Review of the Pacific Fishery Management Council Stock Assessment and Review (STAR) Panel Lingcod and Cabezon Stock Assessments

Dr. Mark N. Maunder

For the University of Miami Independent System for Peer Review

Executive Summary

In general, the cabezon and lingcod assessments appropriately use the available data and adequately represent the uncertainty in the assessments. Modifications and refinements to the assessment methods are possible, but major improvements to the assessments will only be achieved with the collection of additional data. The assessment for cabezon is very uncertain and provides very little information that could be used for management of this stock. Alternative management approaches may be more appropriate for this species until adequate data is available to reduce the uncertainty. The assessment for lingcod is more certain. The lingcod stock has been depleted to low levels and is now rebuilding, but the rate of rebuilding is uncertain.

a) Primary sources of uncertainty
The main uncertainties in the cabezon assessment are the unknown level of historic catch, natural mortality, steepness of the stock-recruitment relationship, the appropriateness of the CPFV CPUE based index of relative abundance, and growth parameters.

The main uncertainties in the lingcod assessment are the values of natural mortality and the steepness of the stock recruitment relationship, the influence of the portion of the population that is in Canadian waters, the lack of fit to the early catch-at-age data in the northern area, the appropriateness of the indices of abundance based on the NMFS triennial survey and trawl logbook CPUE, amount and survival of discards, and the value of the initial exploitation rate or equivalently the historical catch.

b) Strengths and weaknesses of current approaches
The current approaches for both cabezon and lingcod appear to be appropriate. The assessments present a reasonable description of the uncertainty in the assessments. There are some inconsistencies between the two approaches and modifications may provide some improvements to the assessments. The use of Coleraine reduced the flexibility of the lingcod assessment and an alternative method should be considered in the future.

c) Suggested research priorities to improve the stock assessment.
There is no fishery independent survey designed for these species. The current surveys are probably inadequate for these species. The priority for future assessments is to develop a dedicated fishery independent survey for these and other inshore species. Various methods should be considered including the use of trap and hook and line gears. Alternatively, or in addition, a coast wide tagging study should be considered. Other

priorities for research include improving estimates of catch and discards, estimating biological parameters, and collection of catch-at-age data.

Additional information can be found in the STAR Panel reports. In general, this report focuses on topics not covered in the STAR Panel reports and topics that require additional detail.

Background

This section provides a short description of the cabezon and lingcod stocks and assessments. A more complete background can be found in the individual stock assessment and STAR Panel reports.

Cabezon

This is the first quantitative assessment of the population status of cabezon along the west coast of the United States. It is also the first assessment of a species from a group of inshore groundfish that have relatively little data from the fisheries and about their biology. The population was divided into a northern area (Washington and Oregon) and a southern area (California) and these were assessed separately. There is much less data for the northern area and the results for this area were implausible. Therefore, the STAT Team focused on the assessment of the southern area. Due to the lack of survey data and because there has been no previous assessment, the status of cabezon stock along the west coast of the United States has, until now, been unknown.

Cabezon is primarily a nearshore species. The management of cabezon is mainly limited to recent imposed regulations or regulations that are combined with those for other species. Until recently, the majority of removals have been made by the recreational sector. Cabezon have been a component of the recreational sector for over a century, but the amount of removals is very uncertain before 1980. The commercial fishery started catching cabezon in significant amounts in the late 1980s as part of the live-fish hook and line and pot/trap fisheries.

The data available for this species are very limited, and there is little information about the biology of the species. The catch data are very uncertain prior to 1980. There is no fishery independent survey for this species because most surveys are conducted at depths outside the range of cabezon. Possible indices of abundance were based on recreational CPUE, larval abundance surveys, and power-plant impingement rates of juvenile cabezon. However, many of these indices are based on data that have only a 5-10% probability of a non-zero observation. There is no catch-at-age data available, only catch-at-length. Biological information is lacking, and even the growth rate is based on very limited data.

A custom stock assessment model was developed using the AD Model Builder software. The model was an age and sex-structured statistical catch-at-length analysis. Recruitment was modeled as varying around a Beverton-Holt stock-recruitment relationship. The model included separate fisheries for the recreational and commercial sectors.

Lingcod
In 1999 lingcod was designated overfished based on an assessment of the northern portion of the stock. A rebuilding analysis was carried and a rebuilding plan developed. A coast wide assessment was carried out in 2000 confirming the overfished status of lingcod. The rebuilding analysis and rebuilding plan were subsequently updated.

Lingcod have had a relatively long period of management, but it has only been in recent years (since 1995, but more severe since 1998) that management actions have had a substantial effect on the commercial removals.

Lingcod have a long history of both recreational and commercial fisheries dating back to the early 1900s. However, assessment models have only used catch data since 1973. Recreational catch comprises a large proportion of the total catch and is an important component of this stock. The amount of discarding has increased substantially since 1998 due to management regulations.

A lot more is known about the biology of lingcod than cabezon, but there is still lack of information about natural mortality and the steepness of the stock-recruitment relationship. These parameters are influential in the stock assessment model and therefore translate into uncertainty in the stock assessment results.

The population was divided into northern (US-Vancouver, Columbia) and southern (Eureka, Monterey, Conception) areas and assessed separately. There is much less data for the southern area and subsequently the results for the southern stock are more uncertain.

The Coleraine general stock assessment model was used to assess this stock. The model was an age and sex-structured statistical catch-at-age and catch-at-length analysis. Recruitment was modeled as varying around a Beverton-Holt stock-recruitment relationship. The model included separate fisheries for the recreational and commercial sectors. The population was modeled starting in 1973 at an exploited population size. The model is fit to the NMFS triennial survey and trawl logbook CPUE based indices of abundance for both areas and to the WDFW tagging based index of abundance for the northern area. Catch-at-age and length data is available for the commercial and recreational fisheries, the NMFS triennial survey, and the tagging study.

Description of Review Activities
I spent the period of 15th-19th of September 2003 in Seattle at the Northwest Fisheries Science Center (NWFSC). Most of this time was spent attending the Stock Assessment Review (STAR) Panel meeting for cabezon and lingcod. Several background documents were provided, and they are listed in appendix A. The AD Model Builder (ADMB) code and data files were also supplied for both assessments (including the Coleraine user interface files). These documents and stock assessment models were reviewed before the STAR Panel meeting. Several analyses were carried out to investigate the stock assessments for these two species.

Review of Coleraine model for lingcod

The program and data files for the Coleraine application for lingcod were supplied by Thomas Jagielo of the Washington Department of Fish and Wildlife. As the lead programmer for the Coleraine application, I was able to thoroughly review the application of Coleraine to the lingcod stock. In general, the use of Coleraine for lingcod was carried out correctly. However, there were a few changes required. These changes were sent to the STAT team before the meeting and were incorporated in the results presented by the STAT team at the meeting.

A bug (actually, an undocumented known limitation) in the release version of Coleraine causes the parameter phase and bounds for all survey parameters (catchability and selectivity) to be the same as those for the first survey. Due to this bug, the catchability and the full selectivity difference by sex for the tagging data were estimated on the bounds. This problem was overcome by increasing the bounds for the first survey. The problem has been corrected in the most recent unreleased version of Coleraine (personal communication, Billy Ernst, University of Washington). It should be noted that in any application using AD Model Builder (the underlying programming language used to develop Coleraine), the parameter estimates should always be compared to their bounds. Parameter estimates on their bounds indicate that the model is mis-specified and this can cause convergence and estimation problems.

