Preliminary Draft for Full SSC Review

SSC Workshop on Risk Tables for ABC Advice to Council

Compiled by the North Pacific Fishery Management Council's
Scientific and Statistical Committee
June 2021

Contents

Executive Summary and Recommendations

In February 2021, the SSC convened a workshop to: evaluate how the risk table process was working; address consistency issues with the risk table as identified by the GPTs, authors, and SSC; and to provide guidance for moving forward. The SSC appreciates hearing about the challenges the authors and JGPT have found with the risk table process.

The workshop objectives were to:

1. Assess the progress and value of species-specific risk tables for all stocks
2. Evaluate risk table consistency among species and highlight challenges
3. Define "risk" and "uncertainty"
4. Compare ABC and OFL buffers for scientific uncertainty with ABC reductions due to the risk table
5. Discuss future options

The workshop included two plenary sessions and breakout sessions providing open discussion between stock assessment authors, Plan Teams and SSC members (session leads in parentheses, see February Workshop Agenda Appendix 1). Time was set aside for public testimony relevant to the workshop at the end of the workshop. The SSC appreciates the important contributions from the topic session leads and all participants that contributed to the discussions.

In February, the SSC established the following timeline and process for finalizing guidance to stock assessment authors.

Date/Meeting	Action	Who
June 2021	Assemble full workshop report	SSC/Workshop Session lead
June 2021	Preliminary recommendations	SSC/NPFMC
September 2021	Comment and recommendations	CPT/GPT
October 2021	Finalize 2021 recommendations	SSC/NPFMC

This report provides a short written description of each topic session and a summary of the key findings. Based on this summary, the SSC provides the following preliminary guidance on the development and use of Risk Tables.

Preliminary Guidance and SSC Recommendations

- The SSC concluded that the risk table framework is working well. The tables have expanded communication among assessment authors and between assessment authors and ecosystem/process researchers. The framework is intended to provide a clear and transparent basis for communicating assessment-related and stock condition concerns that are not directly captured in model-based uncertainty, the tier system, or harvest control rules.

- The SSC recognizes that within the context of the risk tables, "risk" is the risk of the ABC exceeding the true (but unknown) OFL. The risk tables are intended to inform the process of adjusting the ABC from the maximum when needed. Recommendations of an ABC reduction from the maximum permissible requires justification. The risk tables provide an avenue for articulating that justification.
- The SSC recommends that risk tables are produced each year for all groundfish (and perhaps crab) stocks and stock complexes in the fishery. The SSC requests that the authors consider if there have been any changes to previous conditions and update the tables accordingly. The SSC recommends that authors of stock complexes consider the dominant and weakest members of the complex when formulating advice.
- Stocks for which the expected catch is much lower than the ABC still need a risk table. This should avoid a situation where the SSC assumes, incorrectly, that there is little risk for a given stock, but the Council then sets the TAC higher than expected.
- Risk scores should be specific to a given species or species complex. While comparison across species (e.g., within a tier, with similar life histories) is useful for consistency, the SSC does not support trying to prescribe a common reduction from maxABC for a given risk score across species because the processes underlying the score may differ among stocks. The SSC recommends that considerations of reductions in ABCs below the maximum permissible be made on a case-by-case basis with justification based on risk scoring. The risk table rankings include qualitative information that requires a certain amount of subjective but well-informed interpretation of the available data by the author(s), the Plan Teams and the SSC, and as such the SSC feels that blanket comparisons across species for the purpose of explicitly defining ABC reductions are not prudent.
- The SSC recommends that the fishery/community performance column should focus on information that would inform the biological status of the resource (e.g., an unexplained drop in CPUE that could indicate un-modelled stock decline, or a spatial shift indicating changes in species' range), and not the effects of proposed ABCs on the fishery or communities or bycatch-related considerations. The SSC recognizes that the community impact information is critical for informed decision making for TAC setting and recommends this information be included in other Council documents such as the ACEPO and/or the Economic SAFE.
- The SSC encourages the inclusion of LK/TK/S as a source of knowledge about the condition of the stock.
- The SSC appreciates the discussion of avoiding double-counting information, in the assessment/Tier system and risk table, or among columns of the risk table. The SSC agrees that authors should avoid inclusion of stock trends/processes that are incorporated in the assessment or reflected in the Tier ranking in risk tables. For cases where a process external to the assessment is relevant to two or more risk categories, the SSC recommends that the narrative reflect the interconnected relationships that exist between rankings among risk categories.
- The SSC suggests a potential revision to the category levels: from the existing four to three categories (normal, increased, extreme).
- The SSC reiterates that reductions in ABC below the maximum permissible are intended to be infrequent and only occur under exceptional circumstances. If they begin to become commonplace, that should warrant further review of the assessment and/or the Tier system.

