From:

Status of the Pacific coast groundfish fishery through 2008, stock assessment and fishery evaluation

Stock assessments, STAR Panel reports, and rebuilding analyses

October 2008

Pacific Fishery Management Council
Portland, Oregon

http://www.pcouncil.org/groundfish/stock-assessments/safe-documents/

Canary Rockfish STAR Panel Report

NOAA Western Regional Center
Building 9 Conference Room
7600 Sand Point Way, NE
Seattle, Washington 98115
July 30 – August 3, 2007

Reviewers:

Steve Ralston (panel chair), Scientific and Statistical Committee (SSC) Representative
Patrick Cordue, Center for Independent Experts (CIE, Rapporteur)
Devorah Hart, Northeast Fisheries Science Center (NEFSC)
Jim Ianelli, Alaska Fisheries Science Center (AFSC)
Paul Medley, Center for Independent Experts (CIE)

Advisors:

Brian Culver, Groundfish Management Team (GMT) Representative
Peter Leipzig, Groundfish Advisory Panel (GAP) Representative

Stock Assessment Team:

Ian Stewart, Northwest Fisheries Science Center (NWFSC)

Overview:

The canary rockfish assessment incorporated a variety of data sources into the candidate base model that was presented for review. Those data included landings and discards from trawl and non-trawl fisheries from southern California, northern California, Oregon, Washington, and bycatch in the at-sea whiting fishery. Extensive age and length composition data were also available for most modeled fleets. The principal abundance index used in the model was the triennial shelf trawl survey (1980-2004), although the NWFSC shelf/slope trawl survey (2003-06) was also included. Both surveys collected information on age and length composition. The SWFSC-PWCC/NWFSC coastwide pre-recruit index was evaluated but was initially excluded from the proposed base model.

A number of changes were evident between the last canary rockfish stock assessment conducted in 2005 and the STAT's initial candidate base model, which included:

- Completely revisited age estimation issues (laboratories, bias, and precision)
- Selectivity blocking based *a priori* on management actions (post-1995 only)
- Model-based, not design-based, estimates of survey abundance
- Incorporation of NWFSC shelf/slope survey (indices and compositions)
- Recruitment deviation vector extended back to 1916
- Evaluated effect of pre-recruit survey

Following the STAR Panel review a number of changes were included in the final base model:

- Incorporation of a trawl fishery selectivity block to account for the introduction of roller gear and high-rise nets in 1979
- Recruitment deviation vector starting in 1960
- The pre-recruit survey was included
- sigmaR (σ_r) was set equal to 0.5, based on partial tuning of recruit-deviations
- Due to differences in start dates of the triennial survey and serially correlated residuals in model fit, survey ***q*** was estimated separately for 1980-1992 and 1995-2004
- Fix steepness (***h***) using a revised meta-analysis prior (i.e., darkblotched rockfish removed).

Revisions to the model that were completed by the STAT before the STAR panel review had a marked effect on plausible values for steepness, with the age and length compositions favoring high ***h*** and the triennial survey favoring low ***h***. The Panel and STAT agreed that it was not possible to reliably estimate steepness from the available data. Consequently, steepness was fixed based on a meta-analysis of west coast *Sebastes* that excluded canary rockfish and darkblotched rockfish. Darkblotched rockfish was excluded because the likelihood profile that had been used in developing the prior was considered out of date, i.e., a more recent assessment resulted in a very different view of steepness.

Analyses Requested by the STAR Panel:

Round 1 requests

A: Exploration of triennial survey with regard to seasonal effects and time of day. Produce descriptive plots and/or tables.

Reason: There was a concern that the trend towards earlier start dates for the triennial survey could compromise the time series as an abundance index (e.g., availability and/or vulnerability of canary could vary seasonally).

Response: A plot of the daily distribution of tows within year was presented. It showed that the surveys fell into two blocks: mid July-mid September timing for 1980-1992; and June-mid August timing for 1995-2004. Within the second block there was a trend towards earlier start dates and finish dates with the 2004 survey being the earliest. Plots of catch rates and average catch rates were presented by day of year and time of day. Also, bottom temperature plots were shown.

Discussion/conclusion: The only point of concern was the change in timing of the surveys. However, it was not possible to tell from the data presented whether this was a problem or not. The plot of catch rate by day of year combined over all years showed an increasing trend, but this was because the early surveys covered the "high density" northern strata late in the year. It was suggested that a fixed main-effects GLM could potentially be used to tease out the various effects (year, area, stratum, month or day-of-year).