The STAT team was unable to estimate the right and left limbs of selectivity curve of the surveys due to convergence problems. It was found that estimating these parameters in an additional phase allowed their estimation in the northern area.

The STAT team provided likelihood profiles of several model parameters. These profiles were not smooth indicating convergence problems. Suggested changes in the phase order of the estimated parameters and including the –maxfn 10000 option to increase the number of function evaluations improved the convergence ability of the model.

The aging error matrix was entered the wrong way around. This caused the estimated proportions-at-age to sum to a number other than 1. Transposing the aging error matrix corrected this problem.

Development of a short term Coleraine model for cabezon

The Coleraine stock assessment program was applied to the cabezon data in an effort to get a more realistic description of the total uncertainty. To include the uncertainty about the historic catch that was assumed in the STAT teams assessment model, the Coleraine model was started in 1979 at an estimated exploited abundance level (similar to the lingcod assessment). The CPFV observer data index was the only index of abundance used. The commercial and recreational length-frequency data was used. Uncertainty in natural mortality and the steepness of the stock-recruitment relationship was allowed by including priors for these parameters and estimating them in the model. Using a short term model also implicitly allowed for uncertainty in historic recruitment, which was ignored for the early years of the STAT team assessment. The results of this model showed that the confidence interval on the current depletion level was about 0.2 to 0.7 (Figure 1).

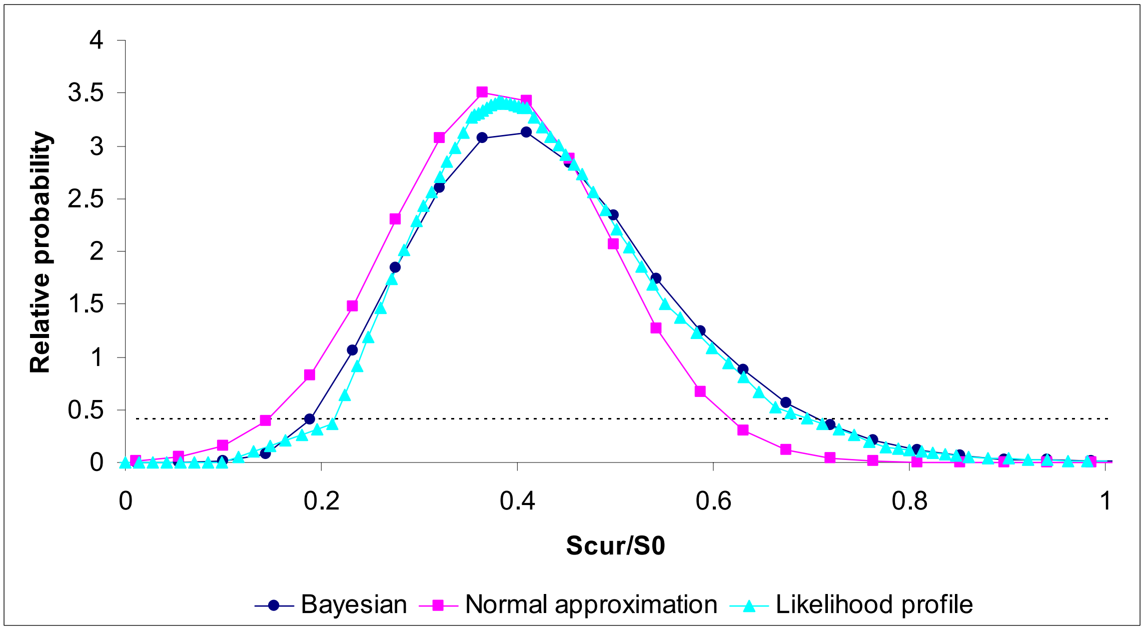

Figure 1. Bayesian posterior distribution, normal approximation, and likelihood profile for the current depletion level of cabezon in the southern area using the Coleraine implementation of a short term model that started in 1979. The horizontal dashed line intercepts the normal approximation distribution at the 95% confidence intervals.

Background documents

Several background documents were provided before the review (see Appendix A for additional details). These documents were read prior to the meeting in Seattle and a list of clarification questions was prepared for the STAT team.

<u>Documents supplied by the STAT Teams</u>

1) "Status and future prospects for the cabezon (*Scorpaenichthys marmoratus*) as assessed in 2003", by Kevin Piner, Jason Cope, Carolina V. Minte-Vera, and Andre Punt. This document is an initial draft of the cabezon assessment. It describes the fishery and historical management, the stock assessment model, the data used in the assessment, biological parameters, and initial results.
2) "Assessment of lingcod (*Ophiodon elongatus*) for the Pacific Fishery Management Council in 2003", by Thomas H. Jagielo, Farron R. Wallace, and Yuk Wing Cheng. This document is an initial draft of the lingcod assessment. It describes the fishery and historical management, the stock assessment model, the data used in the assessment, biological parameters, initial results, and rebuilding

analysis. The document's appendices included the Coleraine manual, the Coleraine output for the northern and southern assessments, and the coast wide lingcod rebuilding analysis.

Other documents

Several other relevant documents related to the assessments were read before the meeting. Those found to be of main interest are listed below.

1) "Movement, mortality, and size selectivity of sport- and trawl-caught lingcod off Washington" by Thomas H. Jagielo. This published paper described the estimation of movement, fishing mortality, and size selectivity based on analysis of tag-recapture data. The results showed that lingcod generally move offshore as they age. The tag-recapture data was shown to be uninformative for natural mortality and growth parameters.
2) "Discard mortality of trawl-caught lingcod in relation to tow duration and time on deck" by Steven J. Parker, Polly S. Rankin, and Robert W. Hannah. This published paper describes an experiment to determine the discard mortality of lingcod. The results showed that lingcod were generally tolerant of being discarded and substantial mortality only occurred with extended time on the deck (>30 minutes) or when crushed because the total catch of all species was high. They stressed that information about how fish are handled on the vessel is needed to determine discard mortality. This study, though informative, suffered from pseudo replication because the replication occurred at the level of the individual fish and not at the level of the trawl tow.
3) "Lingcod (*Ophiodon elongates*)" by Alan J. Cass, Richard J. Beamish, and Gordon A. McFarlane. This published report summarized the biology, stocks, and fisheries of lingcod in Canadian waters. This report provided a valuable source of information about the species.
4) "West coast groundfish harvest policy" by Stephen Ralston. This document gave background on the management of groundfish by the PFMC.

Documents supplied at the meeting

In addition to the background documents provided before the meeting, the following document was provided at the meeting.

1) "Modelling the abundance indexes of lingcod (*Ophiodon elongates*) in Pacific Northwest coast of US", by Yuk Wing Cheng. This document summarized the initial application of General Additive Models to the commercial lingcod logbook data.

STAR Panel meeting

The Stock Assessment Review (STAR) Panel for cabezon and lingcod was held from the 15th-19th of September, 2003 in Seattle at the Northwest Fisheries Science Center. The members of the STAR Panel and STAT Teams are listed in Appendix B. This meeting started with presentations of the two draft stock assessment reports. These presentations included descriptions of the fishery and historical management, the stock assessment model, the data used in the assessment, biological parameters, and initial results. The stock assessments were then discussed, and the STAR Panel created a list of additional requested analyses for each species. An iterative process followed that alternated between the presentation of results and the request for additional analyses. Finally, the STAR Panel generated a report for each species summarizing the assessment, listing the requested analyses, and providing recommendations. These reports include input from the CIE reviewer and should be consulted for additional information. (note: the GAM analysis of commercial logbook data for lingcod was the first item discussed due to its controversial nature and availability of the analyst. There were also several e-mail exchanges between the STAR Panel members before the meeting regarding this analysis)

Additional Analyses

The CIE reviewer carried out several additional analyses before and during the STAR Panel meeting.

