Mid-Atlantic Fishery Management Council
Scientific and Statistical Committee
OFL CV Guidance Document

Approved by Council June 2019
Revised XX 2020

Introduction

The Mid-Atlantic Fishery Management Council's (MAFMC) Scientific and Statistical Committee (SSC) currently uses a control rule to specify the acceptable biological catch (ABC; catch level that sets an upper bound for the Annual Catch Limit) for stocks that have accepted estimates of the overfishing limit (OFL; the catch that is expected to achieve the fishing mortality threshold (FMT)). The control rule is based on the P* (probability of overfishing) approach, which is used to calculate a catch level that is expected to achieve a pre-specified probability (P*) of exceeding the maximum fishing mortality rate reference point. In addition to the P*, which is specified by the MAFMC, the control rule requires a probability distribution for the OFL to describe uncertainty. Because of the difficulty in accurately quantifying the total uncertainty in the OFL, the SSC currently specifies a distribution for the OFL. The point estimate of the OFL from the stock assessment is used as the median of a lognormal distribution with a coefficient of variation (CV) specified by the SSC.

The true uncertainty in the OFL is needed to achieve the MAFMC's goal of a catch limit that meets a specific probability of overfishing. If the CV of the OFL is underestimated, the probability of overfishing will be higher than desired, and, conversely, if the CV of the OFL is too high, then the probability of overfishing will be lower than specified by the Council. The OFL CV is uncertain and difficult to estimate accurately. Three primary sources of uncertainty affect uncertainty in the OFL: uncertainty in the current stock biomass, uncertainty in the FMT and the OFL that is derived from it, and uncertainty from projecting into the future. Uncertainties in biomass and OFL derive from similar sources. Uncertainty is introduced by sampling variability when data are collected. Additional uncertainty is introduced as a result of assumptions and parameter estimates used in the assessment models. Since assessment models are simplifications of real-world, important uncertainties may be entirely uncharacterized. Therefore, the OFL is subject to substantial uncertainty, and the true uncertainty (instead of assessment model precision) is very difficult to estimate.

The ABC is derived from the OFL by assigning an appropriate level of uncertainty. The SSC believes that no single model or even ensemble of models can fully capture the full assessment uncertainty. Rigorous consideration of key assessment parameters and assumptions and comparison among model simulations can improve one's understanding of the true but essentially knowable uncertainty. This document describes the criteria used for determination of bins of uncertainty levels. Ultimately, the final determination is dependent on expert judgement and qualitative evaluation of a suite of factors that affect uncertainty of the OFL.

The MAFMC SSC has used a range of values, 60-150%, for the CV of the OFL distribution in determining the ABC. However, the SSC, MAFMC, and stakeholders have questioned the rationale for various values of the OFL CV that have been applied by the SSC as well as the consistency underlying the decisions about OFL CVs among assessed stocks. When the ABC control rule was initially adopted, a default amount of uncertainty was estimated from a meta-analysis of accuracy of estimates from simulation studies of statistical catch-at-age model performance, including the uncertainty in biomass in the last year, uncertainty in the fishing mortality reference point, and their covariance[1]. This analysis indicated that a CV = 100% was a reasonable value for the average CV of the OFL distribution. Since that time, the SSC has chosen CVs for the OFL distribution that differ among stocks (Figure 1).

The SSC's intent for this document is to elevate confidence in ABC recommendations by establishing a replicable process that meets Council risk policy objectives and identifies relevant components of assessment uncertainty to be provided to the SSC. The approach outlined here will not resolve all scientific uncertainties and problems, and exceptions will arise that are not specifically addressed in this document; however, this approach should help alleviate many issues and provide a clear, consistent, and transparent process that documents the SSC deliberations and conclusions.

The SSC's approach to setting OFL CVs will:

- Result in prudent decisions for catch advice that are consistent in meeting the objectives of the Council's Risk Policy;
- Be based on clear decision criteria that are consistently applied across stocks; and
- Be supportable with evidence.

