BOCACCIO

STAR Panel Report
Southwest Fisheries Science Center
Santa Cruz, California
April 20-25, 2003

STAR Panel Members

Thomas Helser, NMFS Northwest Fisheries Science Center, STAR Chair
Farron Wallace, Washington Department of Fish and Wildlife
Martin Dorn, NMFS Alaska Fisheries Science Center, SSC Representative, Rapporteur
David Sampson, Oregon State University
Patrick Cordue, Center for Independent Experts, University of Miami
Peter Leipzig, Fisherman's Marketing Association, GAP Representative
David Thomas, California Department of Fish and Game, GMT Representative

STAT Team Members Present:

Alec MacCall, NMFS Southwest Fisheries Science Center

Overview

The STAR Panel met to review the 2003 draft bocaccio assessment during the week of April 21-25, 2003, at the Santa Cruz Laboratory of the Southwest Fisheries Science Center. The STAT team (Alec MacCall sole constituent) presented the draft assessment, and assisted the review by conducting analyses and doing additional model runs. The draft prepared for the STAR Panel meeting had many changes from the 2002 bocaccio assessment. These changes included; a lower assumed natural mortality coefficient (reduced from 0.2 to 0.1 per year), recalculated recreational fishery CPUE indices with new data selection criteria, new GLM models for CPUE in numbers rather than weights, and a correction in 2000-2002 to account for the two fish bag limit in 2002 and avoidance of bocaccio in 2000-2001. A new maturity at length relationship was also proposed. Catch in 2000-2002 was adjusted using observer estimates of discard. Length-frequency data sets were adjusted to reflect middle-of-year size as required by the Synthesis model, and the lower two size bins were truncated (<24 cm). Finally, this assessment departed from the catch histories in the 2002 assessment, such that catch histories were used as reported by Ralston et al. (1996) with the addition of the foreign catches of bocaccio from Rogers (2002).

Effects of these changes from the 2002 model on the new model results were not systematically evaluated in the draft document. Accordingly, a major effort of the Panel was to establish an "audit trail" between the model configuration and data in the 2002 assessment and a base run model in the current assessment. Due to the fact that the Stock Synthesis model required approximately 3 hours to converge, time during the meeting was insufficient for the STAT team to step through the four models the Panel deemed necessary to produce such an audit trail. However, after reviewing intermediate model runs with essentially the same data as in the 2002 assessment but with newly modified CPUE indices for the period covered in the 2002 assessment, the Panel was satisfied that the data and methods employed in the updated assessment were relatively consistent. Additionally, the Panel also dedicated major effort to evaluating the new data and different methods (Delta-GLM) used to standardized the fishery-dependent CPUE indices used to track stock abundance.

The bocaccio assessment is complex, reflecting the complexity of the fishery and the diversity of data sources. Notwithstanding the quantity of data used in the assessment, the Panel considered the assessment results to be highly uncertain. While it is clear there has been a significant decline in bocaccio abundance in the past few decades, the estimates of unfished stock abundance as well as the estimates of recent stock depletion are problematic. The lack of reliable fishery-independent indices of stock abundance is a severe limitation for the bocaccio assessment, but this is true for virtually every other West Coast rockfish assessment. While the goal of the stock assessment is to arrive at a risk neutral estimate of stock abundance and trend, extensive use of indices whose properties are not well understood makes it impossible to come to any objective conclusion whether this goal has been attained. Despite these limitations, the Panel concluded that the bocaccio assessment represents the best available scientific information on stock abundance trends and current stock depletion. However, because of

insufficient data prior to 1969 to resolve year-class magnitudes, the Panel recommends against using in rebuilding analyses the Stock Synthesis estimates of steepness or recruitment strength prior to 1970.

Fisheries-independent time series evaluated in the assessment include recruitment indices (midwater trawl, power plant impingement, and recreational pier CPUE) and stock biomass indices (CALCOFI larval index of spawning stock biomass, and NMFS triennial survey). Fisheries-dependent time series include CPUE indices from RecFIN for northern and southern California, CDF&G partyboat CPUE, and commercial trawl CPUE from California logbooks. Length composition data are available for commercial and recreational fisheries and the NMFS triennial survey. The strengths and weaknesses of these data sets are discussed in later sections of the report.