Acknowledgements

The SSC expresses its sincere appreciation for the contributions of the discussion leads and rapporteurs. The SSC is grateful for the leadership of these individuals in preparing for, and contributing to, discussions during the workshop, and their work to develop summaries contained in this report.

Discussion 1: Introduction to Risk Tables: historical background, conceptual framework, synthesis of applications in 2018-2020.

Martin Dorn, Stephani Zador, Kalei Shotwell
(Discussion Leads & Rapporteurs)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Historical background and the genesis of risk tables

An explicit part of the NPFMC stock assessment process is an evaluation of whether it is appropriate to reduce the ABC from the ABC resulting from application of the control rules in the Tier system. As described in both the BSAI and GOA groundfish FMPs, groundfish stock assessments should "*determine whether conditions exist that warrant setting ABC at a value lower than the maximum permissible value (such conditions may include—but are not limited to—data uncertainty, recruitment variability, and declining population trend) and, if so:*

a. document those conditions,
b. recommend an ABC lower than the maximum permissible value, and
c. explain why the recommended value is appropriate.

The above steps are undertaken first by the assessment authors in the individual chapters of the SAFE report. The Plan Team then reviews the SAFE report and makes its own recommendation. The SSC then reviews the SAFE report and Plan Team recommendation, and makes its own recommendation to the Council. The Council then reviews the SAFE report, Plan Team recommendation, and SSC recommendation; then makes its own recommendation to the Secretary, with the constraint that the Council's recommended ABC cannot exceed the SSC's recommended ABC."

The NPFMC tier system is designed to be a precautionary system in which buffers are already in place to achieve a preferred degree of conservatism. Therefore the rationale for a reduction from the maximum permissible ABC should be that there is either additional uncertainty in the assessment and/or additional risks (probability of something bad happening) to the stock that are not adequately taken into account by the default precautionary settings. The risks generally relate to a loss of fishery sustainability and an inability of the stock to perform its role in a functioning ecosystem, such as might occur due to severe decline in stock abundance. This understanding of risk is consistent with how risk is understood in the context of ecological systems (e.g. see Holsman et al. 2017), and the concept is broader than just the uncertainty associated with the assessment, though of course assessment uncertainty is an important element of it.

For example, in 2006 a reduced ABC for EBS walleye pollock was justified in part due to an increase in biomass of juvenile pollock predators and an apparent lack of pollock prey (Zador et al., 2017). The SSC's intent is that setting the ABC below the maximum permissible should be applied sparingly and that the tier system should be regarded as the primary basis for establishing the ABC. It is also important to note that the sloping harvest control rule for the ABC will substantially reduce the target fishing mortality rate when the stock is at a low abundance. This reduction in the fishing mortality rate is intended to address the concerns related to low stock abundance.

In 2018, the NPFMC SSC recognized that the process of considering whether to reduce the ABC below the maximum permissible was a long-standing aspect of scientific advice that is provided to the NPFMC. However, the magnitude of the reduction and the criteria used to justify the reduction had not been standardized across groundfish species. The NPFMC SSC therefore encouraged the development of a more objective and rigorous process for considering ABC reductions that included a review of both stock assessment and ecosystem factors. In February 2018, the NPFMC SSC requested that a workshop be held to address the topic of adjustments made from the maximum permissible ABC, and asked for the identification of clear and transparent rules for defining the specific criteria to be used when adjusting the ABC.

To provide an overview of historical practice of recommending ABCs less than the maximum permissible, Thompson (2018) provided the workshop with a review of the annual stock assessments from 2003 to 2017 and identified all instances when the Plan Teams recommended setting the ABC below the maximum permissible (also see Discussion 3 this report). During the 15 years, the Plan Teams recommended setting ABC below the maximum permissible in a total of 76 instances (roughly five per year). Reasons varied but generally grouped around concerns regarding the stock assessment (e.g., uncertain survey estimates or parameter estimates), population dynamics (e.g., poor recruitment or declining biomass), or ecosystem considerations (e.g., predation pressure or bird die-offs, though reductions due to ecosystem considerations were relatively uncommon). In some cases, economic factors were cited, such as variability in yield or the amount of effort required to catch the ABC. The buffers ranged from less than 10% to greater than 90%, but were most often between 10% and 30%, with a mode at a buffer of 15%.