Decision Criteria

The SSC agreed to consider nine decision criteria to help define an appropriate OFL CV when setting new or revised ABC recommendations. All decision criteria will be considered by the SSC; however, the relative importance and "weighting" of each criteria will be different for each species and consistent with the approaches and analyses evaluated within each assessment framework. In addition, while these criteria were specifically developed to help in SSC

[1] For more information, please see the SSC white paper titled "Description and Foundation of the Mid-Atlantic Fishery Management Council's Acceptable Biological Catch Control Rule" found at: http://www.mafmc.org/s/MAFMC-ABC-Control-Rule-White-Paper.pdf.

deliberations, they may also be helpful to stock assessment workgroups as they consider and evaluate data and model appropriateness and uncertainty.

The nine decision criteria are provided below. In addition, there is also supporting language that generally describes the considerations and information the SSC may utilize when considering each criterion.

1. **Data quality**
   a. Types and quality of available data are primary determinants of the accuracy of any assessment model.
   b. Important fishery-independent data considerations include survey design, coverage (of the unit stock area), and efficiency of survey gear.
   c. Fishery-dependent considerations include accuracy and precision of landings and discards
   d. Availability of age and/or length data for fishery-independent and dependent sources; validity of underlying assumptions and any potential data borrowing (i.e., gap filling)
   e. Information on natural mortality and other assumed model parameters

2. **Model appropriateness and identification during the assessment process**
   a. Model selection process and tests are important for choosing assessment models that are likely to be accurate (e.g., model sensitivities within a given model structure)
   b. Comparison among~~between~~ the assessment baseline model and models with different structures is important to determine the effects of assumptions
   c. Model appropriateness in capturing species and fishery specific traits, such ~~as fleets,~~ biological characteristics, life history patterns, spatial/stock structure, and ~~biological characteristics~~ fleets,
   d. Amount of model testing with consistent or divergent estimates (particularly for management relevant quantities like the OFL or stock status)~~trends among models~~

3. **Informed by retrospective analysis**
   a. Retrospective pattern is ~~direct~~ evidence of model misspecification and suggests directionality of change with respect to "true" or at least improved model rather than an unspecified set of alternative models
   b. Comparison of the adjusted OFL to the uncertainty of the OFL estimated from the baseline model to determine if retrospective pattern is a larger portion of uncertainty

4. **Informed by comparison with simpler analyses**
   a. Swept area biomass or gear comparisons that suggest appropriate minimum scale of population
   b. Comparison with other empirical or simpler measures; e.g., survey Z, Beverton-Holt length-based Z

5. **Informed by ecosystem factors or comparisons with other species**
    a. Stock-relevant ecosystem factors directly included in the assessment model, e.g.,:
        i. Environmentally dependent growth or other population processes
        ii. Factors limiting/enhancing stock productivity (habitat quality, etc.)
        iii. Predation, disease, or episodic environmental mortality (e.g., red tide)
    b. Ecosystem factors outside the stock assessment affecting short term prediction
        i. General measures of ecosystem productivity and habitat stability (e.g., primary production amount and timing, temperature trends, etc.)
        ii. Comparisons among related species; e.g., recruitment, growth, condition patterns across Mid Atlantic fish species stable, varying synchronously, or varying unpredictably
        iii. Climate vulnerability or other risk assessment evaluation of potential for changing ~~increasing or decreasing~~ productivity under changing conditions

6. **Informed by measures of trend in recruitment (primarily affecting the accuracy of forecasts)**
    a. Stanzas of abundance for recruits
    b. Decreasing R/SSB as SSB decreases (evidence of depensation)

7. **Informed by prediction error**
    a. Comparisons of model performance given prior assessments
    b. Consistency among repeated assessments should be considered in light of changes in the best available information or understanding of stock and fishery dynamics

8. **Assessment accuracy under different fishing pressures**
    a. Age-structured assessment approaches are generally more accurate under higher fishing mortality rates relative to natural mortality.
    b. Non-age-structured assessment approaches may require specific patterns in the data to be highly accurate (e.g., high contrast in abundance and fishing pressure for a production model)
    c. Prediction error and dynamic trends (e.g., decadal periods) in fishing selectivity patterns

9. **Informed by simulation analysis or full MSE**
    a. Simulation analyses can be used to test how robust assessment approaches or management strategies are to specific misspecifications in the models or issues in the data.

