$p(\boldsymbol{\theta} \mid \mathbf{y}) = \begin{Bmatrix} \mathbf{Q} \\ \mathbf{R} \\ \mathbf{A} \end{Bmatrix} \begin{matrix} uantitative \\ esource \\ ssessment \end{matrix}$

LLC

Quantitative Resource Assessment LLC

San Diego, CA
USA.

Bering Sea Pacific Cod Stock Assessment Model Scenarios Requested by FLC/QRA

Introduction

Requesting model scenarios for the Pacific cod stock assessments without knowledge of the assessment authors preferred model and alternatives greatly complicates the process. Therefore, we request that the assessment author use his best judgment when interpreting our requests and contacts us with any questions about the scenarios. We also have provided a commentary of the CIE reviewers' reports that the assessment author, Plan Team, and SSC, can use as a guide in creating their own scenarios or interpreting the scenarios that we request (see Appendix).

Scenarios Requested

Scenario 1

The stock assessment authors preferred model with the following changes

Estimate the bias and variance parameters of the aging error matrix inside the stock assessment model.

Scenario 2

The stock assessment authors preferred model with the following changes

Time blocks determined by initially modeling selectivity as a random walk.

Scenario 3

The stock assessment authors preferred model with the following changes

Use the bootstrap method to estimate the samples sizes, scale them so that the average is 300, then use iterative reweighting to update the samples sizes by fitting the Michaelis–Menton equation to the observed and effective samples sizes.

Scenario 4

The stock assessment authors preferred model with the following changes

Model time invariant survey selectivity, but model temporal changes in growth.

Scenario 5

The stock assessment authors preferred model with the following changes

Include the conditional age at length data and the length composition data, rather than the mean-size-at-age data, and estimate the variation in length at age inside the stock assessment model.

Appendix: Report on the Pacific cod CIE review

Summary

The three reviewers generally agree that the Pacific cod assessment is based on the best available science, but there a few areas that need improving through additional research and data collection. The reviewers did not provide any novel suggestions that would greatly improve the assessment or deal with the remaining issues.

The review process followed a set of questions outlined in the terms of reference. I present my summary below based on these questions. I have also added topics addressed by the reviewers that were not included in the terms of reference. My recommendations are provided at the bottom of each section in italics. I also summarize my recommendations that are relevant to choosing the 2011 model assumptions.

Assumptions for 2011 model

Further investigation is needed to determine the appropriate method to model and estimate the aging error and selectivity parameters.

Include the age composition data and the length composition data (or age conditioned on length and length composition) for all years if an appropriate aging error matrix can be generated, otherwise exclude the age data.

Include the conditional age at length data and the length composition data, rather than the mean-size-at-age data, and estimate the variation in length at age inside the stock assessment model.

Keep the current data partitioning.

Use dynamic binning for composition data

Eliminate the pre 1982 survey data.

Time blocks should be determined by initially modeling selectivity as a random walk.

Fix the catchability at the value estimated by Nichol et al. (2007).

Fix natural mortality at the value from Jensen's (1996) equation.

Use the bootstrap method to estimate the samples sizes, scale them so that the average is 300, then use iterative reweighting to update the samples sizes by fitting the Michaelis–Menton equation to the observed and effective samples sizes.

Model time invariant survey selectivity, but model temporal changes in growth.

Fix the standard deviation of recruitment the annual residuals at 0.6 and test the sensitivity of management parameters to 0.4 and 0.8.

Terms of reference topics

Use of age composition data

The reviewers acknowledge that there is aging error/bias. They recommend including the age composition data in the assessment model in conjunction with an aging error matrix. They noted that excluding the aging data caused some undesirable model behavior. The reviewers also recommended continuing the research into the sources of the aging bias.

There was some concern that the age composition data used the same information as the length composition data so that the data was used twice. This needs to be clarified. However, double weighting of data is not too concerning since the weights are arbitrary in the current model. If the weights are "estimated" inside the model, then the issue of double weighting needs to be addressed.

Include the age composition data and the length composition data (or age conditioned on length and length composition) for all years if an appropriate aging error matrix can be generated (see below), otherwise exclude the age data.

Use of mean-size-at-age data

The reviewers recommend excluding the mean-size-at-age data, particularly if temporal variation in growth is not modeled. The mean-size-at-age data is the same data as used in the age composition and length composition data so the data sets are not independent.

Include the conditional age at length data and the length composition data rather than the mean-size-at-age data. This data provides information on variation of length at age, mean length at age, and temporal variation in mean length at age. The appropriate data to include also needs to consider the information required to estimate an aging error matrix.