One shortcoming of the historical analysis is that documentation existed only when there was a recommended reduction. In some cases, an evaluation identified various concerns, but a reduction was not recommended because the conditions were regarded as not sufficiently extreme to warrant a reduction. In other cases, no evaluation was made. It was noted that one advantage to applying a framework consistently for all stocks was that it would establish a stock-specific record of concerns and issues with the assessment, population dynamics, and the ecosystem. There would be supporting documentation in situations where the maximum permissible ABC was considered scientifically appropriate. There are many stocks in the North Pacific with reliable stock assessments, are at healthy levels of abundance, and have no severe environmental/ecosystem concerns, and documentation of these cases is important for a balanced perspective.

In response to the 2018 workshop, an Ad Hoc working group was assigned to develop the risk table framework. At that time, the Ad Hoc working group recommended a framework that distinguishes between three types of considerations (assessment, population dynamics, and environmental/ecosystem). Within each type of consideration, there is a range of concern from level 1 (no concern) to 4 (the highest level of concern) (Table 1). As a standard part of the annual stock assessment process, assessment authors and ecosystem scientists assign risk levels by qualitatively evaluating each of the three types of considerations using available information that is not modeled analytically in the stock assessment model, but which might inform a decision on the ABC in the current year. This distinction is important to avoid double counting. Information in the risk table comes either from the assessment itself, in the case of assessment uncertainty and population dynamics or, in the case of environmental/ecosystem information, from two main sources: the ecosystem status report, and the species-specific ecosystem and socioeconomic profiles (ESP, Shotwell 2018, 2020) that are available for some North Pacific stocks. The ecosystem status reports contain a broad range of ecosystem indicators that reflect ecosystem-level processes. The ESPs contain ecosystem indicators that are linked to the stock through known mechanistic relationships. This information was combined to inform risk tables as part of the harvest specification process, with the caveat that indicators in the ecosystem status reports have to be interpreted with respect to the particular stock. In December 2018, SSC recommended of adding a fishery performance column.

The initial risk levels are assigned by the assessment authors and included in the draft stock assessments. They are then reviewed and adjusted through the same annual review process as the stock assessment. The amount of any recommended reduction needs to be clearly stated along with the risk table, with an explanation of how this value was selected.

A timeline on the introduction and use of risk tables in the NPFMC harvest specification process is as follows:

- In October 2017, following the near collapse of GOA cod stock, SSC specified the need for a formal way to evaluate ecosystem conditions within the stock assessment. "*The SSC also recommends explicit consideration and documentation of ecosystem and stock assessment status for each stock … to aid in identifying stocks of concern.*"
- In summer of 2018 an Ad Hoc working group formed in response to an SSC request to develop a consistent approach to recommending ABC reductions
- Fall 2018: ABC adjustment workshop recommendations presented to Plan Team/SSC, draft risk tables developed for 5 stocks with assessment, population dynamics, and ecosystem columns.
- December 2018: SSC/AP/Council recommended that risk tables be done for all assessments in 2019: *"Additional environmental, ecosystem or other species specific biological concerns that the Plan Team identifies that are not addressed in the stock assessment model should be clearly documented and provided to the SSC for consideration...The Council supports the SSCs recommendation of adding a fishery performance column."* - Council motion Dec 2018.
- June 2019: SSC recommended that *"The combined efforts of developing ESPs for key species, the planned fall and spring meetings of the Ecosystem Status Report team to assess ecosystem change, and the development of risk tables should provide the information needed to inform the NPFMC of relevant ecosystem change...In addition, risk tables only need to be produced for groundfish assessments that are in a "full" year in the cycle."*
- Fall 2019: risk tables are completed for all full assessments, SSC recommends dropping the overall risk score and provides direct responses to ten requests raised by Plan Teams.
- December 2019: Council reiterated the dual purpose of the risk table *"...1) to facilitate further collaboration and communication among stock assessment scientists and those in other disciplines (for example, ecosystem and climate scientists) and 2) to increase transparency and consistency in the rationale for reducing from maximum permissible ABC based on exceptional risks/circumstances that are not already addressed in the stock assessment, tier system, and harvest control rules."*
- Fall 2020: risk tables completed for all full assessments, dedicated risk table workshop slated for the February 2021 SSC meeting and September 2021 Joint Plan Team meeting.