General Framework Discussion Table

The framework table is intended to provide qualitative assessment of the nine criteria and is not to be used to tabulate a specific score. Instead, the table will help document the SSC deliberations, ensure a consistent process is followed for all species and assessments, and help the Council and public understand the rationale for the decision reached by the SSC.

The table currently has OFL CV default values of 60%, 100%, and 150%, and were derived from a variety of simulation analyses, MSE evaluations, and expert judgement by the SSC. As new information, analyses and assessment methods become available, the SSC may modify the default OFL CV bins or recommend a different OFL CV for a specific species assessment. If any changes to the current default OFL CV values are warranted, the SSC will provide justification and supporting documentation as to why a different value was recommended.

The framework table below provides general evaluation metrics associated with the nine decision criteria for each OFL CV bin.

Decision Criteria	Default OFL CV=60%	Default OFL CV=100%	Default OFL CV=150%
Data quality	One or more synoptic surveys over stock area for multiple years. High quality monitoring of landings size and age composition. Long term, precise monitoring of discards. Landings estimates highly accurate.	Low precision synoptic surveys or one or more regional surveys which lack coherency in trend. Age and/or length data available with uncertain quality. Lacking or imprecise discard estimates. Moderate accuracy of landings estimates.	No reliable abundance indices. Catch estimates are unreliable. No age and/or length data available or highly uncertain. Natural mortality rates are unknown or suspected to be highly variable. Incomplete or highly uncertain landings estimates.
Model appropriateness and identification process	Multiple differently structured models agree on outputs; many sensitivities explored. Model appropriately captures/considers species life history and spatial/stock structure ~~(e.g., black sea bass)~~.	Single model structure with many parameter sensitivities explored. Moderate agreement among different model runs~~alternative models~~ indicating low~~or~~ sensitivities of model results to specific parameterization.~~.~~	Highly divergent outputs from multiple models or no exploration of alternative model structures or sensitivities.
Retrospective analysis	Minor retrospective patterns. ~~No retrospective adjustment necessary, or OFL estimate includes retrospective adjustment.~~	Moderate retrospective patterns. OFL estimate includes retrospective adjustment only if outside 95% bounds of non-adjusted terminal B and F.	No retrospective analysis or severe retrospective pattern observed.
Comparison with empirical measures or simpler analyses	Assessment biomass and/or fishing mortality estimates compare favorably with empirical estimates.	Moderate agreement between assessment estimates and simpler analyses.~~Both assessment biomass and/or fishing mortality empirical estimates highly uncertain.~~	Estimates of scale are difficult to reconcile and/or no empirical estimates.
Ecosystem factors accounted	Assessment considered habitat and ecosystem effects on stock productivity, distribution, mortality and quantitatively included appropriate factors~~, reducing uncertainty in short term predictions~~. Evidence outside the assessment	Assessment considered habitat/ecosystem factors but did not demonstrate either reduced or inflated short-term prediction uncertainty based on these factors. Evidence outside the assessment suggests that ecosystem	Assessment either demonstrated that including appropriate ecosystem/habitat factors increases short-term prediction uncertainty, or did not consider habitat and ecosystem factors. Evidence outside the assessment

Commented [1]: I know this is NEFSC practice, but may/may not be appropriate. It means that if the precision of estimates is low enough, there will never be an adjustment for a retrospective pattern. Also, for extremely precise models, there will always be adjustments for retrospective patterns.

The description of retrospective adjustments is out of place here. Maybe move to the first column.

Commented [2]: Do we know how many assessments have this accessory information. If only few do - then aren't we biasing this toward CV=150%? Is this what we want to do?

Commented [3]: I'm not sure how this would be shown.

Commented [4]: I think you could show this by comparing model runs with and without including the relevant ecosystem factor(s); therefore I would suggest leaving this statement in the description