Numerous sensitivity analyses were carried out using the Colerain assessment model for lingcod. This model had the corrections listed above, but did not include the new GLM based commercial logbook CPUE abundance index. These sensitivities looked at dome shaped versus asymptotic selectivity, values of natural mortality, and the steepness of the stock recruitment relationship. All sensitivities showed that the population was highly depleted in the late 1990s, but rebuilding since then was sensitive to the values of natural mortality and the steepness of the stock recruitment relationship. These results were consistent with the sensitivity analyses carried out by the STAT Team.

A simple Pella-Tomlinson production model (with $B_{MSY}/B_0 = 0.3$; see Maunder 2003 for justification of this value) was fit to the index of abundance based on the GLM standardization of the lingcod commercial logbook data for the northern area. This analysis showed that the rebuilding of the stock abundance since its low point in the late 1990s was much faster than the assessments using Coleraine. This analysis provides additional support to the uncertainty about the rebuilding of the lingcod stock.

A retrospective analysis using the STAT Teams Coleraine model was carried out using the data that was available for the 2000 lingcod assessment in the northern area. This analysis used data up to 1999. The results were essentially the same as the assessment using all the data. Unfortunately, the model run did not converge completely and the results should be taken as possibly unreliable. The reason the model did not converge is because the exploitation rate became too high and triggered the maximum exploitation rate penalty that occurs at an exploitation rate of 1. Due to the implementation of this penalty in Coleraine, triggering the maximum exploitation rate penalty does not necessarily mean that the exploitation on any age-class is equal to 1. This only happens when all fisheries have full selectivity at a single age. Triggering the maximum exploitation rate may have also occurred in some of the other model runs that had convergence problems (e.g. estimating the selectivity of the surveys for the southern area).

CIE Report
The period following the STAR Panel meeting was used to synthesize the information and results provided at the meeting to generate this report.

Summary of findings

The final assessments developed during the STAR Panel meeting are considered to be an appropriate use of the data and description of the uncertainty for the cabezon and lingcod stocks. However, there are still several uncertainties that need to be investigated and/or reduced. In particular, there are several differences between the application of stock assessment methods to these two species that highlight areas that need to be investigated in the future. The following gives a summary of the assessments of the two stocks and the areas of concern that need to be addressed, as well as other topics of interest.

Status of the cabezon stock and its uncertainties
It was only possible to produce a credible assessment for the southern area. This is because the abundance indices for the northern area showed an increasing trend and are therefore inconsistent with the assumptions of the stock assessment model (The analysis estimates very large biomass so that the catch has very little effect on the abundance creating a flat biomass trajectory. A flat biomass trajectory is the best fit to the index of relative abundance since the model starts from a virgin state, and biomass is unable to increase due to the trajectory of increasing catch).

The main source of data for the southern area is the Commercial Passenger Fishing Vessel (CPFV) logbook data. However, the index of abundance from this data is inconsistent with the stock assessment model. When the coefficients of variation estimated from the CPUE standardization are used in the stock assessment, less than half of the 95% confidence intervals for each year cover the predicted values. In addition, there is substantial autocorrelation in the residuals. When a scaling parameter is estimated within the stock assessment model to increase the coefficients of variation for the CPFV logbook index, predicted values are contained within the confidence intervals, but the autocorrelation in residuals increases. There is much more uncertainty when the scaling parameter is estimated and the population is estimated to be less depleted (the confidence interval on depletion is about 0.2 to 0.4 without the scaling parameter and 0.3 to 0.9 with). One surprising result was that the model with the scaling parameter estimated was generally not sensitive to the sensitivity analyses. It is uncertain what the reason for this is, but this result provided additional support for estimating the scaling parameter. The STAR Panel agreed with the STAT Team that the scaling parameter should be estimated to ensure that the assessment estimated an appropriate level of uncertainty.

I applied a short term model to the cabezon data using Coleraine to investigate the full uncertainty due to unknown historical catch, steepness of the stock recruitment relationship, natural mortality, and variation in historic recruitment. To include the uncertainty about the historic catch that was assumed in the STAT teams assessment model, the Coleraine model was started in 1979 at an estimated exploited abundance level (similar to the lingcod assessment). The CPFV observer data index was the only index of abundance used. The commercial and recreational length-frequency data was used. Uncertainty in natural mortality and the steepness of the stock-recruitment relationship was allowed by including priors for these parameters and estimating them in the model. Using a short term model also implicitly allowed for uncertainty in historic recruitment, which was ignored for the early years in the STAT team's assessment. The results of this model showed that the confidence interval on the current depletion level was about 0.2 to 0.7 (Figure 1). It is expected that the final Bayesian analysis of the STAT Team that combines posteriors of several sensitivities to the steepness of the stock recruitment relationship and natural mortality will produce similar results.

The main uncertainties in this assessment are the unknown level of historic catch, natural mortality, steepness of the stock-recruitment relationship, the appropriateness of the CPFV CPUE based index of relative abundance, and growth parameters.

See the cabezon STAR Panel report for additional information.

Status of the lingcod stock and its uncertainties

The results of the stock assessment and the numerous sensitivity analyses support the conclusion that the lingcod stock has been depleted to a low level and is now rebuilding. However, there is uncertainty in the exact level of depletion and the rate of rebuilding. In particular, the current depletion level and rebuilding rate is sensitive to the natural mortality and the steepness of the stock recruitment relationship. Neither of these parameters could be estimated by the stock assessment model. However, a steepness of 0.9 was associated with a better fit to the data for the northern stock compare to the lower values that were considered. One of the main causes of the uncertainty is the lack of information about recruitment for the most recent years (since 1993 for the northern stock). The lack of information is due to no apparent modes in the age or length-frequency data and because the tagging data, which is considered an index of recruitment for the northern area, stops in 1992.

The main uncertainties in this assessment are the values of natural mortality and the steepness of the stock recruitment relationship, the influence of the portion of the population that is in Canadian waters, the lack of fit to the early catch-at-age data in the northern area, the appropriateness of the indices of abundance based on the NMFS triennial survey and trawl logbook CPUE, amount and survival of discards, and the value of the initial exploitation rate or equivalently the historical catch.

See the lingcod STAR Panel report for additional information.

Short term versus long term models

The assessments for the two species used different methods to cope with uncertainty in the historic catches. The cabezon assessment assumed recreational catch from 1930 to 1979 and applied a sensitivity analysis to this level while the lingcod assessment started in 1973 at an exploited population size. The choice between these two different methods can have implications on the estimates of depletion and the associated uncertainty (Maunder 1998; see figure 2 of Maunder 1998). First, estimates of depletion from a long-term model will often have much less uncertainty than from a short term model. Therefore, sensitivity analysis of the historic catch or inclusion of the uncertainty of the historic catch in the analysis (see Appendix C) should be carried out. Second, the estimates of current depletion will be based on different assumptions and data. Current depletion requires the estimation of both current biomass and virgin biomass. In the short term model, the estimate of biomass at the start of the modeling period is dependent on the catch-at-age or length data and the estimate of virgin biomass is dependent on the assumptions about the stock-recruitment relationship. The estimate of depletion level from the long term model can be dependent on the uncertain historic catch. The depletion level is used in management to determine if a rebuilding strategy is required and what exploitation rate should be applied. Therefore, the best estimates of depletion level and the associated uncertainty are required. It is recommended that research be carried out to determine the most appropriate method to investigate uncertainty in historic catch.