B: Plots/tables of mean length in the trawl fishery data.

Reason: Some Panelists wanted to see the extent of the changes in the different time series (as a broad indicator of depletion).

Response: Plots were produced for each fishery. Declines in mean length were seen in all of the fisheries up until about 1995, with declines greater in the south than the north.

Discussion/conclusion: Regulations impacting the fisheries after 1995 may have been sufficient to change selectivities to such an extent that trends in mean length after that period would not be indicative of depletion levels. Trends before that period may be indicative of total mortality and it was suggested that Beverton-Holt estimates of Z be produced from the mean length data for each fishery (see Request K).

C: Follow up on the at-sea hake data determining if any other data are available (e.g., mean weight and/or length before 2003).

Reason: Any additional data sources are of interest, if only for qualitative corroboration of model results.

Response: Mean weight data were presented from at-sea sampling of the hake fishery for

1990-2007.

Discussion/conclusion: The mean weight data showed no trend over the period. It appears that the largely mid-water fishery has always caught a few very large canary rockfish.

D: For the base model produce a likelihood profile over $\boldsymbol{R_0}$. Tabulate and graph likelihood components. Also, profile over $\boldsymbol{R_0}$ with steepness = 0.6 (mean of Dorn prior).

Reason: A profile over $\boldsymbol{R_0}$ is often useful for revealing "tensions" in the model, in terms of which data sets are better fitted with high or low biomass.

Response: Plots of likelihood components were produced (standardized so that each component had a minimum of zero).

Discussion/conclusion: As is common with many stock assessments, some tensions in the model were confirmed (from previous profiles on steepness) with survey abundance indices favoring lower $\boldsymbol{R_0}$ than length and age data. The length data showed the strongest preference for higher $\boldsymbol{R_0}$.

E: Explore the relative proportion of tows in juvenile and adult habitat in the NWFSC and Triennial trawl surveys.

Reason: The NWFSC survey catches a higher proportion of small fish than the Triennial survey. It was thought that part of the explanation could be a higher proportion of tows in "juvenile habitat".

Response: A cut-off depth was determined for "juvenile" habitat. Plots of the proportion of tows within depth strata in juvenile and adult habitat were produced for both surveys. Cumulative distributions of tow proportions relative to depth were also displayed.

Discussion/conclusion: The NWFSC survey had a somewhat larger proportion of tows within shallower depths than the Triennial survey. It was concluded that this could be partly, but not wholly, responsible for the larger number of small fish caught in the NWFSC survey.

F: Distribute status report on Canadian stock assessment.

Reason: Additional background material

Response: The report was distributed to Panel members who wanted it.

<u>Round 2 requests</u>

G: Turn off all priors and estimate survey $\boldsymbol{q}$s analytically (median unbiased option).

Reason: Simply a tidying-up exercise (to avoid unnecessary computations).

Response: The changes were made and, as expected, almost identical results were produced.

H: Start a block in 1979 in fishery selectivities to allow for transition to roller gear/high-headline nets (steepness = 0.6).

Reason: Steepness was set to the mean of the Dorn prior (***h*** = 0.6) as there was no other reasonable basis for choosing steepness. A new break-point for selectivities was introduced to allow for known changes in fishing gear (albeit, changes occurred over a number of years).

Response: The increased value for steepness resulted in lower estimated virgin biomass and a substantially higher estimate of current depletion. The extra selectivity block had only a minor effect.

Discussion/conclusion: As expected, the change in steepness produced a large change in estimated depletion. The basis for choosing steepness was further discussed. Concern about the prior was raised by the STAT and, in particular, the inclusion of darkblotched rockfish, for which steepness estimates had recently been substantially revised.

I: Add in the coastwide pre-recruit survey from 2001-2006.

Reason: To explore the effect of including the indices and to raise the issue of whether they should be in the base model.

Response: The indices were included. As expected they had little effect on estimated biomass trajectories, but did alter the estimates of recent recruitment (which could be important in projections).