	suggests that ecosystem productivity and habitat quality are stable. Comparable species in the region have synchronous production characteristics and stable short-term predictions. Climate vulnerability analysis suggests low risk of change in productivity due to changing climate.~~positive impacts on productivity from changing climate~~.	productivity and habitat quality are variable, with mixed productivity and uncertainty signals among comparable species in the region. Climate vulnerability analysis suggests moderate risk of change in productivity~~neutral impacts on productivity~~ from changing climate.	suggests that ecosystem productivity and habitat quality are variable~~quality variable~~ and degrading. Comparable species in the region have high uncertainty in short term predictions. Climate vulnerability analysis suggests high risk of changing ~~negative impacts on~~ productivity from changing climate.
Trend in recruitment	Consistent recruitment with no trend~~OFL estimates adjusted for recent trends in recruitment.~~	OFL estimates adjusted for recent trends in recruitment. ~~No recruitment trend or uncertain. Insufficient evidence to adjust OFL estimate based on recruitment information available.~~	Recruitment highly inconsistent, trend not considered or no recruitment estimate.
Prediction error	Low estimate of recent prediction error.	Moderate estimate of recent prediction error.	High or no estimate of recent prediction error.
Assessment accuracy under different fishing pressures	High degree of contrast in landings and surveys with apparent response in indices to changes in removals. Fishing mortality at levels expected to influence population dynamics~~Observed high fishing mortality near the target~~ in recent years.	Moderate agreement in the ~~contrast in~~ surveys to changes in ~~and~~ catches. ~~"One-way" trips for production models.~~ Observed moderate fishing mortality in fishery (i.e., lack of high fishing mortality in recent years).	Relatively little change in surveys or catches over time. Low precision of estimates. Low fishing mortality in recent years. "One-way" trips for production models.
Simulation analysis/MSE	Can be used to evaluate different combinations of uncertainties and indicate the most appropriate OFL CV for a particular stock assessment.		

A worked example evaluation of the nine criteria provided in the table above is provided for summer flounder (see page 8).

Process for OFL Determination

The SSC's consideration, evaluation, and discussion of the nine decision criteria in determining the appropriate OFL CV level could potentially become cumbersome and time-consuming to be handled effectively during an SSC meeting, particularly if multiple species-specific ABC recommendations are required. In an effort to add efficiency to the ABC-setting process while still allowing for extensive SSC input and discussion, the SSC will develop a pre-decisional, non-binding document evaluating the nine decision criteria ahead of the SSC meeting. This document will then be posted as part of the SSC meeting materials and available to SSC members for review ahead of the meeting in which an ABC recommendation is required. The process for developing the pre-decisional document and the SSC's OFL CV determination will follow the steps outlined below:

- Upon completion of a stock assessment, the appropriate SSC species lead ~~will work with the stock assessment lead and Council staff to~~ evaluate the nine decision criteria and

Commented [5]: Why would this affect CV?

Commented [6]: agreed, reworded to align risk of climate driven change in productivity (increase or decrease) with CV: low CV with low risk of change, high CV with high risk of change

Commented [7]: I suggest swapping these two descriptions. I think the best case scenario is no trend in recruitment. That means our previous estimates are relevant for projections. A trend in recruitment means that things are changing and that our past estimates may not be relevant for the future. To me, this implies more uncertainty.

complete a draft framework discussion table, seeking input from Council Staff and the stock assessment lead as necessary.

- The draft framework table would then be provided to a sub-group of SSC members (initially comprised of the OFL CV workgroup members) for review and feedback. This group would meet to discuss the draft characterization via webinar prior to the SSC meeting ~~and would likely meet via webinar~~. This sub-group will help ensure consistency in the interpretation and evaluation of the decision criteria. The SSC species lead, and the sub-group will then develop a draft narrative summarizing the key findings based on the draft framework table. This narrative will include information on the most critical and important decision criteria specific to the species and stock assessment reviewed and highlight any other areas of extended deliberation by the sub-group. The narrative will also recommend an appropriate OFL CV level for SSC consideration. The framework table, narrative and OFL CV recommendation will all be labeled as draft and are pre-decisional and non-binding. The SSC will use the assembled information as a guideline for decision making following a rigorous an open discussion during its meetings.
- The draft framework table and narrative will be provided to the full SSC and posted as part of the meeting materials in advance of the meeting in which the ABC recommendations will be made.
- During the SSC deliberations to address the ABC Terms of Reference, the SSC species lead will provide an overview of the decision criteria evaluation, narrative, and OFL CV recommendation.
- SSC members present at the meeting will then discuss and deliberate any/all information available in order to make an OFL CV recommendation. The SSC meeting summary report will contain both the framework table with an evaluation and rationale of the nine decision criteria and a summary narrative. Providing both the framework table and narrative in the meeting summary will help provide a comprehensive record of the SSC's deliberations and justification for their recommendation for future reference.