Synthesis of Stock Responses

In preparation for the workshop on risk tables, we collated the minutes from previous Plan Team, SSC, AP, and Council meetings regarding risk tables and their development from 2017 to 2020. These minutes were provided to the SSC prior to the start of the workshop (Shotwell, 2021). Additionally, each SAFE report that conducted a risk table contains between two to twelve pages on risk table evaluation (not including the risk categories definition table). We generated a summary table of the risk table scores for the stocks that have completed risk tables from 2018 through 2020 (Table 2). To date, there have been fifty-three total risk tables completed, five in 2018, nineteen in 2019, and twenty-nine in 2020, indicating that the number of risk tables completed each year has increased since the introduction of the risk tables in 2018. Over that time, only six stocks have used the risk tables to reduce from maximum ABC, four in 2018, four in 2019 and three in 2020. Only two stocks (Alaska sablefish and EBS pollock) have proposed a reduction in all three years. There have also only been three stocks with levels greater than 2 (Alaska sablefish, GOA Pacific cod, and BSAI blackspotted and rougheye rockfish), and one (GOA Pacific cod) was downgraded by the SSC from a level 4 to a level 2. In contrast, there have been twenty-two stocks with all categories at level 1 and fourteen stocks with no reduction that had at least one level greater than a level 1. This summary demonstrates the rarity of reductions from maximum ABC using this tool.

Lessons Learned

After three years of implementing risk tables with stock assessments, there are some general lessons we have learned. First, using the standard risk table format whether or not a reduction was ultimately recommended has provided a level of transparency that did not exist with the previous ad-hoc method, where there was no way to compare one instance where a reduction was applied to another because rationales were specific to a stock. The succinct synopsis of concerns for multiple categories provided in risk tables has allowed stakeholders to easily compare concerns across stock assessments. Previously, such comparisons were more difficult as it required in depth understanding of multiple stock assessments as both the types of concerns and levels of explanation provided varied among assessments. Anecdotally, this transparency has created some uncomfortable moments where stakeholders asked questions about apparent differences in responses to similar concerns across assessments that authors could not easily answer. We believe that the risk tables will facilitate research on these types of questions that may allow us to move towards standardizing maxABC reductions across assessments.

Second, risk tables have established a record of concerns that were considered in developing scientific recommendations, whether or not a reduction was recommended. Previously, concerns were not recorded if no reduction was recommended. Analysis of historical reductions as related to concerns could only be informed by data conditioned on there being a reduction, excluding the same concerns that existed for an assessment model but did not justify a reduction (“zeros”). Over time, these records will allow analysis of historical decision-making that can inform future decision-making.

A third benefit of developing risk tables is that the process of building the tables has fostered collaboration among assessment authors and ecosystem scientists. A process has developed that assigns at least one Ecosystem Point of Contact (POC) to each stock assessment. After meetings to exchange information gaps and needs, the Ecosystem POC compiles relevant information and text for the author to use as needed to complete their risk tables. Grouping meetings among stock assessments authors of stocks with similar life histories and using standardized categories of predators, prey, competitors, and environmental processes for ecosystem information supports further collaboration and knowledge-sharing. The previous ecosystem text within stock assessments varied among assessments in both breath and how often they were updated to address current state of knowledge. As ESPs are produced, they will replace the old ecosystem text. But risk tables will always be a record of the current state of knowledge, including information from ESPs and ESRs, to inform the current ABC.

A fourth benefit to the risk tables is that they can document unusual or unexpected observations that are not addressed in the stock assessment model. Recent phenomena such as marine heatwaves and changes in fish distribution have led to observations/data that assessment models cannot fit well. In these cases, having a standard place to record these observations—essentially caveats to the estimated maxABC—serves as a record of the best available science at the time of the assessment, which fisheries management relies upon. There is a lag inherent in the scientific process; observations/data need to be analyzed before relationships can be quantified and incorporated into stock assessment models. Usually, mechanistic relationships between environmental variables and biological responses are considered stationary. Thus, ignoring current observations that seem to indicate that underlying biophysical mechanistic relationships have changed (ie., evidence of non-stationarity) could lead to inaccurate predictions. The documentation of unusual/unexpected observations in risk tables included in the stock assessments provide scientific context for future development of stock assessment models. This could be especially important in years when an assessment model is not fully updated. A risk table in a year without a full update could document concerns that did not exist when the stock assessment model was last fully updated. This would serve to be a transparent record of best available science, that over time could help to prioritize research, explain past decision-making, and other efforts to build trust in the robustness of our fisheries management process.