In reviewing the documents after the STAR Panel meeting, it was unclear why the lingcod assessment used a short-term model starting in 1973. Previous reports indicated that data is available prior to 1973, but it has not been used in the assessments. I assumed that this data is highly uncertain, and that is why the lingcod STAT Team chose to ignore the data. However, this should be made explicit in the stock assessment report. The lingcod STAT Team fixed the initial exploitation rate in the short-term model and looked at sensitivity to the initial exploitation rate. This is similar in application to the cabezon STAT Team who assumed historical catch and looked at sensitivity to the historical catch. The advantage of the short-term model occurs because the initial exploitation rate is a parameter in the model that can be estimated and/or used to explicitly include the uncertainty in the Bayesian analysis in addition to sensitivity analyses.

Sensitivity analysis versus Bayesian analysis

Assessments for both stocks use (or propose to use) Bayesian analysis to represent the uncertainty in the analyses. However, the application of Bayesian analysis is unable to represent all the uncertainty in the analysis. This is because the implementation for lingcod and cabezon does not facilitate the inclusion of all the uncertainty. This difficulty can be divided into three components: 1) uncertainty in catch data; 2) uncertainty in model structure; and 3) inability to estimate some parameters in the model. One option is to use sensitivity analysis to investigate the additional uncertainty. However, the combination of uncertainty from sensitivity analyses is problematic. The cabezon STAT Team intends to combine posterior distributions of the different sensitivity analyses to combine the uncertainty. This method is somewhat ad hoc, and another method may be more appropriate.

1) There are several methods that can be used to include uncertainty in the catch data. Three of these methods are described in appendix C.

2) There have been several studies that have included structural uncertainty in Bayesian analysis (Sainsbury 1988; Patterson 1999; Parma 2001). The method of combining posteriors weighted by the prior probability for each model appears appropriate for inclusion of uncertainty in model structure (e.g. include uncertainty in the structure of the stock recruitment relationship, Ricker versus Beverton-Holt).

3) ADMB uses the mode of the posterior distribution and the variance-covariance matrix to develop jumping rules for the MCMC algorithm. When the model is unable to converge or if a parameter is estimated on the bound, the MCMC algorithm will not work or will work improperly. One method to overcome this problem is to put a penalty on the parameter that is causing the problems during the phase that estimates the mode of the posterior and then remove the penalty during the MCMC phase [if (!initial_params::mc_phase) f += square(parameter-prior);]. It should be noted that combining posteriors of sensitivity analyses on multiple parameters (e.g. natural mortality and the steepness of the stock recruitment relationship) ignores the correlation between these parameters and it is preferable to include the parameters in the Bayesian analysis.

GAM analysis of lingcod commercial logbook data
Indices of abundance are often the most important and influential data included in stock assessments. Therefore, it is important that these indices are as accurate and precise as possible. This generally means that the index is proportional to abundance. However, many reviews of stock assessments focus on the stock assessment model and the weighting factors of the different data sets, but only superficially review the abundance indices. It was encouraging that this review included an in-depth review of the lingcod commercial logbook CPUE analysis. Unfortunately, the other indices were not reviewed in detail. In particular, the CPFV logbook index for cabezon is problematic.

The year effect was included in the GAM analysis as a continuous variable. Using year as a continuous variable smoothes the year effect. The population model can also be considered a smoothing function. Therefore, if the GAM analysis is used to develop an index of abundance, the year information in the CPUE data is smoothed twice. If it is thought desirable to smooth the full data set, the CPUE standardization can be integrated with the stock assessment model (Maunder 2001a). However, this produces a very computationally intense procedure and may not be desirable. It is recommended that year be included as a categorical (factor) variable.

The early data, before 1980, do not have latitude and longitude information. Therefore, the GAM did not use these data in the analysis. However, the GAM was used to predict the year effect for 1975-1979. This required extrapolation outside the range of the data using the GAMs cubic splines. The population dynamics model within the stock assessment should be a better extrapolator (otherwise we would not be using a stock assessment model at all; the stock assessment model also uses other data to provide information on abundance), and therefore the years 1975-1979 should not be included in the GAM. Alternatively, latitude and longitude could be excluded from the GAM analysis and the data for 1975-1979 included. This would produce four more years for the index of abundance. There is a tradeoff between including latitude and longitude, which should reduce bias in the index, and obtaining four more years for the relative abundance index. The decision on which method to use should consider how much of the total variability in CPUE that the explanatory variables latitude and longitude explain, and the differences in the year effects for years after 1979, when these two variables are excluded. The STAT team and STAR Panel agreed to drop latitude and longitude from the analysis and to include the additional years (However, one and two of these years were removed from the analysis due to limited sample size in the northern and southern areas, respectively; see below).

The GAM method is new and was not understood by the majority of the meeting participants. Therefore, it is suggested that the analysis be carried out using a GLM with year included as a categorical variable. Each categorical variable requires a base (comparison) category to be identified. The base category does not change the relative index, only the confidence intervals. This is because the estimated year effects are a measure of the differences from the base year. Therefore, if the base year has large uncertainty, confidence intervals for all years (the relative difference from the base year) are large. In stock assessment models, only the relative variation is needed because stock assessments usually estimate a scaling parameter for the variance. There are several methods that can be used to overcome the choice of base year (Maunder and Starr 2003). Two simple methods are: 1) chose the base year as the year with the lowest variance, which is typically a function of the number of observations for the year and the size of the effect for the year. In this case, when including the index in the stock assessment, the base year is given a variance equal to the lowest estimated variance; and 2) weigh the variance by some function (e.g. square root) of the number of observations for each year. A scaling parameter for the variance on relative abundance indices derived from GAM or GLM analyses of CPUE data should always be estimated in the stock assessment model or fixed at the appropriate level different from that estimated in the GAM or GLM.

Trip limits in 1998 greatly reduced the catch rates. The catch limits in 1995-1997 were much less restrictive and probably did not influence the catch rates significantly. Therefore, year effects for 1998 and later should not be used in the stock assessment model. To be consistent and eliminate the influence of data from these years, they should also be excluded from the GLM or GAM analysis.

The STAT Team and STAR Panel agreed to use a GLM with year as a categorical variable, excluding latitude and longitude as explanatory variables, use data from 1978 and 1977 in the south and north, respectively (the first few years had very low sample size and the variance was included in the assessment as being constant over time), up to 1997.

Custom models versus packages

The assessments for the two species use different models. The cabezon assessment uses a custom built model using the ADMB programming language. The lingcod assessment uses a general stock assessment model, Coleraine, developed by researchers at the

University of Washington (Hilborn et al. 2000). Both models were developed using ADMB and have similar model structure. However, the cabezon model is designed to more appropriately represent the dynamics and data of the cabezon stock. The previous lingcod assessment model was a custom built model developed in ADMB. However, the previous STAR Panel raised concerns about the custom build model being invalidated and the lack of ability to fit to length-frequency data. For these reasons, the lingcod STAT Team decided to use a general package that has been used in previous assessments and allows for the use of length-frequency data.