Given the additional work and meetings necessary prior to a scheduled SSC meeting as outlined above, increased coordination between the NEFSC, Council staff and the SSC will be critical to ensure stock assessment documents and information are available in a timely manner. Ideally, stock assessment documents and any other pertinent information would be available three weeks prior to the scheduled SSC meeting. The SSC species lead and the sub-group would meet two weeks prior to the scheduled SSC meeting to review the framework table and develop the draft narrative. The sub-group documents would then be available to the SSC and posted to the meeting materials at least one week prior to the scheduled SSC meeting. Delays in any part of this process could result in a number of implications ranging from inefficient and extended SSC meetings to potential delays in the making ABC recommendations. In addition, continued SSC involvement in the SAW/SARC process (i.e., chairing SAW/SARC assessment reviews, embedding with the assessment work group) will play a critical and informative role in the process to help ensure the timing and deadlines are achieved.

Commented [8]: I have flipped this sentence - the original suggested the SSC lead had to work with the assessment lead and Council staff - this edit indicates that is "may" work......

Commented [9]: This part still raises concerns
1) It adds an additional step between receiving the assessment and the SSC meeting. Experience indicates that assessment teams are often up against deadlines at the best of times, and this may challenge our ability to have pre-decisional meetings
2) Concerned over the optics of this since this is a large group of the SSC having a non-publicly advertised meeting to make decisions/recommendations prior to the SSC meeting.

Worked Example

Below is a worked example for summer flounder based on the results of the 2018 benchmark assessment. The worked example includes the SSC OFL recommendation, an evaluation of the nine decision criteria as outlined in the framework table and a short narrative documenting key conclusions.

Based on an evaluation of the nine decision criteria, the SSC recommends a CV of 60% be applied to the OFL estimate as an appropriate ABC for summer flounder in fishing years 2019-2021.

Decision Criteria	Default OFL CV=60%	Default OFL CV=100%	Default OFL CV=150%
Data quality	Two synoptic surveys (fall and spring) are available for all years in assessment. Additionally, 13 regional surveys are used in model tuning. Time series for R/V Albatross and R/V Bigelow treated separately for spring and fall trawl surveys. Bigelow estimates adjusted for results of cooperative research studies on gear efficiency. Age data available for all years in surveys, and age-length keys from surveys were applied to commercial landings, recreational landings, and commercial discards. Recreational and commercial discards are low and measured with good precision. Sex-specific information available for growth. Newly revised historical MRIP catch estimates were used in assessment.		
Model appropriateness and identification process	Models incorporating age and sex-specific growth and mortality rates were developed, tested, and reviewed. Multiple models by different assessment teams were considered. ASAP was preferred assessment model but SS and other statistical catch-at-age models were considered. These include models with age and sex-dependent rates of natural mortality, growth, and fishery selectivity. However, additional work on the more complicated models is needed to appropriately evaluate to the single sex models.		
Retrospective analysis	Retrospective pattern in current assessment is minor with retrospective errors over the last 7 terminal years averaging -4% for F, +2% for SSB, and +2% for recruitment. These retrospective errors are about one-tenth as large as their magnitude in the previous benchmark assessment.		