Discussion/conclusion: Nobody objected strongly to the use of the indices. There was some concern that the time series was too short for the indices to be fully validated against model estimates of recruitment. However, it was noted that the four year NWFSC shelf/slope survey time series was included in the base model without validation. It was also pointed out that if the canary assessment was updated in two years, that if the time series was not included in the base model then it could not be used in the update (when the recruitment indices would contain two more points). However, the converse argument also applies: according to the update rules, the recruitment indices would have to remain in the base model (whether that seemed desirable or not at the time).

J: Try a range of initial sigmaR values (0.4, 0.8, and 1.2) and calculate tuned output values.

Reason: To explore the effect of using different initial values of sigmaR.

Response: The three runs were compared with a sigmaR-tuned model (input sigmaR = 0.3, output = 0.28). Higher input values of sigmaR gave higher output values and resulted in higher estimated virgin biomass and lower estimated depletion.

Discussion/conclusion: Although the alternative start values for sigmaR gave quite different results to the tuned model it was apparent that the tuned result was independent of the initial value of sigmaR. However, Panel members were not convinced that full tuning of sigmaR was appropriate as it tends to underestimate recruitment variability. It was suggested that a better approach was to use a high initial value and then simply fix sigmaR at its output value (i.e., partial tuning). (See Request P.)

K: Calculate Beverton-Holt total mortality (Z) estimates: $Z = K \cdot (L_\infty - l_{bar})/(l_{bar} - l_c)$, where l_c = full selection cutoff (pick by inspection where selectivity = 0.75), l_{bar} = mean length above l_c, K and L_∞ = von Bertalanffy growth parameters. Apply to all fisheries and surveys where there are adequate data.

Reason: Some Panel members were interested in seeing these estimates and using them as a diagnostic (vis-à-vis the plausibility of model estimates of exploitation rates).

Response: Plots of the annual estimates were produced by sex for each fishery at two different values of l_c.

Discussion/conclusion: The estimates varied by sex, fishery, and the value of l_c. The annual estimates generally showed an increasing trend up until about 1995, after which trends varied depending on the fishery. The Panel members who had requested these estimates were satisfied that they were consistent with the SS2 model results.

L: Convert the length frequency data from 2003-2006 for the hake fishery to weights and compare with mean weight plots.

Reason: To check if the weight samples in 2003-2006 were consistent with the length samples which had been included in the model.

Response: The length frequency data were converted to weights and shown to be consistent with the mean weights in those years.

Discussion/conclusion: The consistency of the mean weight and length frequency data raised the issue of whether the model predictions of mean weight for this fleet were consistent with the observed mean weights (see Request V).

M: Produce the "equilibrium yield" figure with a "real fishery" selectivity (rather than an "avoidance fishery"), steepness = 0.35. Also do with steepness = 0.6.

Reason: To see how the reference points for this assessment were affected by the selectivities and steepness used in the calculations.

Response: The plots were produced as requested (showing equilibrium yield vs relative depletion, with estimated depletion and depletion reference points marked). The big impact was from steepness, though the equilibrium estimates of yield were altered somewhat depending on the selectivities used.

Discussion/conclusion: The method of calculating reference points was discussed with particular reference to when and how the selectivities should be chosen for different calculations. The key point was to avoid using estimated selectivities from an "avoidance fishery" in calculations that were pertinent to a "target fishery". The STAT asked if people thought that the plot format was useful to managers and other users of stock assessments. There was general agreement that the plot format was useful.

N: Explore why the length data shows a strong preference for high $\boldsymbol{R_0}$. (Look at the likelihood components in the existing $\boldsymbol{R_0}$ profiles).

Reason: It was not clear why the length data should be fitted so much better with high $\boldsymbol{R_0}$.

Response: Likelihood components were presented for existing profiles on steepness.

Discussion/conclusion: This revealed/confirmed tensions between data sets but also within the length data (with different components pulling in different directions). Profiles on $\boldsymbol{R_0}$ would have been more revealing perhaps, but this issue was considered low priority.

O: Request a new canary prior for steepness from Martin Dorn.

Reason: The STAT had expressed concern that the existing prior contained a steepness profile for darkblotched rockfish from the 2005 assessment which was very different from that obtained in the 2007 darkblotched assessment.

Response: Martin Dorn supplied a new prior which excluded the 2005 darkblotched rockfish profile (the 2007 darkblotched profile was not available as the assessment was still being finalized). The new prior had a mean of 0.52 (the mean of the old prior was 0.6).