Unfortunately, Coleraine is unable to provide all the flexibility needed for the assessment of lingcod. For example, the right limb of the selectivity curve has to be the same for both sexes. In the previous custom built model, the right limb of the selectivity curve could differ between the sexes. In addition, the selectivity curves for the two sexes both have a maximum of one. However, some of the data suggest that the overall selectivity is less for one sex.

It is not guaranteed that a general software package can eliminate all errors in the assessment. For example, the initial Coleraine model for lingcod had a few errors in its implementation (the aging error matrix was transposed and the bounds set to narrow for some parameters). These errors were due to the lack of familiarity with Coleraine and the inadequate documentation for some of Coleraine's "features". The STAT Team has experience with these types of models that reduced the number of errors in implementing the Coleraine model. A recent review of another assessment developed by a less experienced stock assessment modeler discovered a much greater number of errors in the implementation of Coleraine. There may also be errors in the general model that have yet to be discovered because not all options have been validated.

The custom model is more flexible, but more prone to errors. The cabezon STAT Team applied the Stock Synthesis general model to the data to validate the results of the custom model. This approach is commendable, however it does not necessarily validate the results as not all components and uses of the model are compared (e.g. rebuilding).

I consider that the best method of developing a model is to first write down all the appropriate equations, review them, and have general agreement on their appropriateness. Next, double code the model using two independent programmers (It is difficult to get two independent programmers, so I often program the model in Excel and ADMB. These are two very different programming styles and the probability of making the same errors is reduced). If a general model is to be used, it needs to be very general or in a form (source code available) that allows the analyst to modify the code. However, these two approaches still have the possibility of errors due to either not all the options being validated or the code for the modifications having errors.

The STAT Teams should consider investigating the MULTIFAN-CL software that is now freely available (http://www.multifan-cl.org/). This software is continually being developed and has been used for several stock assessments (Fournier et al. 1998; Hampton and Fournier 2001; Kleiber et al. 2003). One feature of MULTIFAN-CL that is lacking in most assessment models is the inclusion of tagging data and spatial structure in the population dynamics (Hampton and Fournier 2001).

Combining two management areas and ignoring Canada

The cabezon and lingcod assessments are both carried out independently (with some sharing of information for cabezon) for the northern and southern areas. This separation is generally made based on management rather than biological considerations (this may be appropriate due to the limited movement of these species; see spatial model recommendation). In both cases, there is a different amount of data in each area and therefore the ability to assess them differs. This makes the combining of the two areas, as required for management, difficult. Under the current rebuilding procedure, it is recommended that combining of areas should be carried out after the rebuilding analyses have been carried out independently for each area. It is possible that the uncertainty in the area with less information will overpower the other area. Doing the rebuilding analyses separately and presenting both the separate and combined results will ensure that this is recognized.

The methods used to manage these stocks should be re-evaluated to incorporate the differences in uncertainties between the two areas. In addition, the stocks are not independent from those in the adjacent Canadian waters, and the Canadian data should also be considered in the assessments. For lingcod, one of the high areas of catch rates is right on the boarder with Canada.

Asymptotic selectivity

The assessments of the two species differed in their assumptions about the selectivity curves. The cabezon selectivity was assumed to be asymptotic, and the lingcod assessment allowed the selectivity to be dome-shaped if there were appropriate support in the data. The lingcod assessment showed that there was correlation between the value of the steepness of the stock recruitment relationship and the right limb of the selectivity curves. A steepness of 0.9 (as used in the base-case) caused the estimates of selectivity to be asymptotic in the northern assessment. It is often assumed that at least one fishery should have asymptotic selectivity to provide a useful stock assessment. However, the sensitivity analyses for both species showed that the results were not sensitive to the assumption of asymptotic and dome-shaped selectivity. Methods to choose between domed shaped and asymptotic selectivity need to be developed and may involve ad hoc methods based on the biology, stock distribution, or gear characteristics.

Recruitment

The biology of cabezon and lingcod differs from many other commercial species because the males guard the eggs and the female produces fewer eggs (Cass et al. 1990). These differences mean that the relationship between stock and recruitment and the variability in recruitment may be different compared to that of other commercial species. It might be expected that there is higher survival of eggs that are guarded, and thus recruitment variability would be less, and recruitment may be more correlated with stock size. This is contrary to the high stock recruitment steepness (0.9) used in the lingcod base-case, but it may be more consistent to the steepness used for cabezon (0.7). However, there is still a larval stage after the eggs hatch, and significant density independent mortality may occur then. Due to the high mortality of unguarded eggs and because males guard a single female's eggs (Cass et al. 1990), the stock recruitment curve should be related to both the male and female biomass. This method has been used for the stock assessment of eastern Bering Sea snow crab (Benjamin Turnock, personal communication).

Discards

There are several management controls that have been implemented in recent years that have promoted the discarding of these species. Discards are now a very large portion of the total catch for these two species. If the mortality due to discarding is not taken into consideration, the ability of the stocks to rebuild will be over-estimated. A recent study for lingcod (Parker et al. 2003) showed that lingcod have very high survival rates if discarded soon after capture. This study also showed that survival was reduced due to crushing if lingcod were caught in trawl sets with large catches (including all species) or if they were left on the deck for long periods (e.g. 30 minutes). Therefore, it is important that records are kept of the size of the catch and the amount of time discarded fish are kept on the deck before being returned to the ocean. In the case of recreationally caught fish, it is expected that there would be high survival unless there is significant levels of other mortality (e.g. gut hooking) or if recreational fishers discard small dead fish and replace them with larger fish when their bag limit has been exceeded.

Inconsistencies in recreation catch estimates

The recreational catch estimates for 1980 were much higher than in previous years for both lingcod and cabezon. It is my understanding that the method used to estimate recreational catch changed in 1980 for some of the areas. This change in catch indicates that there is large uncertainty in the historical catch as emphasized by both STAT Teams and, in some cases, has a large influence on the results. It also indicates that current estimates of recreational catch may also be highly uncertain. The recreational catch comprises a large proportion of the total catch. Therefore, it is important that the quality of the estimates of recreational catch is investigated and improved, and that the uncertainty in the estimates of recreational catch are included in the analysis (see Appendix C).

The standard deviation of length-at-age

It is becoming common to include catch-at-length data directly into stock assessment models (e.g. Fournier et al. 1998). Historically, catch-at-length data was converted into catch-at-age data using several different approaches. However, the development of new modeling approaches and efficient parameter estimation techniques has allowed the integration of catch-at-length data into the stock assessment model (Fournier et al. 1998). The integrated method is considered more appropriate because the uncertainty in the method to convert length to age is included in the analysis and the proportions-at-age are constrained by the population dynamics model. The most common method is to use an age-structured model to predict catch-at-length data. This method requires a distribution of length-at-age. The distribution is usually assumed to be normal with known mean and standard deviation. The mean is taken from the growth model and the standard deviation is often assumed to be a linear function of the mean. The cabezon STAT Team found that the results were, in some cases, sensitive to the assumed standard deviation of the length-at-age distribution. In some applications, it has been possible to estimate the parameters of the linear relationship between the standard deviation and mean (Hampton and Fournier 2001). However, this is not always the case and unrealistic estimates have been obtained (e.g. Watters and Maunder 2002). It has become apparent that methods need to be developed to include information about the distribution of length-at-age into stock assessment models. Maunder (2002) integrated age-length data into the stock assessment of yellowfin tuna to provide information of length-at-age. However, the age-length data were not randomly collected and thus provide a biased representation of the distribution of length-at-age. Collection of length-at-age data need to be designed so that the information about the distribution of length at age is maintained (i.e. random and not systematic sampling of length categories).