Analyses Requested by the STAR Panel:

1) Compare length composition by year for the Northern California RECFIN and CDFG length composition data.

These data sets were combined in the initial model runs. The Panel was concerned that there might be systematic differences between these data sets. There was apparent contamination of RecFIN length data with CDF&G data. This needs to be investigated further, but is likely to have minimal impact on the assessment results.

2) Compare model runs for unadjusted length composition and length composition adjusted to the middle of the year.

This analysis was not performed due to insufficient time during the STAR Panel meeting.

3) Requests for new analyses of methods used to develop selection criteria of bocaccio trips in the RecFIN data base and subsequent standardizing RecFIN CPUE data.

3a) Calculate species coefficients for pre- and post-1990 periods to evaluate the temporal consistency of the pattern of species association.

3b) Calculate CPUE indices using the GLM model for the Southern California RECFIN data using the two sets of coefficients (pre- and post-1990).

The new method relied on the use of species composition information to identify trips that are more highly associated with catches of bocaccio. In the statistical model, a linear combination of species-specific coefficients (effects) predicted the probability of observing bocaccio in the trip's catch, and a "threshold" probability level was chosen as a means to include trips in the subsequent Delta-GLM analysis. Panel members expressed concerns that the RecFIN index generated by applying the new statistical methodology for trip selection and subsequent Delta-GLM standardization may be sensitive to changes in species composition. The Panel noted that depending on the selection criterion used

(the estimated probability that the trip would catch bocaccio) CPUE at the very beginning and end of the time series varied, although the trends over the entire time series were largely the same.

4) Report sample sizes for the bag limit correction analysis.

The frequency distribution of bag sizes showed a peak at 10 fish (the previous bag limit) but no attempt was made to adjust for possible effects of this earlier bag limit.

5) Calculate the bag adjustment using the ratio of bags from two to nine to one bags rather than bags two or larger to remove the effect of previous 10 fish bag limit.

A bag limit of two fish for bocaccio was imposed in 2002. A regression of mean bag size against the ratio of bags two or larger to one fish bags was used to obtain a correction for CPUE data in 2000-2002. Although the two fish bag limit was not imposed until 2002, the regression analysis suggested avoidance of bocaccio in 2000 and 2001 (in fact a bag adjustment between 2000-2002 was applied to the RecFIN CPUE data).

The bag limit correction has a large impact on the CPUE index values for 2000-2002 (increasing the raw CPUE index by almost a factor of two). The Panel agreed that if CPUE were accepted as an unbiased abundance index, then some adjustment was required to account for changes in fishing behavior due to the change in regulations. Still, any attempt to correct CPUE indices for avoidance of bocaccio is problematic.

6) Provide the estimated trend for a GLM analysis that uses only presence/absence information.

This index would not be affected by the bag limit instituted in 2002. The prevalence of positive catches showed less of an increase in 2000-2002 relative to the CPUE index produced by the full GLM analysis.

7) The Panel requested further information about the derivation of the new maturity schedule.

A new maturity schedule, proposed in the new draft assessment, showed a decrease in the age at 50% maturity from 5 years of age based on the Wyllie Echeverria (1987) results used in the 2002 assessment to 3 years based on new port sampled data. The STAR Panel, recognizing the potential impact on the new assessment results, requested more information on the data and protocols used to determine the new maturity schedule. The STAR Panel was informed by the STAT team that the previous maturity schedule from Wyllie Echeverria (1987) was consistent with the port sampler maturity data after recalculation of maturity at length using appropriate staging codes.

8) It was requested that figures of recreational CPUE indices include both bag limit adjusted and unadjusted indices.

This was not done at the meeting, but the Panel expects the request to be completed in the final draft stock assessment report.

9) Conduct exploratory GLMs using CalCOFI data with a) line-month interaction, b) line-year interaction. Provide the interaction coefficients, and residual deviance.

The Panel was concerned that if there were significant line-month or line-year interactions, then methods used to calculate year coefficients to track abundance could be biased. However preliminary analysis did not suggest significant interaction terms.

10) Document a stepwise transition between last year's assessment (data and model assumptions) and this year's base model runs. Four runs were requested.