Discussion/conclusion: The removal of darkblotched as the basis for revising the prior was queried as other species which contributed to the canary steepness prior had also been assessed in 2007. However, the only substantial revision had been to darkblotched (Martin Dorn, pers. comm.). It was agreed that the revised prior was preferable to the old prior.

The STAT also produced results for a run which estimated recruitment deviations from 1960 rather than the start of the model. This request was anticipated by the STAT as a similar request for arrowtooth flounder had revealed that the assessment was very sensitive to the year in which recruitment deviations were first estimated. However, the canary results were not sensitive to this choice (for the assumed low value of sigmaR).

<u>Round 3 requests</u>

P: Form a new candidate base model:

- all priors uniform;

- $\boldsymbol{q}$s estimated analytically with median unbiased option;
- use pre-recruit time series;
- steepness = 0.511 (mean of middle 50% of new Dorn prior);
- recruitment deviations estimated from 1960;
- sigmaR chosen by using an input value of 1.2 and then setting it to the output value;
- re-tune if necessary.

Reason: The above specifications were consistent with the results of previous discussions. The level of sigmaR was not specified as it was thought that the other changes might lead to a somewhat different value than that previously obtained (by the specified method).

Response: The candidate base model was not constructed as requested because of an unexpected result for sigmaR. An input value of 1.2 returned an output value of almost 1 which was considered unrealistic by the STAT. The STAT performed a number of runs at a range of sigmaR values in combination with different start years for estimating recruitment deviations and presented the results graphically. On the basis of these results the STAT tentatively defined the candidate base model with a sigmaR value of 0.5. The candidate base model also included the selectivity break-point in the trawl fisheries at 1979 and in addition included estimation of the selectivity of the smallest length bin in the NWFSC survey.

Discussion/conclusion: The discussion concentrated on the value of sigmaR to use in the base model. The results presented showed that the output value of sigmaR depended on the input value and on the number of recruitment deviations estimated. With all deviations estimated an input of 1.2 produced an output of 0.5, but with deviations estimated from 1960, an input of 1.2 returned an output of almost 1.0. The attempt to get the "data to speak" by specifying a large input sigmaR had produced an unexpected result. From a pragmatic point of view it was decided that the STAT's choice of 0.5 was acceptable, being not too large relative to other values used for rockfish and not too inconsistent with the tuned value of 0.3. There is also some evidence in the literature that tuned values tend to under-estimate sigmaR.

Q: Sensitivity to candidate base model: no recruitment deviations estimated and steepness fixed; also with steepness estimated.

Reason: Given the inability to track cohorts "by eye" in either the length or the age data the issue arose as to whether the estimation of recruitment deviations could be justified in terms of an improved fit (i.e., did the addition of the extra parameters give a sufficient decrease in the total negative log-likelihood).

Response: The total negative log-likelihood for the sensitivity was larger than that of the base model by almost three times the estimated number of recruitment deviations. The estimated value of steepness was about 0.4.

Discussion/conclusion: By "rule of thumb", the estimation of the deviations was justified by the increase in likelihood. It was interesting to note that all of the improvement was in the length data. The age data had almost exactly the same likelihood. The estimate of steepness was not very different from that in the original base model proposed by the STAT, which had relied on likelihood from the Triennial survey. It was noted that there was a clear signal in the data with regard to steepness under the assumption of deterministic recruitment (structural assumptions can impart contrast in likelihood surfaces even in the absence of "real information" in the data).

R: Sensitivities to candidate base model: single ***M*** (no ramping, same for both sexes).

Reason: This was an initial step in the exploration of whether the lack of females in the observed data could be explained by selectivity in the absence of sex-specific mortality (with higher female ***M***).

Response: Two runs were presented with fixed values of ***M*** at 0.06 and 0.07.

Discussion/conclusion: The runs showed substantially degraded fits from the comparative run with ramped female ***M***. However, this was not unexpected and it was noted that sex-specific selectivity would be needed if comparable fits were to be obtained. (See Request X).

S: Sensitivity to candidate base model: split triennial times series into two blocks (1980-92 and 1995-2004.

Reason: A continuing concern about potential changes in availability due to the change in survey timing.

Response: The time series was split as requested and separate ***q***s were estimated (with the same selectivity).