Separation of CPUE into proportion positive and catch-rates of positive observations
The delta lognormal method for standardizing CPUE has become the preferred method of standardizing CPUE when a large portion of the observations have zero catch. This method combines two separate models. The first model determines the probability of obtaining a non-zero catch and the second model determines the size of the catch if the catch is non-zero. In the lingcod assessment, the year effect of the positive catches generally remained constant, but the proportion of positive catches reduced over time. This meant that the index of relative abundance was driven by the proportion of positive catches. An ecological explanation for this phenomenon is that as the abundance of lingcod decreases, the group size stays the same; it is the number of groups that decreases. An alternative explanation is that the fishermen reduced the proportion of trips that they targeted lingcod. If the size of lingcod groups does not reduce as the abundance decreases, it is possible that efficient searching methods may keep the catch-rates high even as the abundance declines. A more detailed investigation into the difference between the year effects of the proportion non-zero catch and the catch rates of the non-zero observations is needed.

Water hauls in the NMFS triennial trawl survey index
The lingcod STAT team provided estimates of relative abundance from the NMFS triennial trawl survey including and excluding water hauls. There was a large difference in the abundance index for 1980 when the water hauls were excluded due to one large catch that was classified as a water tow. In addition to changing the index, the variance of the estimate was also greatly reduced when the water tow was excluded. This suggests that the survey is not appropriate for lingcod. The survey may not be measuring the main

component of the lingcod population and only a low density fringe portion of the population.

Effective sample size for catch-at-age and catch-at-length data

It is well recognized that the actual sample size for catch-at-age and catch-at-length data is higher than the effective sample size that should be used in statistical catch-at-age and catch-at-length models. The effective sample size is lower because 1) the samples are often correlated due to fish in the catch grouping by size or age and 2) the population dynamics model does not include all the processes that cause the age and length distributions to change over time (e.g. temporal variation in growth or natural mortality). The cabezon STAT Team used (a single iteration of) the method of McAllister and Ianelli (1997) to estimate the effective sample size and, based on the cabezon assessment, the STAR Panel recommended that the lingcod STAT Team divide their sample size by ten. The cabezon STAT Team used a single sample size for each gear and did not iterate the method until the effective sample size converged. A more appropriate approach would be to use a linear regression of effective sample size versus actual sample size and iterate the process several times (Alec McCall, personal communication). However, I am not aware of any studies that have determined if this method improves the parameter estimates and the estimates of uncertainty.

Management of low information species

The PFMC bases its management on the results of stock assessments and rebuilding analyses. Unfortunately, many inshore groundfish species that require management have little data (both abundance and biological) leading to very uncertain assessments. The cabezon assessment, the first assessment of this species, is a prime example. The confidence interval on the depletion level is 0.2 to 0.7 and provides very little advice on the status of the stock. It would be useful for the PFMC to re-evaluate the methods that it uses to manage these species. There is a substantial amount of recent literature on alternative methods to manage low information species (Starr et al. 1997; De Oliveira et al. 1998; Hilborn et al. 2002), and these methods should be considered by the PFMC. This is particularly important given the lack of success using traditional approaches (e.g. $F_{35\%}$ as a proxy for F_{MSY}; Ralston 2002).

STAR Panel process

In general, the STAR Panel meeting was a success, and the final stock assessments were appropriate for each species and adequately described the uncertainty. However, it was only on the last morning of the meeting that the base-case for the lingcod assessment was determined. This late determination of the base-case resulted from the sensitivity analyses having been carried out using different configurations of the model, making it difficult for the STAR Panel to refer to the sensitivities with respect to the base case. This highlighted the possibility that one week may not provide sufficient time for the assessments to be appropriately reviewed. Two possible solutions to avoid this undesirable outcome are to: 1) ensure that the stock assessment reports and associated model program files are distributed at least two weeks in advance, including any supplementary reports (e.g. the standardization of CPUE); or 2) get one of the STAR Panel members to work with the STAT Team the week before the STAR Panel meeting. This would ensure that most errors are removed from the analysis, problem areas are highlighted, and that the appropriate sensitivity analyses are carried out before the meeting. An example of how this would have benefited the meeting is demonstrated in the review of the GAM analysis of lingcod commercial logbook data. At the start of the meeting, the STAR Panel considered the GAM analysis inappropriate, and it was replaced with a GLM analysis. However, due to the change in methodology, neither the STAT Team nor the STAR Panel had time to provide a thorough review of the new GLM analysis.

Many of the sensitivity analyses carried out by the lingcod STAT Team were only presented on the computer projection screen. This made it difficult for the STAR Panel to discuss the results when the STAT Team was away producing more model runs. In future STAR Panel meetings, it is recommended that either printouts of all results be provided or that the presentations be kept on a computer at the meeting for later viewing.

It was encouraging to see the involvement of industry both on the STAR Panel and at the meeting. Industry members, particularly those with a historical knowledge of the fishery, are able to provide information to help interpret the data and stock assessment results. The scientists that are analyzing the data often do not have this type of knowledge.

Conclusions/recommendations

In general, the cabezon and lingcod assessments appropriately use the available data and represent the uncertainty. Modifications and refinements to the assessment methods are possible, but major improvements to the assessments will only be achieved with the collection of additional data. The STAR Panel reports list numerous recommendations for both species and these reports should be referred to for the majority of recommendations. Additional recommendations can be found above in the findings section. The priority for future assessments is to develop a dedicated fishery independent survey for these and other inshore species. Various methods should be considered including the use of trap and hook and line gears. Alternatively, or in addition, a coast wide tagging study should be considered. Other priorities for research include improving estimates of catch and discards, estimating biological parameters, and collection of catch-at-age data. Here, I focus on those recommendations that I consider require additional details.

The highest priority for both species is a fishery independent index of abundance. The current NMFS triennial survey is probably inadequate for both species. A new coast wide survey that focuses on inshore species should be developed. Various methods should be considered including the use of trap and hook and line gears. Alternatively, or in addition, a coast wide tagging study should be considered. A tagging study may, in addition to abundance, provide information on movement, growth, and natural mortality. These species are particularly applicable to tagging studies because of their tolerance of being captured (Parker et al. 2003). The tagging study could be combined with the survey, and all or a portion of the survey caught fish being tagged and released, providing additional information at less additional cost. Due to the tolerance of being captured, it may be possible that multiple recaptures for the same individual can be achieved providing more information than the standard tag-recapture study. Recording of multiple recaptures would require the participation of recreational and/or commercial fishermen who would need to keep logbooks of all fish caught, including tagged individuals that are released. Any survey or tagging program should be carried out over multiple years and coordinate with Canada. Tag release and recapture data for Canada would allow the determination of the dependence of the US portion of the stock on the Canadian portion of the stock. Surveys and tagging studies need to be carried out in the areas that are closed to fishing and where it is illegal to retain these species. This is important because when an area is closed, the fishery dependent data coming from those areas ceases, and any changes in those areas may not be represented in the stock assessments.