The Panel was greatly concerned that there was no audit trail to document the effect of changes from the 2002 assessment specifications and data on model results. This was important since the draft assessment presented at the meeting contained numerous model and data changes. The STAT team showed results of runs that suggested that there was little effect from moving to the recalculated indices over the time period covered by the 2002 assessment. Changes in stock status can therefore be attributed to new indices not included in the 2002 assessment (particularly the adjustment for bag limits in recent years), not to changes in model structure or assumptions.

11) Two base run models were suggested.

The triennial trawl survey biomass index appears to be contradictory to the RecFIN CPUE indices in recent years. While both indices show a strong decline since the mid-1980s, the RecFIN CPUE index has increased substantially over the last three years but the triennial survey index remains flat. On one hand, fishery dependent CPUE indices can mask real declines in abundance if fishers are able to redirect effort to areas of high density. On the other hand, the triennial trawl survey may be less efficient at low stock abundance because bocaccio preferentially occupy untrawlable habitat (varying q with stock abundance). Because of these concerns, the Panel sought to bracket the uncertainty with two base run models. Both models have the following features in common:

Catch history as in Ralston et al. (1996) with the addition of foreign catches of bocaccio reported in Rogers (2002), and 2000 t annually of historical catch prior to 1950.

Updated length composition data and abundance indices with revised methods for calculation, but remove the recruitment indices. The Panel was concerned the recruitment indices were not sufficiently comprehensive spatially to provide a reliable index of year-class strength.

Maturity schedule as in the 2002 assessment, based on Wyllie Echeverria (1987).

Set the natural mortality coefficient equal to 0.15 (with sensitivity analyses for M = 0.1-0.2). The Panel did not consider the rationale presented in the draft assessment for M=0.1 to be convincing, and opted instead for the value in Ralston et al. (1996) as a more

reasonable choice. There wasn't sufficient time during the STAR Panel meeting to do a thorough sensitivity analysis on natural mortality.

Estimate selectivity patterns as in the 2002 assessment.

Estimate recruitment freely from 1960-2001, but use constant background recruitment in 1950-1959.

Base model 1: RecFIN CPUE indices with bag limit adjustment, but remove the triennial trawl survey (biomass index and length composition data).

Base model 2: Use triennial trawl survey data, but remove the RecFIN CPUE indices. Remove the triennial trawl survey length frequency data from before 1989 because of irregular survey coverage for earlier years.

12) Do an additional pair of model runs to evaluate the influence of estimated recruitments during the early time period (1951-69).

These models were similar to base models 1 and 2 but extended background recruitment to 1969 so that annual recruitment strengths would only be based on years for which length composition data were available. The only major effect of extending the backgound recruitment was on the level of stock depletion estimated for the early years of the modeled period.

13) Do an additional model run with the S/R steepness parameter fixed at 1.0, rather than 0.39.

The Panel was concerned that essentially fixing the steepness parameter at a low value was causing Stock Synthesis to estimate a ridiculously high level for unfished stock biomass. The model run with steepness = 1 still resulted in high stock depletion in the first year (4% of unfished in 1951). Later it was discovered that when "background" recruitment is used for the initial recruitment the Stock Synthesis model ignores the reported unfished level (R. Methot pers. commun. April 28, 2003).

14) Prepare bubble plots of length composition residuals and overlays of predicted and observed length frequency.

A pattern of large positive residuals was observed for the main modes in the series of length frequency data, suggesting that the CVs in the growth model were too large, at least for the younger fish. Variable growth by cohort could be responsible to producing such large CVs. This could be affecting the estimates of recruitment strength and variability.

15) Produce a table showing average recruitment, B_0, Spawning biomass per recruit at F=0, stock depletion in 1951 and 2003 for the initial base runs and with modifications for extended background recruitment.