Discussion/conclusion: The bad residual pattern in the fit to the Triennial time series was eliminated with the second segment of the series being fitted almost exactly. The Panel recommended, and the STAT agreed, that the split time-series be adopted as the base model because of the concerns about surveying timing (and the poor residual pattern if the series was not split). The Chair expressed concern about the precedent set by adopting the split in the Triennial time series (with regard to assessments of other stocks which rely on it as an abundance index).

T: Explore potential seasonal effects for the Triennial trawl survey (use a simple fixed main-effects GLM).

Reason: A continuing concern about potential changes in availability due to the change in survey timing.

Response: This was not presented for canary rockfish.

Discussion/conclusion: There was a presentation with regard to arrowtooth flounder which indicated that there were some problems with the balance of the sampling – so that it appeared difficult to get meaningful estimates of seasonal effects. It was agreed that the original GLMM approach used to derive the Triennial biomass indices could perhaps be adapted to include seasonal effects. The arrowtooth-STAT agreed to attempt the desired GLMM analysis for canary as well as arrowtooth. (This was done, but was never formally presented – being somewhat peripheral to the assessment given that the split had already been agreed to for the base model.)

U: Profile over $\boldsymbol{R_0}$ for candidate base model (or an alternative model if preferred).

Reason: To explore which likelihood components showed contrast across $\boldsymbol{R_0}$.

Response: Profiles were produced and presented graphically.

Discussion/conclusion: Most contrast was shown by the age data. The indices were fitted better at lower mean recruitment. The length data were fitted better overall by higher mean recruitment (this did not apply to all length times series). Age data were best fitted somewhere in between. Depletion estimates were positively correlated with $\boldsymbol{R_0}$ as were estimates of M for older females.

V: Candidate base model: compare model predictions of mean weight to observed mean weights in the at-sea whiting fleet.

Reason: It was of interest to see if the model was qualitatively consistent with the additional data.

Response: Predicted mean weights were calculated and graphically compared with the observations.

Discussion/conclusion: The predictions were consistent with the observations showing a slight trend with a decline and then an increase in comparison to the very flat observed mean weights.

W: Candidate base model (or an alternative model if preferred): do the "no fishing" run.

Reason: It is of interest to see what biomass would have been present, using the estimated parameters (in particular, recruitment deviations) in the absence of fishing (i.e., no catches removed).

Response: The run was done and the biomass trajectory was compared with that of the candidate base model. The "no fishing" trajectory was relatively flat.

Discussion/conclusion: There was discussion about how exactly the run was implemented in SS2. Two variations are conceivable. SS2 simply applies the recruitment deviations, as estimated, to the stock recruitment curve in the appropriate years. Compared with the base model, which has removals due to fishing, the deviations are applied to higher biomass. An alternative formulation is to use the estimated number of recruits – in some sense attributing recruitment to purely environmental factors. There was also discussion on the use of such runs to construct alternative "dynamic B_0" reference points (i.e., annual depletion being measured by biomass divided by the un-fished biomass in the same year). No conclusions were agreed upon.

X: Continuation of R: explore different values of ***M*** and alternative selectivities (e.g., sex specific).

Reason: See Request R.

Response: Two different values of ***M*** were tried but there was insufficient time to explore further.

Y: Sensitivity to candidate base model: use full set of conditional age-at-length data.

Reason: The STAT expressed some desire to do this run and some members of the Panel were interested to see the results.

Response: This was not done due to lack of time and some diminishment of the STAT's desire to see the results. (It is moot whether the addition of a large amount of extra conditional age-at-length data will lead to a better assessment, but it does hugely extend the required runtime.)

<u>Round 4 requests</u>

Candidate base model:

As in request P with:

- sigmaR = 0.5;
- selectivity for NWFSC survey freely estimated on minimum size bin;
- 1979 selectivity split in Oregon and Washington trawl fisheries;
- split triennial time series.

Z: Run the candidate base model estimating recruitment deviations from 1950. Examine standard deviations of estimated recruitment deviations and total likelihood. Compare the 1950 and 1960 runs. Choose a year to start estimating recruitment deviations. The base model is then fully defined.

Reason: This was a final check to make sure that the choice of 1960 was appropriate.

Response: The 1950 run was done and compared to the 1960 run.

