r$  based on the RMSE value from the same model.

Best likelihood values for  $M$  were high at 0.2 for males and 0.26 for females. Best fits to Petersen tag abundance only were for lower values of  $M$ , but the STAT thought that those values for  $M$  seemed unrealistic. The Panel suggested that  $M$  was a primary source of model uncertainty, and that it might be possible to select a range of  $M$  values that could be used for sensitivity testing that could be the same for the northern and southern black rockfish assessments. The Panel also noted that  $M$  values used for southern black rockfish were generally lower than those used for northern.

The Panel suggested value of 0.6 instead of 0.7 for steepness for consistency with the southern black rockfish assessment.

## Requests (1):

The STAR Panel requested a new base case and some sensitivity runs as follows:

Base Case:

- a) Increase CV on length at age (change CV\_growth\_pattern in the control file to 0)
- b) Investigate freeing  $l_{\min}$ ,  $l_{\max}$  and  $K$  for growth
- c) Set steepness to 0.6
- d) Add 0.2 as an adjustment to the calculated CVs on tag abundance
- e) Free up peak parameter for selectivity (and perhaps fix other appropriate ones) for all fisheries
- f)  $M$  ramp from 10 to 15 for females (no change).  $M$  male 0.14 and ramp to 0.2 for females
- g) All  $\lambda$  values set to 1
- h) Include 1986 and 1987 trawl mean size at age
- i) Re-weight  $\sigma_r$ , length and age compositions. Calculate sd of the Pearson residuals for age and length frequencies.

Sensitivities:

- a) Remove length compositions and CPUE for the tagging fleet to 1990
- b) Low  $M$  of 0.1 ramping to 0.16 and high  $M$  of 0.18 ramping to 0.24.
- c) An alternative catch history is to be developed that accounts for higher trawl catch in the 1940s.

Response to Request (1)

A modified base case was presented that did not include base case options (b) or (i). For base option (e), the peak was freed on trawl, survey and sport with a fixed width. Non-trawl was not freed. It still showed a lack of fit to older age classes. Expected length frequencies now look normal.

Option (a) and (b) in the sensitivities were not yet explored.

An alternative catch history (c) was constructed. It included 90% of the Astoria landings from 1936 to 1950 plus 10% of the rockfish catch from the trawl fishery off Washington in the 1936 to 1950. Catches from 1915 to 1936 were set to 0. Initial fishing mortality was set to 0. Using this catch series in SS2 does not alter the initial biomass or current depletion substantially as there was sufficient time since the large trawl catches for the population to recover.

## Requests (2)

Base:

- a) Free  $l_{\min}$ ,  $l_{\max}$  and  $K$  for growth
- b) Down-weight length compositions to better fit the tag abundance index, if it won't fit, reduce the index CV.

Sensitivities:

- a) Remove early length compositions and associated CPUE
- b) Low  $M$  of 0.1 ramping to 0.16 and high  $M$  of 0.18 ramping to 0.24.

Response to requests (2)

Trawl size composition for 2002 was removed, and also 1987 mean size at age for trawl because these sizes were much larger than those seen even in 1986 and were difficult to fit.

No convergence problems were experienced for the base case and the jitters also worked.

Fits to low natural mortality were not as good as other scenarios (and the hessian didn't invert for low). Also did  $M$  0.16 and 0.22 as an alternative base case. The STAT thinks that higher  $M$  values are more plausible as they better match the fishing mortality rates off Newport indicated by tagging. The STAT is essentially using the  $q$  for the tagging index as a diagnostic reality check, which the Panel suggests would be better implemented as an informed prior.

Removal of early tagging length composition data improves the tagging abundance and CPUE index fits. Estimation of  $K$  improves the fit to age compositions but not length. The overall likelihood was improved substantially through estimation of  $K$ .

The base case is still not fitting relatively narrow peaks in observed female age compositions, but fits to older females generally improved.

The sport fishery lengths do not fit the mode prior to about 1995 when there was a regulation change, so time blocks for selectivity might improve the fit.

The STAT doesn't believe that there is any good reason to leave out the early tagging data.