Use of ageing bias as an estimated parameter

The reviewers did not agree on whether estimating the aging bias in the assessment model was the best approach. The models run during the review were not adequate to determine if the aging bias could be estimated appropriately. More research is needed on the form of the aging error and bias and whether it can be estimated within the stock assessment model.

The aging error comes from at least two source: 1) variability in reading the ages as indicated by double reading and 2) bias due to "false" rings being formed or the edge effect. An appropriate functional form for the aging error needs to be developed that can accommodate these two sources of error. We need to obtain the model files and investigate the appropriate method to model and estimate the aging error.

External estimation of between-individual variability in size at age

All three reviewers suggest estimating the variation of length at age outside the stock assessment model. This is partly due to undesirable model behavior when it was estimated inside the model.

The model does not include age conditioned on length data and therefore ignores some of the information available to estimate variation in length at age. Estimating variation in length at age outside the model does not take account of the aging error or selectivity. Variation in length at age should be estimated inside the model while including the age conditioned on length data. The development of a more appropriate growth model should also improve the models estimates of variation in length at age.

Catch data partitioned by year, season, and gear

The reviewers consider that the current catch data partitioning is appropriate. One reviewer noted that there is uncertainty in the catch estimates and this should be investigated.

Keep the current catch partitioning. Consider investigating a model that combines all catch into a single fishery for each season (it might be appropriate to reduce or eliminate the number of seasons) and use time varying selectivity for the fishery (the approach used by Ianelli for assessing pollock). The length composition data would need to be raised to the total catch within each fishery and summed across fisheries.

Size composition data partitioned by year, season, gear, and 1-cm size intervals

The reviewers consider that the current size composition data partitioning is appropriate. They recommended using dynamic binning to reduce the number of zeros in the likelihood function.

Keep the current size composition partitioning and use dynamic binning.

Age composition data partitioned by year, season, and gear

The reviewers consider that the current age composition data partitioning is appropriate. The reviewers were ambivalent about the use of the pre 1982 survey data because it probably does not influence the results.

Keep the current age composition partitioning and eliminate the pre 1982 survey data.

Functional form of the length-at-age relationship and estimating the parameters thereof

The reviewers noted the poor performance of the Richards growth curve due to the need to constrain one of the parameters to be positive.

A new growth curve needs to be developed for the Pacific cod assessment and implemented in Stock Synthesis.

Number and functional form of selectivity curves estimated, including assumptions regarding which selectivity curves should be forced to exhibit asymptotic behavior

The reviewers suggested that at least one selectivity curve should be asymptotic. They also suggested that a random walk should be used to model time varying selectivity to identify changes is selectivity and use this to determine where the time blocks should be applied.

The reviewers did not understand the types of selectivity curves available in Stock Synthesis. A selectivity curve can be created as a random walk over age (or length). This would allow a bimodal selectivity curve. The parameter for each age (the age offset) can also be modeled as a random walk over time, as can the parameters for functional forms.

A more robust approach is needed to model selectivity and determine which selectivity curves are asymptotic. Time blocks should be determined by initially modeling selectivity as a random walk.

Fixing the trawl survey catchability coefficient for the recent portion of the time series such that the average product of catchability and selectivity across the 60-81 cm size range equals the point estimate obtained by Nichol et al. (2007)

The reviewers recommended fixing the catchability at the value estimated by Nichol et al. (2007). They noted that when estimated, the estimate was similar to the Nichol et al. (2007) value. They also recommended collecting more tagging data to improve the estimate.

Fix the catchability at the value estimated by Nichol et al. (2007) and encourage further data collection.

Fixing the natural mortality rate at the value corresponding to Jensen's (1996) Equation 7

The reviewers recommended that the value of natural mortality continue to be fixed at the value from Jensen's (1996) equation. They also noted that it should be updated once the aging bias has been addressed and that age-specific natural mortality should be investigated.

Fix natural mortality at the value from Jensen's (1996) equation until the issues in the stock assessment model have been addressed, then estimate natural mortality within the stock assessment model and consider age specific natural mortality.

Input sample sizes for size composition and age composition data, and input log-scale standard deviations for survey abundance data

The reviewers recommended that the standard errors used for the survey index of abundance likelihood function should be reevaluated based on the survey design. The reviewers generally agreed with the bootstrap method used to estimate sample sizes, but noted that rescaling the averages to 300 caused the samples sizes to be lower than that suggested by the model fit to the composition data (effective sample sizes).

Use the bootstrap method to estimate the samples sizes, scale them so that the average is 300, then use iterative reweighting to update the samples sizes by fitting the Michaelis–Menton equation to the observed and effective samples sizes.