Discussion/conclusion: The 1950 run had below average recruitment estimated from 1950-58 and correspondingly higher recruitment estimated from 1990 onwards. The plot of the standard deviations of the recruitment deviations suggested there was "information" about the early recruitment deviations. This may have been an artifact of the zero sum imposed on recruitment deviations. A better fit to the data was perhaps achieved by making the later recruitments larger, given there were sufficient early recruitment deviations to balance the zero sum. However, the early low recruitment may also have been more consistent with the early fishery length frequency data. The STAT chose the original 1960 start and the Panel agreed.

AA: Profile on steepness, being sure to include the means of the lower and upper 25% tails from the new Dorn prior (i.e., low, high steepness values).

Reason: Steepness appeared to be the best candidate as a dimension of uncertainty.

Response: Results were presented for seven runs extending from the low to the high values of steepness.

Discussion/conclusion: The greatest contrast was shown by the length and age data, both of which fitted better at high steepness. The biomass indices showed very little contrast (this was a change for the Triennial survey – splitting the series had reduced the contrast significantly). There was further discussion on whether these data contained any real information on steepness and the relevance of the results to adopting a higher steepness value in the base model (compared to the 2005 assessment). Certainly, these results lend no support to the low value of steepness used in the 2005 assessment (as the only likelihood components with a "preference" favor high steepness).

AB: Sensitivity to base model: completely remove the influence of the stock recruitment relationship. Compare runs, in particular recruitments and asymptotic confidence intervals on spawning biomass. Plot recruitment versus spawning biomass and overlay alternative stock recruitment relationships (base, low, high).

Reason: It was suggested that it would be useful to see what estimated recruitments fitted the data best in the absence of an imposed stock recruitment relationship.

Response: The requested run was done and the results presented as requested. In addition, the recruitment estimates from the base model were also plotted. The freely estimated recruitments followed the same pattern as those in the base model but became increasing large and variable in the later years (from 1980 onwards).

Discussion/conclusion: There was much discussion about the correct interpretation of these results. This option, as currently implemented in SS2, does not produce interpretable reference points or estimates of virgin biomass (the estimated recruitments, in this case, had a mean level which was far larger than the estimate of $\boldsymbol{R_0}$). However, it was argued that it was still possible to interpret the estimated annual recruitments as those which gave the best fit to the data. The plot of recruitment versus spawning biomass gave no support to a Beverton-Holt stock recruit relationship (there was no indication of lower recruitment at lower stock size). There was agreement that there were other model configurations which should at least be considered in sensitivity runs in the future (and not just for canary). Two options were suggested, neither of which would impose a stock recruit relationship: estimation of an initial age structure at the start of the period where data are informative (with no need for a full catch history); or, retaining a full catch history, estimation of recruitment deviations in an internally consistent manner with average recruitment. In both cases, stock recruitment relationships can be derived from model outputs (e.g., if needed for derivation of reference points).

Description of final base model:

The final base model included all data from the original base model with the addition of the pre-recruit survey indices: catch history 1916-2006; fishery age and length data 1968-2004; NWFSC trawl survey 2003-2006; triennial survey 1980-2004; and coast-wide pre-recruit indices 2001-2006.

The final specification included:

- all priors uniform;
- ***q***s estimated analytically with median unbiased option;
- steepness = 0.511 (mean of middle 50% of new Dorn prior);
- recruitment deviations estimated from 1960;
- sigmaR = 0.5 (no tuning)
- estimate the selectivity in the smallest bin size for the NWFSC survey
- include an extra time block for fishery selectivity (break-point 1979)
- split the triennial survey abundance indices (break-point 1995)
- tune CVs and effective sample sizes.

We note here that one panelist was concerned with the decision to split the triennial survey due to the precedent this would set for other assessments. The majority of the panel felt that there were too many uncertainties regarding the seasonal distribution of canary rockfish to treat the survey as a single time series, and thus implicitly assume that the change in timing of the survey does not affect survey catches. The majority of the panel felt the case for splitting the time series was especially compelling given the residual pattern that occurred when no split was made, which closely matched the shift in survey timing. All the panel members agreed that the change in triennial survey timing needs to be considered seriously in all future assessments that use this survey, but that the decision to split the survey for canary rockfish does not necessarily imply that this is the most appropriate action in all cases.