### Requests (3)

Base case:

- a) Male  $M$  0.16, old female  $M$  0.22
- b) Free  $l_{\min}$ ,  $l_{\max}$  and  $K$
- c) No removal of early tagging data
- d) Trawl mean size at age data included.

Sensitivities:

- a) Low  $M$  of 0.12 ramping to 0.18 and high  $M$  of 0.19 ramping to 0.25.
- b) Free up parameters for trawl selectivity
- c) Increase weight on tag abundance index
- d) Larger historical catch.

Response to requests (3)

Earlier  $K$  was mis-specified. Model fits age and size at age are now better than earlier base, but worse fits for abundance indices and length comps. The overall fit is however improved.

There is a tradeoff in fit between the 1986 and 87 mean size at age and the length frequencies for the trawl fishery.

Trawl selectivity is tending towards a gradual increase from small to large which seems implausible.

There is conflict between fitted growth for recent and earlier periods.

A high weight ( $\lambda=350$ ) was applied to tag abundance resulted in no significant improvement to the fit to tag abundance.

### Request (4)

Base case:

- a) Don't include trawl mean size at age
- b) Fix trawl selectivity width.

Response (4)

The overall fit was improved, although with worse fit to trawl length frequencies.



### **Request (5)**

Base case:

- a)  $M$  0.16 ramping to 0.22
- b) All tag data included
- c) Don't include trawl mean size at age data.

Response (5)

With an input  $\sigma_r$  of 0.5, the RMSE is 0.35. This model produces  $q$  values of 0.737 for the tag abundance. The STAT team believes that this value is too high and should be in the order of 0.2 to 0.4.

### **Request (6)**

New base case as above with  $\sigma_r$  0.30.

Sensitivities:

- a)  $M$  low 0.12 to 0.18, high 0.19 to 0.25
- b) Larger historical catch.

Response (6)

Overall likelihood across  $M$  has tightened. Other runs were presented by the STAT that reduced the ramp to 0.04, and with a range of male  $M$  values of 0.14, 0.18, and 0.21. The natural mortality analysis presented by the STAT earlier indicated that the spread should be about 0.04. A value of 0.18 was the indicated  $Z$  for 1980. The Panel was more comfortable with lower  $M$  values due to the longevity of the species. The current  $q$  value for the tagging abundance is coming out at about 0.7 and the STAT believes that value should be 0.3 based on the fraction of the area where the survey is carried out. The Panel would be happy to include an informed prior on  $q$  based on an analysis, but such an analysis to develop an informed prior has not been done. The STAT feels that the stock is not overfished, and the Panel preferred lower value of  $M$  produces an overfished stock which is implausible. The Panel suggests that this is not necessarily a problem, and that a range of  $M$  values should capture the range of uncertainty. There was a discussion about the role of the Panel and what level of guidance in development of base cases can be imposed, and how much the assessment becomes a product of the Panel. The STAT also felt that higher  $M$  values better match those used in previous assessments and also those produced from catch curve analyses from 1980. The Panel pointed out that these were  $Z$  values, and therefore the  $M$  value should be lower as the stock was not unexploited at the time. The STAT agreed to use the suggested range of lower  $M$  values for sensitivity but to modify the difference value to 0.04 based on STAT analyses.

## Request (7)

Base case:

$M$  0.16 (males and young females) ramping to 0.20 (old females)

Sensitivities:

- a) Low  $M$  of 0.12 ramping to 0.16 and high  $M$  of 0.19 ramping to 0.23.
- b) Alternate catch series with base  $M$ s.

Response (7)

The index for tagging abundance is noisy and the trend almost misses all confidence intervals. Effective  $q$  for the tagging index is 0.83 and the STAT thinks that this is perhaps twice what it should be. The Panel pointed out that the SS2 value of  $q$  is a function of selectivity which is strongly dome shaped for the associated fishery. Without an objective evaluation of an informed prior on  $q$  it is difficult to compare a prior conception of  $q$  based on tagging and the one estimated by SS2. The Panel and STAT agree that this is the best assessment available at the moment, but there are reservations about the  $q$  for the survey and that this dimension has not been explored. The STAT is happy to proceed with this base case and range of  $M$  values.

## Description of base model and alternative models used to bracket uncertainty

The following was the final base case and sensitivity tests agreed by the Panel and STAT.