The STAT team produced a table with the following results:

Model	$R_{1951\text{-}1986}$	SPR(F=0)	B_0	B_{2003}/B_0	B_{1951}/B_0
B1-base	5364	2.500	13412	8.5%	26%
B1-constant recr. to 1969	4729	2.787	13180	10.0%	59%
B2-base	5269	2.648	13952	3.9%	22%
B2-constant recr. to 1969	5230	2.498	13063	5.6 %	54%

Estimates of average recruitment were similar for all models runs, and estimates of unfished spawning output (B0) obtained by multiplying SPR (F=0) by average recruitment during 1951-1986 were more plausible than the estimates of B0 from the Stock Synthesis model. (Note: it was determined after the meeting that when Synthesis is configured to use background recruitment in the initial age composition, as had been done in the bocaccio model, the parameter value for virgin recruitment must be manually set equal to the background recruitment for Stock Synthesis to calculate a value of B0 consistent with the background level of recruitment.). Base runs showed higher initial depletion than the new runs, implying that the base model required a higher initial depletion in order to match the large increase in the CalCOFI time series in the late 1960s.

To bracket uncertainty, the Panel decided to use only two runs for ease of interpretation. Three factors were considered: depletion in the first year of the model, depletion in the last year of the model, and early recruitment strength. The third factor was included because of concern that estimating relative year-class magnitude using only the CalCOFI spawning biomass index gives implausible results, namely, a spike in recruitment in 1962 more than twice as large as than any year-class for which adequate length composition data are available. The Panel chose the B1-base and B2 with constant recruitment to 1969 as models that most reasonably incorporated these three major sources of uncertainty. However the true level of uncertainty is unknown, but very likely of larger magnitude.

Technical merits and/or deficiencies of the assessment

The STAR Panel commends the STAT team for their extraordinary effort during the STAR Panel meeting. It was also clear that considerable effort has been devoted to assembling diverse sources of information on trends in bocaccio abundance. The Panel considered the new CPUE analysis based on species associations an innovative approach, but concludes that more work should be done to evaluate whether such a fishery-dependent index is appropriate for tracking stock abundance. The approach may have

wide applicability for numerous west coast rockfishes and the Panel suggests further investigation and possible submission to a peer-reviewed journal. Species associations in the catch could be corroborated by applying the method to data sets with additional information (location, time of day, number of anglers) and the robustness of the approach could be tested using simulation.

Long run times (3 hrs) for the bocaccio Stock Synthesis model hampered a more thorough investigation of model(s) during the STAR Panel meeting. Slow convergence could be indicative of a poorly parameterized model, such as freely estimated parameters without adequate information, or sets of parameters that are highly correlated. It is difficult when using Stock Synthesis to develop a well-supported population model when the data begin very sparsely and then become abundant in later years. Model complexity should depend on the availability of information, but Stock Synthesis was unable to handle this transition smoothly. Consideration should be given to moving to a more flexible modeling environment.

The Panel had many questions relating to how the data and results changed since last year. Basic exploratory data analyses, such as comparison of survey indices, CPUE indices, catch time series, selectivity patterns, year class abundance, biomass trends, etc. would have greatly assisted the work of the STAR Panel. There was also a general lack of description of data sources, sample size information, etc., which should be a required component of any stock assessment document. A pre-assessment meeting could have addressed many of the issues that the Panel was forced to deal with.

Generalized linear models (GLMs) with mixture error distributions (binomial-lognormal, or binomial-gamma) were extensively used in the bocaccio assessment to estimate annual biomass indices. The STAR Panel wishes to emphasize that GLMs are MODELS. Accordingly, diagnostics should be employed as a matter of course to evaluate the adequacy of model fits, error assumptions, and model structure. The STAT team typically fit both delta-lognormal and delta-gamma models and selected the model with the lowest CV for the year effects. The STAR Panel is uncomfortable with this approach and recommends that more formal goodness of fit tests be used for model selection. Systematic and thorough exploration of main effects and interactions is another essential part of fitting GLMs that was not reported in the draft assessment.

The STAR Panel is concerned about the unstated and unevaluated assumptions needed to consider CalCOFI larval counts as an index of spawning biomass. Ralston et al. (2003) notes that changes in the timing of spawning can have a large influence on a biomass index based on larval abundance. A GLM model consisting of only main effects could either overestimate or underestimate stock decline if the stock contracts into more favorable habitat at low stock size (MacCall 1990). The CalCOFI survey is a highly unbalanced design. There is an extensive time series of surveys in southern California, but the area north of Point Conception was surveyed primarily in the early years of the time series, and only infrequently in later years. A more troubling concern is the potential for bocaccio egg production to be influenced by adverse environmental conditions as has been shown for other rockfish during El Nino episodes (Eldridge and

Jarvis 1994). One potential explanation for the low values of the CalCOFI index in the 1990s is that bocaccio egg production was adversely affected by the anomalously warm water temperatures during this decade. Because of these concerns the STAR Panel recommends that the CalCOFI index be used with caution and only to indicate general trends in abundance. Unfortunately, the CalCOFI index is the only information prior to the 1970s, and so the estimated population trends in this early period follow the CalCOFI index very closely. The STAR Panel was unconvinced that the CalCOFI index is so closely associated with bocaccio spawning stock biomass.