Comments on the Technical Merits and/or Deficiencies of the Assessment:

Technical Merits:

- The preparation of documentation for the panel was excellent, and greatly facilitated the panel's ability to review the assessment.
- The assessment was based on SS2 software, which has been well tested. Using this software increased the confidence of the panel in the analysis and results.
- The method including age observations as conditional on the length was considered a better way to include age data in the model, thereby avoiding *ad hoc* weighting.
- The data have been improved considerably since the last assessment, especially

improvements in the estimation of age data precision and bias.
- The method for defining time blocks for selectivity based on information independent of the data itself improves on the previous *ad hoc* choices.

Technical Deficiencies
- Given the use of conditional age-at-length data it is not clear with this approach how to calculate effective sample size and jointly tune the model/data.
- An *ad hoc* method was used to weight the commercial age and length data.

Explanation of areas of disagreement regarding STAR Panel recommendations:

Areas of disagreement among the members of the STAR Panel:
There were no areas of disagreement among the five panelists.

Areas of disagreement between the STAR Panel and the STAT team:
There were no areas of disagreement between the STAR panel and the STAT team.

Unresolved Problems and Major Areas of Uncertainty:

Without doubt the value of steepness was the major uncertainty in this assessment. The lack of a recent directed fishery combined with the limited amount of survey data made it impossible to reliably estimate this quantity. Moreover, this same lack of recent information implies that estimated rate of rebuilding is primarily controlled by steepness. The Dorn prior suggests that values of steepness between 0.3 and 0.7 are reasonable. Such a range implies great uncertainty regarding the extent of current rebuilding and even greater uncertainty in forecasts. The Dorn prior is based on estimates of steepness in other rockfish assessments, and is sensitive to the inclusion or exclusion of stocks such as darkblotched rockfish. The estimated values of steepness in these other assessments are themselves highly uncertain, and even if they were precisely determined, the hypothesis that steepness in canary rockfish can be inferred by those in other rockfish stocks may be questionable.

Other issues that deserve further consideration include:
- the triennial survey is inefficient for canary; hence, this assessment is really predicated on catches, low natural mortality, and the assumed value of steepness. The survey data and compositional information may not permit reliable estimation of stock status.
- the possibility of a seasonal effect on $\boldsymbol{q}$ from the triennial survey should be evaluated.
- the relationship between the pre-recruit survey and recruitment deviations needs to be verified.
- stock structure and movement is poorly understood.

Concerns Raised by GMT and GAP Representatives During the Meeting:

The GAP and GMT representatives did not object to the discussion and outcome of the STAR panel review but noted the following points:

- there is a need to undertake a comprehensive analysis of all historical catch data and to assemble the information into a reference data set
- streamlined access to NWFSC data sets would be desirable
- greater examination and utilization of logbook data is encouraged

Recommendations for Future Research and Data Collection:

For the next canary rockfish stock assessment

- Assumptions about stock structure and distributional boundaries should be reviewed in light of information on Canadian/Alaskan catches.
- A catch history should be reconstructed 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. As has been previously recommended, the reconstruction needs to be done comprehensively across all rockfish species to ensure efficiency and consistency.
- Evaluate the feasibility of a bi-lateral assessment with Canadian scientists, perhaps through the TSC (Technical Subcommittee of US Canada groundfish working group).
- Investigate the importance of calendar date and other covariates on catch rates from the triennial survey and propose adjustments to account for seasonal and other variation in selectivity/availability.

Generic issues for groundfish assessments

- Establish a meta database of all data relevant to groundfish stock assessment. The database should include enough detail about the nature and quality of the data that a stock assessment author can make a well informed decision on whether it could be useful for their stock assessment.
- Establish accessible online databases for all data relevant to groundfish stock assessment, so that assessment authors can obtain the raw data if required.
- Establish a database for historical groundfish catch histories, “best” guesses and estimates of uncertainty (and processes for updating and revising the database).
- 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.
- Develop standard and appropriate methods for modeling age and length data, including choice of distribution, initial variance assumptions, and tuning methods (current methods can and should be improved).
- Routinely produce and present supporting documentation for any derived indices which are included in a stock assessment model (e.g., GLMM derived trawl survey abundance indices).

Acknowledgments:

The STAR panel would like to thank the NWFSC, especially Stacey Miller, for coordinating the meeting and the review of the canary rockfish stock assessment. Ian Stewart is also to be highly commended for his positive attitude, efficiency, and his willingness to endure the week-long

scrutiny by the Panel.