Base Case (with reference to original draft base case):

- Increased CV on length at age
- Free  $l_{\min}$ ,  $l_{\max}$  and  $K$  for growth
- Steepness 0.6
- Include sport fishery mean size at age data from 2001 and 2002
- Free up peak parameter for selectivity, fix width for trawl
- $M$  ramp from age 10 to 15 for females.  $M$  male 0.16 and ramp to 0.20 for females
- All  $\lambda$  values set to 1.0 except for 0.1 for length compositions
- Re-weight  $\sigma_r$ , length and age compositions.

Sensitivities:

- An alternative catch history that accounts for higher trawl catch in the 1940s and catches back to 1915
- Low  $M$  of 0.12 ramping to 0.16, and high  $M$  of 0.19 ramping to 0.23.

## Comments on the assessment

The presented assessment was structurally quite different to the previous one for the same stock presented in 2003. The STAT is commended in their efforts to move the assessment into the SS2 framework, and the means used to retain tagging abundance and CPUE data within the assessment.

Merits:

- SS2 was used which brings the advantage of standards and a well tested package
- Tagging data has been brought into the model

Deficiencies:

- Tagging is not dealt with in the model as a tagging experiment (this is not possible with current SS2, but is being considered)
- Uncertainty in  $q$  was not explored. Uncertainty could have been expressed as a profile. The assessment would be improved if there was an informed prior on  $q$ .
- Non-independence of the length/age compositions
- Non-independence of the tagging abundance and CPUE series
- Sex-specific selectivity has not been explored as an alternative to elevated  $M$  for females as a means to produce less older females in the population
- The full uncertainty in the catch history has not been explored

## Explanation of areas of disagreement regarding STAR Panel recommendations

### A. Among STAR Panel members (including GAP and GMT representatives)

There were no areas of disagreement.

### B. Between the STAR Panel and STAT Team

There were no areas of disagreement between the STAR panel and the STAT team at the end of the STAR panel meeting. However, after the STAR panel meeting, the STAT produced an alternative proposed base case which is included in the assessment document. This alternative base sets  $M$  at 0.16 for males and young females, as in the base case agreed upon at the STAR panel, but ramps up to an  $M$  of 0.24 for old females (instead of 0.20). The rationale given for this alternative model is that the overall statistical fit is better and that the resulting  $q$  for the tagging study is closer to 0.3. The STAR panel did not have a chance to review this alternative model. It should be noted, moreover, that it was based upon the STAT recommendation that the difference between the male (and young female) and old female  $M$  should only be about 0.04 that the base model old female  $M$  was reduced from 0.22 to 0.20 towards the end of the STAR panel meeting.

## **Management, data, or fishery issues raised by the GMT or GAP representatives during the STAR Panel.**

No issues were raised.

## **Unresolved problems and major uncertainties**

- The major uncertainties are  $q$ ,  $M$ , historical catch and sex-specific selectivity.

## **Recommendations for future research and data collection**

The Panel reiterates research and data collection required to improve the assessments for all rockfish, and also makes specific recommendations for northern black rockfish.

### Generic (all rockfish) recommendations

- Development of fishery independent time series using fixed sites and volunteer fishers properly supervised using standard protocols
- Establish a database for historical rockfish catch histories, “best” guesses and estimates of uncertainty (and processes for updating and revising the database).
- 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.
- Mapping of rockfish habitat – quantitative estimates of area (which will inform CPUE  $qs$  and tagging  $qs$ ).

### Northern black rockfish recommendations

- Development of informed priors for tagging and recreational CPUE  $qs$  (see Appendix 1).
- Age validation study
- Reader to reader comparisons are needed between States (Oregon and Washington).

## 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  $q$ s 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,

$$E(X) = qN$$

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:

$$E(X) = \alpha d$$

where  $d$  is the average density across “fishing spots” (i.e., the specific areas which are fished) and  $\alpha$  is a proportionality constant. Note the distinction between  $q$  and  $\alpha$ ; 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  $\alpha$  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  $\alpha$  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,  $\alpha$  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  $\alpha$  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  $qs$  could be “borrowed” (this could be possible for  $q$  if the areas are comparable, otherwise it could be done for  $\alpha$  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.)