The Panel was concerned about the initial equilibrium assumptions in the model. Before 1951, the model assumes that a constant annual "historical catch" has caused the population to reach an equilibrium age structure in 1951. This creates two problems. First, there may not have been sufficient time, since the beginning of the fisheries, for the population to have reached "equilibrium". Second, the appropriate level of "historical catch" is unknown.

Model estimates of historic depletion are very sensitive to the assumed level of historical catch. The Panel recognized the early catch history was highly uncertain due to lack of information (little or no species composition samples). In the black rockfish assessment, the Panel worked with the STAT team to develop a plausible catch history, which avoided the need to assume historical catch and equilibrium conditions in the first year of the assessment. Also, when age or length data were unavailable, recruitment was based on a stock-recruit curve with fixed steepness and a profile on the steepness parameter was conducted to obtain the most plausible value. If this approach had been used in the bocaccio assessment the logical contradictions in the current assessment would have been avoided. However, for the bocaccio assessment there was insufficient time to do this during the meeting.

STAR Panel did not have an opportunity to review a bocaccio rebuilding analysis. We offer the following comments on the use of assessment results in rebuilding analyses. Sampling from recruitments (or R/S) should be confined to years with reliable recruitment estimates, i.e., after 1969 when length composition data are available. Of particular concern is the estimate of recruitment in 1962 that is more than twice as large as any other year class. This recruitment estimate is based solely on an apparent increase in the CalCOFI index (which indexes stock spawning output). Also, we recommend that rebuilding analyses not be based on the Stock Synthesis estimate of steepness. In the Stock Synthesis runs there was almost nil emphasis on the stock-recruit likelihood components, hence the estimated value for the steepness parameter is not reliable. Steepness for bocaccio must be higher for the stock to have been at the levels seen during the 1950s to 1970s.

Base model 1 vs. Base model 2 with constant recruitment to 1969.

Comparison of the likelihood components from both models do not indicate that one model provides a significantly better fit to the data than the other model. The only basis

for evaluating uncertainty is by comparison of point estimates between alternative model runs. This method of bracketing uncertainty doesn't capture the true level of uncertainty

Explanation of Areas of Disagreement Regarding STAR Panel Recommendations

Since the assessment was not completed at the end of the meeting, the Panel is unsure whether there will be significant areas of disagreement concerning the final version of the assessment. Based on our discussions with the STAT team on the last day, it appeared that the STAT team did not agree with the two final models proposed by the STAR Panel to bracket uncertainty. The STAT team placed greater weight on the RecFIN CPUE indices and gave more credence to the CalCOFI index than the Panel considered justifiable. It appeared that the STAT team intended to complete a Stock Synthesis run which included both the RecFIN and the triennial survey indices, an approach which the Panel rejected because the two sources of information were contradictory. Further, it appeared there would be disagreement concerning the time period over which the STAT team intended to resample recruits (or R/SSB) for use in the rebuilding analysis (i.e., recruitment prior to 1969).

Unresolved Problems and Major Uncertainties

Triennial survey selectivity is implausible. The selectivity curve of the triennial survey appeared nearly uniform over all sizes, which appeared very unlikely for a research bottom trawl survey.

A rebuilding analysis was not brought forward, and was not reviewed. The Panel provided in this report what it feels are appropriate recommendations for the parameters and historic recruitments to be used. Specifically, the Panel recommends B1-base and B2 models with constant recruitment to 1969 as alternative model scenarios (equal probability) with recruitment resampled only back to 1970. The Panel re-emphasizes its recommendation against using the Stock Synthesis estimates of steepness or recruitment strength prior to 1970 in rebuilding analyses.