Challenges of producing risk tables have also come to light during the three years of their implementation. As mentioned above, the increased transparency and ease of comparing across assessments has highlighted some past inconsistencies across assessments in the level of response (ie., amount of reduction from maxABC) to the level of concerns. There has been interest in having a standard reduction level in response to concerns. Risk tables provide the baseline data to inform research on the feasibility of developing standardized responses. It is also possible that the increased transparency of responses to concerns as summarized in risk tables may encourage communication among stock assessment authors and review bodies that would result in a shift over time to more consistent levels of response.

An additional set of challenges that have come to light are how to produce risk tables for bycatch/non-target stocks, stock complexes and Tier 5 and 6 stocks. The question arose as to whether it is appropriate to provide justification for or against reduction against maxABC (i.e., complete a risk table) when the stock is not targeted or caught only as bycatch. In this case, concerns would be relative to the overall impact of the maxABC on the stock, even if it is known that the catch will not approach the final ABC. Producing risk tables for stock complexes was challenging due to a number of factors related to the differences in the amount of scientific knowledge for individual species within the complex. One strategy included focusing on a single species in the risk table to represent the complex, such as Dover sole in the GOA deepwater flatfish stock assessment. Similarly, producing risk tables for Tier 5 and Tier 6 stocks were challenged by both the limited amount of scientific knowledge of the stock and the discontinuity between the ABC and the catch.

Discussions have also taken place about how to know which information goes into which column of the risk tables. For example, heatwave-level temperatures could lead to below average survival of larval/age-0 fish, which is a population dynamics concern. However, heatwave-level temperatures could also have a negative impact on prey availability for all age-classes in that same stock. This would be an environmental/ecosystem concern. Is having the same temperature time series noted in two risk table columns double-counting? The concern about double-counting is based upon an assumption that having variables listed multiple times would artificially elevate the importance of that variable. However, an alternate explanation is that if that variable has impacts in more than one distinct aspect of a fish stock's structure and dynamics as intended to be modelled well in the stock assessment model, then the variable can be listed in multiple columns. This type of double-counting is distinct from the double-counting that is explicitly not allowed for something that is already addressed in the stock assessment model. For example, in the pilot year for risk tables, the GOA Pacific cod risk table included the heatwave as part of an overall level 4 ecosystem concern on the stock size. However, as the precipitous decline in stock size was already captured by the stock assessment model, the level of concern was downgraded to a 2.

Discussion and Questions to Consider

The SSC has provided guidance on the use of risk tables several times during the last several years (see collated minutes). In some cases, this guidance has altered the responsibilities of assessment authors and Plan Teams, and it is important to consider the implications of these changes. The first issue is whether assessment authors and Plan Teams should feel bound to provide a recommendation on an ABC reduction when the risk table indicates an increased concern. The SSC minutes of December 2018 advised that the author and Plan Team do not have to recommend a specific ABC reduction, but should provide a complete evaluation to allow the SSC to come up with a recommendation. Although the FMPs for groundfish do give assessment authors and the Plan Team a role in making ABC recommendations, authors may prefer to avoid making non-model based recommendations and defer this role to the SSC. *Does the SSC have the desire and capacity to take on this larger role? Does the SSC depend on authors and Plan Teams to generate options for consideration, even if the SSC chooses a different course?*

A second related issue is whether an increased risk requires a reduction in ABC. The SSC minutes of December 2019 advised that adjustment from maxABC in response to levels of concern should be left to the discretion of the author, the Plan Team, and/or the SSC, but should not be mandated by the inclusion of a >1 level in any particular category. The stated purpose of the risk table to support a decision whether or not a reduction in the maximum permissible is needed. *What is the intended message if the table is filled out appropriately, and an increased risk is identified, but no reductions are recommended? That increased risk does not matter?* In several cases, authors decided not to recommend an adjustment in the ABC when recent catches were far below the ABC, since there was nothing that would be accomplished by a reduction in the ABC. Some guidance on this situation may be helpful since finding a defensible basis for an ABC reduction can be difficult and time consuming.

The section leads provided a series of question prompts at the end of the presentation for discussion in the break out groups. Many of these questions are related to issues described above. Responses to these prompts are embedded in the reports from the break out groups.