A high proportion of the total catch for these species comes from the recreational sector. Therefore, it is important that accurate accounting of the recreational catch occurs. There is substantial uncertainty in the historic estimates of recreational catch. There is also likely to be substantial uncertainty in the current estimates. Accurate estimates of total catch are vital to assessment and management of these stocks. Therefore, it is recommended that efforts be made to improve the estimates of historical catch and to improve the reporting or estimation of future catch. This is important for both the recreational and commercial fisheries.

Recent management actions have greatly increased the amount of discarded fish. Knowledge of the amount of discards and the survival rate is vital for determining the rebuilding rates of these stocks. A recent study has identified some of the factors affecting survival rates of discards (Parker et al. 2003). This work should be continued. In addition, it is recommended that a better record of the amount and context (e.g. total catch of all species in the trawl shot, time left on deck) of discards be collected for these stocks.

The main index of abundance for the cabezon assessment was the CPFV logbook CPUE index. The model diagnostics suggested that this index was not appropriate due to high levels of autocorrelation in the residuals. It is my understanding that the CPFV logbook CPUE data has only been used in one other assessment (cowcod). Most other assessments use the CPFV observer data. There are many uncertainties in the CPFV logbook CPUE and a more detailed examination of this data is required. It was suggested that a workshop be held on calculating abundance indices from recreational data. I support this suggestion and add that it also incorporate the estimation of recreational catch. These two subjects are highly related, and both are of high importance for the inshore species.

The available data for lingcod suggest that they have limited movement, but show some offshore movement with age and a spawning migration (Jagielo 1999). Distributions of catch show that there are hotspots that are isolated from each other. This information indicates that smaller scale assessments and management may be more appropriate. However, the sub-areas may be linked due to the dispersal of larvae. It is recommended that a spatial model be considered for this stock. If the model is not used for stock assessment and management, the spatial model will still help in the understanding of this species. The model may also help understand the dependency of the US portion of the stock on the Canadian portion of the stock. The tagging study suggested above would provide useful information for a spatial model (e.g. Maunder 2001b, Hampton and Fournier 2001).

References

Cass, A.J., Beamish, R.J. and McFarlane, G.A. (1990) Lingcod (*Ophiodon elongates*). *Canadian Special Publication of Fisheries and Aquatic Sciences* **109**.

De Oliveira, J.A.A., Butterworth, D.S., and Johnston, S.J. (1998) Progress and problems in the application of management procedures to South Africa's major fisheries. In fishery stock assessment models, edited by F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.J. Schweigert, P.J. Sullivan, and C.I. Zhang, Alaska Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks, pp. 513-530.

Fournier, D.A., Hampton, J. and Sibert, J.R. (1998) MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Science* **55**, 2105-2116.

Hampton, J. and Fournier, D.A. (2001) A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine and Freswater Research* **52**, 937-963.

Hilborn, R., Maunder, M.N., Parma, A., Ernst, B., Paynes, J. and Starr, P.J. (2000) Documentation for a general age-structured Bayesian stock assessment model: code named Coleraine. Fisheries Research Institute, University of Washington, FRI/UW 00/01.

Hilborn, R.H., Parma, A., and Maunder, M.N. (2002) Exploitation rate reference points for west coast rockfish: are they robust and are there better alternatives? North *American Journal of Fisheries Management*, **22**: 365-375.

Jagielo, T.H. (2003) Movement, mortality, and size selectivity of sport- and trawl-caught lingcod off Washington. *Transactions of the American Fisheries Society*, **128**: 31-48.

Kleiber, P., Hinton,M.G., and Uozumi, Y. (2003) Stock assessment of blue marlin (*Makaira nigricans*) in the Pacific using MULTIFAN-CL. *Marine and Freshwater Research* **54**(4): 349 – 360.

Maunder, M.N. (1998) Problems with using an environmental based recruitment index: examples from a New Zealand snapper (Pagrus auratus) assessment. In fishery stock assessment models, edited by F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.J. Schweigert, P.J. Sullivan, and C.I. Zhang, Alaska Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks, pp. 679-692.

Maunder M.N. (2001a) A general framework for integrating the standardization of catch-per-unit-of-effort into stock assessment models. *Canadian Journal of Fisheries and Aquatic Science* **58**, 795-803.

Maunder, M.N. (2001b) Integrated Tagging and Catch-at-Age ANalysis (ITCAAN). In: *Spatial Processes and Management of Fish Populations* (Proceedings of the Symposium on Spatial Processes and Management of Marine Populations, October 27-30, 1999, Ancorage, Alaska) G.H. Kruse, N. Bez, A. Booth, M.W. Dorn, S. Hills, R.N. Lipcius, D. Pelletier, C. Roy, S.J. Smith, and D. Witherell, eds. Alaska Sea Grant College Program Report No. AK-SG-01-02, University of Alaska Fairbanks, pp. 123-146.

Maunder, M.N. (2002). Status of yellowfin tuna in the eastern Pacific Ocean. *Inter-Amer. Trop. Tuna Comm., Stock Assessment Report*, **3**: 47-134.

Maunder, M.N. and Starr, P.J. (2001) Bayesian Assessment of the SNA1 snapper (*Pagrus auratus*) stock on the northeast coast of New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **35**: 87-110.

Maunder, M.N. and Starr, P.J. (2003). Fitting fisheries models to standardised CPUE abundance indices. *Fisheries Research* **63**: 43-50.

Maunder, M.N. and Watters, G.M. (2003). A-SCALA: An Age-Structured Statistical Catch-at-Length Analysis for Assessing Tuna Stocks in the Eastern Pacific Ocean. *Inter-Amer. Trop. Tuna Comm., Bull*. **22**(4) xx-xx.

McAllister, M. K., and Ianelli, J. N. (1997) Bayesian stock assessment using catch-age data and the Sampling/ Importance Resampling Algorithm. *Canadian Journal of Fisheries and Aquatic Sciences* **54**, 284-300.

Parker, S.J., Rankin, P.S., and Hannah, R.W. (2003) Discard mortality of trawl-caught lingcod in relation to tow duration and time on deck. *North American Journal of Fisheries Management* **23**: 530-542.

Parma, A.M. (2001) Bayesian approaches to the analysis of uncertainty in the stock assessment of pacific halibut. *American Fisheries Society Symposium* **24**, 111-132.

Patterson, K.R. (1999) Evaluating uncertainty in harvest control law catches using Bayesian Markov chain Monte Carlo virtual population analysis with adaptive rejection sampling and including structural uncertainty. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 208-221.

Ralston, S. (2002) West coast groundfish harvest policy. *North American Journal of Fisheries Management* **22**: 249-250.

Sainsbury, K.J. (1988) The ecological basis of multispecies fisheries, and management of a demersal fishery in tropical Australia. In: *Fish Population Dynamics* J.A. Gulland, ed., 2nd edition. Wiley, New York. pp. 349-382.

Starr, P.J., Breen, P.A., Hilborn, R., and Kendrick, T.H. (1997) Evaluation of a management decision rule for a New Zealand rock lobster substock. *Marine and Freshwater Research* **48**: 1093-1102.

Watters,G.M. and Maunder, M.N. (2002). Status of bigeye tuna in the eastern Pacific Ocean. *Inter-Amer. Trop. Tuna Comm., Stock Assessment Report*, **2**: 147-246.

Appendix A: Bibliography of materials provided

"Status and future prospects for the cabezon (*Scorpaenichthys marmoratus*) as assessed in 2003", by Kevin Piner, Jason Cope, Carolina V. Minte-Vera, and Andre Punt.