Biomass and recruitment prior to 1970 are highly uncertain since the only available time series is the CalCOFI index, which may not be reliable, and in any case would be unable to resolve the relative strength of individual year-classes.

The RecFIN CPUE indices and the triennial survey trends are contradictory. Fishery-dependent CPUE indices can mask real declines in abundance if fishers are able to redirect effort to areas of high density. Similarly, the triennial trawl survey may be less efficient at low stock abundance because bocaccio preferentially occupy untrawlable habitat (varying q with stock abundance). Generally, the Panel felt that data sources with conflicting information should not be used together in the assessment.

In general, Stock Synthesis predicted modes within the size composition data for bocaccio reasonably well, but had a tendency to consistently under-fit the magnitude of the modal size and overestimate the dispersion about the mode. The residual pattern

from the fit to the length frequency data is unusual and indicates systematic lack of fit. Its effect on the assessment results is unknown.

Recommendations

Due to the extensive fishery closures and regulations prohibiting retention of catch in excess of the legal limits, fishery CPUE indices in the future will be biased indices of abundance. The Council and NMFS need to consider to how to monitor bocaccio status in the future. The CPFV data set consisting of reef-specific indices of abundance from partyboats is extremely valuable for evaluating of local fishing effects and as an index of overall abundance. Reef-specific CPUE is not as subject to the typical limitations of fishery CPUE data. A program of exempted fishing permits for partyboats with observers to monitor stock status should be considered.

More attention needs to be given to how growth is modeled in the assessment. A model with time varying growth or cohort-specific growth may improve the fit to the length frequency data. Alternative ways to model variation in length with age should also be considered. Also, the Panel recommends that ageing of bocaccio be re-visited. A modest ageing sample could be used to evaluate whether the linear trend in the coefficient of variation (CV) of length with age in Stock Synthesis is a reasonable assumption, as well as confirming the model estimates of growth.

The Stock Synthesis model apparently does not perform well with the diverse data sets used to assess bocaccio. Consideration should be given to moving the bocaccio assessment to a new modeling environment, ideally one with optimization routines using automatic differentiation rather than numerical differentiation as in Stock Synthesis.

Early catch history of bocaccio is a significant source of assessment uncertainty. Focused research on historical catch is needed. A comprehensive approach should be taken where historical catches of all West Coast groundfish species are investigated at the same time. Assessing historical effort in West Coast groundfish fisheries may be more successful as a collaborative undertaking between an expert in historical research and a stock assessment scientist.

Work needs to be done to figure how to the start the model with appropriate initial conditions and with sensible initial depletion which is consistent with the data.

The relationship between the CalCOFI index and climate should be evaluated. Two analyses are suggested. The first is to compare the residual patterns in model fits to an environmental index such as the Scripps Pier water temperatures. Adding an environmental covariate to the CalCOFI index catchability coefficient may improve the model fit to the index if annual egg production is influenced by environment conditions. A second analysis would be to compare biomass trends to indices associated with regime-scale environmental variability to see if significant correlations exist that would help explain long-term abundance trends.

References

Eldridge, M.B., and B.M. Jarvis. 1994. Temporal and spatial variation in fecundity of yellowtail rockfish. Trans. Am. Fish. Soc. 124:16-25.

MacCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle, 153 p.

Ralston, S., J.R. Bence, M.B. Eldridge, W.H. Lenarz. 2003. An approach to estimating rockfish biomass based on larval production, with application to *Sebastes jordani*. Fish. Bull. 101:129-146.

Ralston, S. J.N. Ianelli, R.A. Miller, D.E. Pearson, D. Thomas, and M.E. Wilkins. 1996. Status of Bocaccio in the Conception/Monterey/Eureka INPFC areas in 1996 and recommendations for management in 1999. In: Status of the Pacific Coast Groundfish fishery through 1996 and recommended acceptable biological catches for 1997, stock assessment and fishery evaluation, Appendix B. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, Oregon 97220.

Rogers, J.B. 2003. Species allocation of Sebastes and Sebastolobus sp. caught by foreign countries off Washington, Oregon, and California, U.S.A. in 1965-1976. NOAA Tech.Memo.

Wyllie Echeverria, t. 1987. Thirty-four species of California rockfishes: maturity and seasonality of reproduction. Fish. Bull., U.S. 85:229-250.