- *How to distinguish double-counting robustly?*
- *How to weigh multiple indicators time series?*
- *How to choose the level value? Current descriptions are not clear enough. How to balance prescriptive rules with new information?*
- *Can the same data be interpreted two ways or is this double counting?*
- *Is there value to adding an "unknown" level?*
- *Should we keep risk tables for tier 1-3 stocks only? Non-targets?*
- *Should there be an overall score? If so, how to weigh across categories?*
- *Should PT/SSC provide explicit risk scores and justifications? When?*

Discussion 2: Frameworks for Addressing Uncertainty and Risk

Martin Dorn

(Discussion Lead & Rapporteur)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Genesis of the Tier System

The groundfish tier system is used by the NPFMC to specify the OFL and ABC for stocks and stock complexes in the groundfish fisheries management plans for the Bering Sea and Aleutian Islands and the Gulf of Alaska. The tiers range from one to six and are structured according to the availability of information about the stock, and the ability to reliably estimate management quantities such as maximum sustainable yield (MSY) and stock-recruit relationships (Table 1). Most stocks with age-structured assessments are in tier 3, where the OFL and ABC are calculated using proxies for the MSY fishing mortality rate based on spawning biomass per recruit.

The tier system was developed during the period 1992-1996 when there were many iterations of early versions of the system. There was robust discussion between the plan teams and the SSC about harvest strategies, and controversies over appropriate limit and target fishing mortality reference points, and over appropriate proxies for FMSY and BMSY. Papers by Clark (1991, 1993) were extremely influential in shaping the tier system. The current groundfish tier system in its present form dates from 1999. A structured approach for providing management advice that deals with availability of information and assessments of different types was, at the time, a novel approach, and one that served as a template for national guidance in the development of harvest strategies. In addition, application of the tier system by the NPFMC over the past twenty years has proven successful in maintaining productive fisheries in the North Pacific that are widely recognized both nationally and internationally as examples of sustainably managed fisheries.

The reauthorization of the Magnuson-Stevens Fishery Conservation and Management Act (MSA) in 2006 changed the requirements for how management actions are developed for U.S. fisheries. Councils were required to set annual catch limits (ACLs) for all managed stocks that are "in the fishery." National Standard Guidelines developed to assist in the implementation of the reauthorized act (Federal Register, 2009) defines two sources of uncertainty that must be considered when establishing ACLs: 1) scientific uncertainty, including error pertaining to both the data and to parameter estimation; and 2) management uncertainty, which represents uncertainty in the efficacy of management practices that are designed to ensure that harvest limits are not exceeded (Figure 1).

As stated in the National Standard 1 guidelines, *"The Acceptable biological catch (ABC) is defined as the level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of OFL and any other scientific uncertainty, and should be specified based on the ABC control rule….NMFS believes that determining the level of scientific uncertainty is not a matter of policy and is a technical matter best determined by stock assessment scientists as reviewed by peer review processes and SSCs. Determining the acceptable level of risk of overfishing that results from scientific uncertainty is the policy issue. The SSC must recommend an ABC to the Council after the Council advises the SSC what* *would be the acceptable probability that a catch equal to the ABC would result in overfishing."* (Federal Register, 2009).

Although the National Standard 1 Guidelines adopted the terminology used in the tier system, the ABC and the OFL were given new definitions with a broader understanding of scientific uncertainty than was originally envisioned in the tier system. The new definitions are most consistent tier 1, where the buffer between OFL and the ABC control rules varies directly with the amount of scientific uncertainty. However, in general the primary type of uncertainty addressed in the tier system is the uncertainty in the stock production curve (i.e., the shape of the stock recruit relationship). This is true for tier 1, tier 3, and tier 5, all of which have an OFL/ABC calculation based on FMSY or proxies thereof. Currently, ABCs are calculated using the point estimate of stock size (usually the MLE). For tiers (1-5), if a reliable pdf of B is available, the preferred point estimate is the geometric mean of its pdf, but this use of this provision currently limited to tier 1 stocks. Thus, for most stocks, the estimate of the buffer between ABC and OFL for a given tier does not vary in response to changes in the uncertainty in stock size and/or status. At the time, there was recognition that this was a potential shortcoming, but it was considered preferable to grandfather in tier system under the new MSA requirements, with the understanding that any potential issues could be addressed at some later time.

What additional role do risk tables accomplish?