	Historical retrospective comparisons show general trends of fishing mortality, stock biomass, and recruitment have been consistent since the 1990s assessments.		
Comparison with empirical measures or simpler analyses	Assessment biomass and/or fishing mortality estimates compare favorably with empirical estimates. Results of cooperative research gear experiments were used to adjust scale of biomass indices used in model tuning.		
Ecosystem factors accounted		Aspects of the ecosystem seem to be changing in recent years. Fall ocean bottom and surface temperatures are increasing, and salinity is at or near the historical high. These physical data series may have shifted around 2012, the warmest year on record for this ecosystem. Spring chlorophyll concentrations, a measure of bottom-up ecosystem production in the summer flounder stock area, are variable, but the fall time series has been decreasing, especially during 2013-2017. Spring abundances for key zooplankton prey are variable and may be worth examining alongside recruitment patterns for future research. Both probability of occurrence and modeled habitat area show similar patterns of increases from the 1990s to the present, which suggests, despite reduced abundance in the past five years, the distribution footprint of summer flounder has not contracted.	
Trend in recruitment		Average recruitment from 1982 to 2017 is 53 million fish at age 0. Recruitment has been below average since 2011, averaging 36 million fish. Overall recruitment variability is modest and it is not possible to determine if recent decline is statistically significant. Projections do not account for recruitment trend.	
Prediction error	Prior to the 2018 benchmark, comparisons of annual forecasts of stock biomass with realized estimates of stock biomass in subsequent assessments reveal a one-year ahead forecasting error with a CV=14%. For two-year forecasts the CV is 26% and for 3 year forecasts the CV= 26%. The average percentage difference between the projection and the subsequent estimate for 1-, 2-, and 3-yr projections was +12%, +23%, and +24%, respectively. Inclusion of the revised MRIP data		

	increased the population scale, rendering prediction comparisons less useful as a metric of model performance.		
Assessment accuracy under different fishing pressures		Fishing mortality has varied over a 6-fold range over the assessment period with major decline since imposition of effective management measures around 2000. This range of fishing mortalities, subsequent fluctuations in total abundance, and success of management changes suggest a moderate level of confidence in assessment results.	
Simulation analysis/MSE	No formal MSE-type analyses have been conducted for this stock.		

Example OFL CV Recommendation Narrative

This is a data rich stock assessment and one of the most comprehensive in the Northeast US. Two synoptic surveys (fall and spring) are available for all years and multiple regional surveys are used in model tuning. Age data available for all years in surveys, commercial landings, recreational landings, and commercial discards. Recreational and commercial discards are low and measured with good precision. The newly revised MRIP catch estimates were incorporated into the assessment for the first time. Extensive work on alternative model formulations (including size and sex-based models) have been conducted by independent assessment teams. Spatial variations in catch rates by sex and fisheries have been examined. Multiple model formulations have been systematically evaluated. More complicated models have not been judged superior to single-sex models. The retrospective pattern for the current assessment is exceptionally low and comparisons of biomass estimates across historical assessments show good agreements with trend. Estimates of prediction error for 1- to 3-year forecasts are less than 25%. The stock has experienced a wide range of fishing mortality rates and appears to have responded as predicted by theory to aggressive management measures in the early 2000's. This suggests a high level of confidence in the results.

Consideration of ecosystem factors apart from the model suggest some cause for concern as increases in temperature and salinity have occurred especially since 2012. It is too early to tell if changes in chlorophyll indices and zooplankton abundance are related to recent reductions (about 31% decline) in average recruitment in this same period.

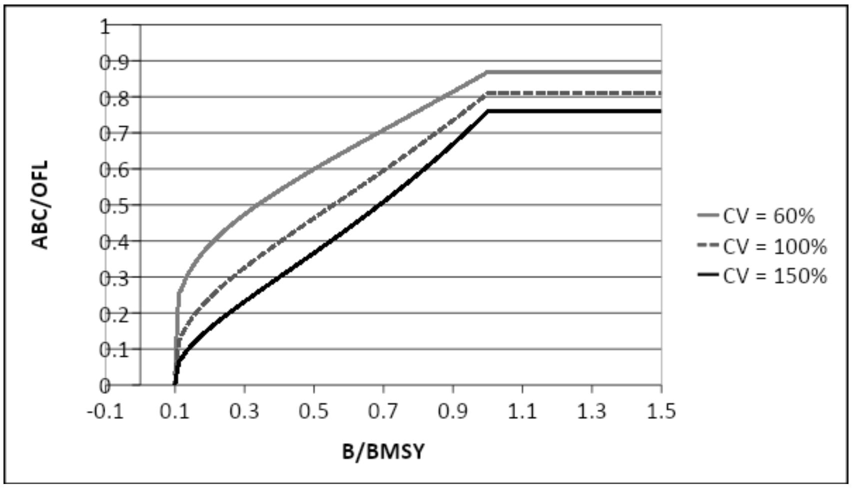

Figure 1. Effect of different CV values currently selected by the MAFMC SSC on the ratio of ABC to OFL for varying levels of biomass relative to the B_{MSY}.