Allowing for annual variability in trawl survey selectivity

The reviewers questioned the need for annual variability in survey selectivity. However, they did recognize that catchability might change over time due to environmental factors such as bottom water temperature.

One reason for allowing the trawl survey selectivity to change over time is to accommodate changes in mean size at age for the young individuals. Temporal changes in catchability could also be due to abundance of different types of prey.

Model time invariant survey selectivity, but model temporal changes in growth.

Setting the input standard deviation of log-scale recruitment (σR) equal to the standard deviation of the estimated log-scale recruitment deviations

The reviewers were not conclusive about how to deal with the standard deviation of recruitment residuals. A value of 0.6 is supported by meta-analysis.

Fix the standard deviation of recruitment residuals at 0.6 and test the sensitivity of management parameters to 0.4 and 0.8.

Use of survey data and non-use of fishery CPUE data in model fitting

The reviewers recommended continuing to exclude the fishery CPUE data from the estimation of model parameters. One reviewer recommended excluding them completely because they might cause confusion. They recommended that the fishery CPUE data should be standardized.

Standardize the fishery CPUE indices and continue to include them in the assessment model, but not contributing to the estimation of parameters.

Other topics

Standardizing the survey

One reviewer suggested that the survey index of abundance be standardized for factors such as vessel, temperature, bottom type, location, and depth using a GLM or GAM. This reviewer also suggested mapping the habitat to improve the survey design.

Standardizing the survey for factors such as vessel, temperature, bottom type, location, and depth is a reasonable approach, but it might be better to post stratify by temperature, bottom type and depth each year rather than simply using a GLM. The habitat mapping could be used in this approach.

Jittering

Jittering the initial starting values of the estimated model parameters came up several times in the reviews. Jittering is a method to make sure that parameter estimates are the best values given the data and model assumptions. This is done because several years ago the model put forward had not converged properly and a better set of parameter values was found prior to the SSC meeting. The sensitivity of results to initial parameter values is probably caused by the selectivity curves. The need to jitter the starting values greatly increases the amount of time needed to do the assessment.

The model needs to be made more stable so it does not need jittering. This might be achieved by developing more robust selectivity curves.

Year to year changes in the model structure

The reviewers questioned the changes in model assumptions from year to year and suggested that the model structure should be fixed for a few years and the assessment only include updated data. In the years between “benchmark assessments” research could be carried out to improve the model.

Fixing the model structure for a few years is a reasonable approach to deal with several practical issues, but it would require the existence of a reasonable model. Unfortunately, and despite the substantial progress made on the Pacific cod assessment, there are still a few major issues that need to be resolved.

Tagging studies

The value of tagging studies came up several times in the reviewers’ reports. The obvious need is to determine catchability for the survey using archival tags. However, well designed conventional tagging studies could be used to provide information on selectivity and natural mortality, validate aging, estimate abundance and exploitation rates, and evaluate stock structure.

A well designed and comprehensive tagging study is highly recommended.

Alternative modeling environments

The reviewers noted that alternative modeling environments might be useful to either customize model assumptions or double check model results. Developing a completely new customized assessment model for Pacific cod with all the functionality needed for sensitivity test would be a substantial task. It would be much better to request that the Stock Synthesis code be modified into a form that makes customization easy. Stock Synthesis can be configured to replicate either exactly or approximately many other stock assessment models and it would be better to apply simplifications of Stock Synthesis rather than using other models. The main reason to use another model is to identify programming errors in Stock Synthesis.

Request that Stock Synthesis becomes more user customizable.

Over parameterization

The reviewers mentioned several times that the models are over parameterized or nearly so. I doubt if this is correct. The issue is more likely related to poor model structure and parameterization (i.e. the selectivity curves).

The models are not over parameterized, but work needs to be carried out to make the model more stable.

Aleutian Islands

The reviewers suggest that the Aleutian Islands should be considered a separate stock.

I don’t understand this issue well enough to make a recommendation.

Management strategy evaluation

The reviewers recommend continuing the management strategy evaluation work.

Management strategy evaluation (MSE) is very useful, but time consuming. Solving some of the issues in the assessment model are higher priority than the MSE work.

Diagnostics

The reviewers suggested several diagnostics that should be applied to the stock assessment including retrospective analysis, residual analysis, and evaluation of the correlation matrix to identify parameters that are highly correlated.

These are useful diagnostics and could be used to help select which model assumptions are appropriate. Retrospective analysis should not be used to determine the size or direction of the bias, only that some form of model misspecification exists.

Dynamic reference points

One reviewer noted that auto correlated recruitment may cause the abundance to drop below management reference levels even if the fishing mortality is relatively low.

Consider instituting dynamic reference points that take account of variation in recruitment