The NPFMC tier system uses the buffer between the OFL and ABC to implement precautionary management. As the SSC has repeatedly emphasized, the SSC's intent is that the tier system should be regarded as the primary basis for establishing the ABC. The sloping harvest control rule for the ABC will substantially reduce the harvest rate when the stock is at a low abundance, and provide a built-in response to concerns related to low stock abundance. The risk table evaluates whether there is either additional uncertainty in the assessment and/or additional risks (probability of something bad happening) to the stock that are not adequately taken into account by the default precautionary settings. The risks generally relate to a loss of fishery sustainability and an inability of the stock to perform its role in a functioning ecosystem, such as might occur due to severe decline in stock abundance. One challenge to application of the risk table is that the tier system focuses primarily on uncertainty in the productive capacity of the stock, as discussed above. Consequently, there is no clear guidance on how other kinds of scientific uncertainty should be taken into account, or how concerns related to the ecosystem and environment are intended to be dealt with in the management system.

Development and application of the P-star approach to account for scientific uncertainty in setting the ABC (Shertzer et al. 2008)

A P-star (P*) approach has been adopted by other Councils, including the Pacific Fishery Management Council (Ralston et al. 2010) and Mid-Atlantic Fishery Management Council to deal with Magnuson-Stevens Fishery Conservation and Management Act (MSA) requirements to account for scientific uncertainty. A P-star approach has also been used by the NPFMC for crab harvest specification, though SSC routinely reduces the ABC from the values obtained by the P-star approach.

Briefly, the implementation of the P-star approach requires input from both the SSC and Council. SSC adopts or specifies some level of uncertainty (sigma) (usually uncertainty in the OFL, but uncertainty ending biomass is also used). The Council specifies its P-star value, which is the acceptable probability of exceeding the OFL, which needs to be less than 0.5 to be in compliance with National Standard Guidelines. These two assumptions, along with an assumption about the form of a probability density function, usually lognormal, produces a unique result for the buffer between OFL and ABC (Figure 2).

The approach taken by the SSC of Pacific Fishery Management Council was to quantify scientific uncertainty by using variation between repeated assessment or between assessment variation (Figure 3). Ralston et al. (2009) estimated the coefficient of variation (CV) of the among-assessment variation in estimates of historical biomass, based on 81 assessments of 15 groundfish and 2 coastal pelagic stocks. Since there seemed to be similar levels of variability for the different stocks, an overall sigma of 0.36 was applied (with the proviso that if estimated uncertainty in ending year was larger than a CV of 0.36 the actual value would be used)

In an approach similar to the NPFMC tier system, PFMC stocks are grouped according to three stock categories:

Category 1: Data rich, Age/size structured assessment with year-class estimation.

Category 2: Data moderate, Aggregate production model, M*survey biomass, year classes not resolved, or highly uncertain category 1 assessment.

Category 3: Data poor. Average catch assessment.

Since the derivation of sigma used only category 1 stocks, the default sigma was used only for those stocks. For category 2 stock a sigma of 0.72 (i.e., twice the sigma for category 1 stock) was used, while for category 3 stocks, a sigma of 1.44 (i.e., four times the sigma for category 1 stocks) was used. While the approach for category 2 and category 3 stocks is somewhat arbitrary, assessments for those stocks clearly have greater uncertainty than category 1 stocks, so this approach is logically consistent and has the outcome of more precautionary management for stocks whose assessments are regarded as being more uncertain. SSC also informed the Council that any P-star greater than 0.45 as would not be considered a meaningful response to MSA mandate to account for scientific uncertainty in setting the ABC. The Council adopted a P* = 0.45 for all category 1 assessments, and P* = 0.40 for category 2 and 3 assessments.

There have been several recent refinements to the P-star approach used by PFMC. The first refinement was to base sigma on uncertainty in the projected OFL instead of ending year biomass (Privitera-Johnson and Punt 2020). This resulted in a new sigma of 0.50. The second refinement was to account for the number of years since the assessment (Wetzel and Hamel 2019). Assessments for many West Coast stocks are done infrequently, so therefore it was considered important to account for the increased uncertainty as assessments become progressively less indicative of current status. While both of these refinements are clear technical improvements in the treatment of uncertainty, the consequence is the buffer between the OFL and the ABC has become progressively larger given the same Council decision on P-star (Figure 4). As might be expected, this ratcheting effect was not welcomed by the fishery managers on the Council. Furthermore, there are still sources of uncertainty that are not adequately addressed by current stock assessment approaches, so additional increases in the buffer could occur as the scientific community develops techniques for a more comprehensive evaluation of uncertainty.