"Assessment of lingcod (*Ophiodon elongatus*) for the Pacific Fishery Management Council in 2003", by Thomas H. Jagielo, Farron R. Wallace, and Yuk Wing Cheng.

"Modelling the abundance indexes of lingcod (*Ophiodon elongates*) in Pacific Northwest coast of US", by Yuk Wing Cheng.

1999 STAR Panel/STAT Team Reports

1999 C.I.E. Review Reports

2000 STAR Panel Reports

2000 C.I.E. Review Reports

"Southern lingcod stock assessment in 1999", by Peter B. Adams, Erik H. Williams, Kelly R. Silberberg, and Thomas E. Laidig.

"Assessment of lingcod (*Ophiodon elongates*) for the Pacific Fishery Management Council in 2000", by Thomas Jagielo, Deborah Wilson-Vandenberg, John Sneva, Sandra Rosenfield, and Farron Wallace.

Groundfish stock assessment and review process during 2003

Code and data files for the cabezon assessment

Code and data files for the lingcod assessment

"Influence of improved performance monitoring on the consistence of a bottom trawl survey" by Mark Zimmermann, Mark E. Wilkins, Kenneth L. Weinberg, Robert R. Lauth, and Franklin R. Shaw. ICES Journal of Marine Science, 60: 818-826.

Appendix B: Members of the STAR Panel and STAT teams

STAR Panel
Han-Lin Lai, NOAA/Northwest Fisheries Science Center, PFMC SSC (Chair)
Chris Legault, NOAA/Northeast Fisheries Science Center
Mark Maunder, Inter-American Tropical Tuna Commission (CIE Reviewer)
David Smith, Primary Industries Research Victoria, Australia
Tony Smith, CSIRO Marine Research, Australia

PFMC
Tom Barns, California Department of Fish and Game, PFMC GMT
Tom Ghio, PFMC GAP

Cabezon STAT Team
Kevin Piner, NOAA/Northwest Fisheries Science Center
Jason Cope, University of Washington, SAFS
Carolina Minte-Vera, University of Washington, SAFS
Andre Punt, University of Washington, SAFS

Lingcod STAT Team
Thomas Jagielo, Washington Department of Fish and Wildlife
Farron Wallace, Washington Department of Fish and Wildlife
(Henry) Yuk Wing Cheng, Washington Department of Fish and Wildlife

Appendix C: Methods for including uncertainty about historical catch

There are at least four methods that could be used to include uncertainty about the historic catch.

1) Ignore the unknown historical catch by starting the model at an exploited level in the first year catch is known more accurately. This involves estimating an initial age-structure that is different from the unexploited equilibrium assumed in the cabezon assessment. There are several methods that can be used to estimate the initial age-structure and the method used will depend on the data available: a) Estimate the abundance in each age-class as a free parameter. b) Estimate an initial fishing mortality parameter that is multiplied by the selectivity-at-age and added to the natural mortality to determine the total mortality (Maunder and Starr 2001). The initial age-structure is assumed in equilibrium using this total mortality (as used in the lingcod assessment). This method can be extended as in the Coleraine model (Hilborn et al. 2000) by estimating a plus group scaling parameter and a penalized deviate for each age-class (similar to recruitment deviates). If a stock-recruitment relationship is assumed, it may be desirable to estimate a different average recruitment (different from virgin) to compensate for the exploited population size. c) A method similar to (b) can be used, but rather than estimating the initial exploitation rate, set it equal to the exploitation rate in the following few years (Hampton and Fournier 2001; Maunder and Watters 2003). This method only works if using the Fournier et al. (1998) method, which estimates catch from the effort fishing mortality relationship (see below). If a stock-recruitment relationship is used, an analytical formula can be used to calculate the initial recruitment (Maunder and Watters 2003).

2) The method of Fournier et al. (1998) fits to the catch data conditioned on effort rather than assuming it is known perfectly. This method estimates an effort deviate for each time period and uses the relationship between effort and fishing mortality, and the Baranov catch equation to predict the catch.

$$F_{t,a} = s_a q E_t \exp\left(\varepsilon_t\right)$$

$$C_t = \sum_a \frac{F_{t,a}}{F_{t,a} + M_a} N_{t,a} \left[1 - \exp\left(-F_{t,a} - M_a\right)\right] w_a$$

Where C_t is the total catch in weight for time t, F_t is the time component of fishing mortality in time t, q is the catchability coefficient, E_t is the effort in time t, ε_t is the annual effort deviate, s_a is the selectivity for age a, $N_{t,a,}$ is the numbers in age-class a at time t, M_a is the natural mortality rate for age a, and w_a is the weight of an individual of age a.

A likelihood function comparing the predicted catch to the observed catch and a penalty on the effort deviations are added to the overall objective function.

$$-\ln L\left(C^{obs} \mid \theta\right) = \sum_t \frac{\left(C_t^{obs} - C_t\right)^2}{2\sigma_C^2}$$

$$-\ln P(\varepsilon) = \sum_t \frac{(\varepsilon_t)^2}{2\sigma_E^2}$$

The size of the standard deviation of the effort deviation (in addition to the size of the standard deviation of the catch, which is usually set to a small value so that the catch is predicted almost exactly) determines the amount of information in the CPUE data about abundance. If no effort data is available, constant effort can be used and the standard deviation of the effort deviation set to a large number. If a standardized CPUE index is available, this can be incorporated by calculating an effort time series by E = C / CPUE. The standard deviation used in the catch likelihood will determine the amount of uncertainty in catch and can be specified on an annual basis to include the inter-annual variation in the catch uncertainty.

3) The uncertainty in catch can be included in the Bayesian analysis by sampling the catch from the appropriate distribution each time the objective function is evaluated in the Bayesian analysis.

4) Include a parameter in the model to scale all the catch that is uncertain and estimate this parameter in the Bayesian analysis. This method assumes that all the error is completely correlated among years.

Appendix D: Statement of work

STATEMENT OF WORK

Consulting Agreement Between the University of Miami and Dr. Mark Maunder

October 17, 2003

General

The consultant will participate in the Stock Assessment and Review (STAR) Panel of the Pacific Fishery Management Council (PFMC) from September 15-19, 2003, in Seattle, Washington. This assessment will provide the basis for management of the Lingcod and Cabezon, two groundfish species important to the nearshore and recreational fisheries. This is a first time assessment for Cabezon, for which an independent review shall improve quality and credibility of these assessments.

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

The consultant will be provided with the following documents:

1. Current draft Cabezon and Lingcod stock assessments reports;
2. Most recent previous Lingcod stock assessment;
3. An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Specific

1) Become familiar with the draft Lingcod and Cabezon stock assessments and background materials;

2) Participate in the STAR Panel to be held in Seattle, Washington, from September 15-19, 2003;

3) Understand the primary sources of uncertainty in the assessments;

4) Comment on the strengths and weaknesses of current approaches;

5) Recommend alternative model configurations or formulations as appropriate during the STAR panel;

6) No later than October 3, 2003, submit a written report[1] consisting of the findings, analysis, and conclusions, addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Die, via email to ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani, via email to mshivlani@rsmas.miami.edu.

Signed________________________

Date_________

[1] The written report will undergo an internal CIE review before it is considered final. After completion, the CIE will create a PDF version of the written report that will be submitted to NMFS and the consultant.

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

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

2. The main body of the report should consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The report should also include as separate appendices the bibliography of all materials provided and a copy of the statement of work.