The SSC of the Mid-Atlantic Council has also used a P-star approach to account for scientific uncertainty and make ABC recommendations. The SSC developed a framework table similar to the risk table in which assessments are evaluated according to nine criteria and are assigned to one of three categories on a spectrum of good to poor assessment performance for each of the nine criteria (Table 2). Each of the three categories has a default CV (or sigma) value associated with it, ranging from 0.6 for the “good” assessments to 1.5 for the “poor” assessments. These CV values were loosely based on simulation results, MSE evaluations, and expert judgement. One notable feature of the framework is that no overall scoring is done. The SSC reaches a consensus on the overall classification of the stock into one of the three categories, which is then used by the Council’s P-star strategy to provide the buffer between the OFL and the ABC.

There are several distinguishing features of these applications of the P-star approach in the Council process. First is that applying the approach requires strong engagement of both the SSC and Council, where SSC provides an approach to characterize assessment uncertainty, and the Council decides on a risk policy by choosing a P-star. This type of decision is foreign to the usual gamut of Council actions, and typically some education and guidance is needed to get the Council up to speed on risk and uncertainty. Another feature is that the approach to characterizing uncertainty tends to be a mix of technical analysis and expert judgement, reflecting both the complexity of the problem, and the incompleteness of current scientific approaches. A final observation is that the SSCs for PMFC and MAFMC (like the NPFMC SSC) are review bodies whose principal role is to review analyses used to support fisheries management decision-making by the Council. However, in both examples presented here, the SSCs took the lead in developing their approach to consider scientific uncertainty in the setting the ABC, which is defined as a critical role of the SSC in reauthorized MSA.

Discussion 3: Quantifying the importance of assessment risk

Grant G. Thompson
(Discussion Lead & Rapporteur)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The SSC Workshop on Risk Tables included separate presentations for each column of the risk table, with a particular focus on *quantifying* the importance of the risks associated with that column. This section of the report summarizes the presentation for the "assessment" column of the risk table.

Enumerating the risk factors

To begin the process of understanding how risks associated with the assessment column are currently quantified, the most recent versions of the risk tables for BSAI groundfish assessments were examined in detail. A total of 23 BSAI groundfish assessments have presented a risk table at least once. Of those, the most recent version for 21 assessments appears in the 2020 SAFE Report, while the most recent version for the other two (Alaska plaice and northern rockfish) appears in the 2019 SAFE report.

Of these 23 tables, the risk "levels" for the assessment column were distributed as follows: two tables assigned a level of 3 (sablefish and blackspotted/rougheye rockfish), five tables assigned a level of 2 (northern rock sole, Pacific ocean perch, northern rockfish, other rockfish, and sharks), and the remaining 16 tables assigned a level of 1. There was almost no correlation between risk level and harvest control rule tier ($\rho = -0.044$), suggesting, perhaps, that authors were basing their assignments of risk level primarily on either: 1) previous versions of the same assessment, or 2) assessments of other stocks or stock complexes within the same tier.

Each determination of risk level is, of course, accompanied by a rationale supplied by the assessment author, and each rationale lists one or more risk factors that contributed to the author's determination. For this exercise, an overall list of risk factors was compiled from the 23 most recent risk tables. Risk factors that appeared to be essentially similar were grouped together. (It should be noted that this involved a degree of subjectivity as to what constituted "essentially similar," meaning that different analysts could easily have arrived at a somewhat different final list of risk factors.) Given that level 1 is defined in the risk table template as being the "normal" level of risk, the risk factors listed by the authors were divided into those that tended to favor a "normal" level of risk (i.e., level = 1), and those that tended to favor a higher level of risk. For ease of reference, the former were termed "positive" risk factors (lower risk) and the latter were termed "negative" risk factors (higher risk).

Twenty-five positive risk factors were identified, while 23 negative risk factors were identified. Table 5 shows the 10 positive risk factors that were listed more than once, and Table 6 shows the 10 negative risk factors that were listed more than once.

Tangible steps toward quantifying risk

The SSC offered the following guidance for the 2018 Groundfish Plan Team workshop that ultimately led to the development of the risk table:

- "The SSC recommends identification of clear and transparent rules for defining the specific criteria to be used when adjusting the recommended ABC." (SSC minutes, February 2018)

In an attempt to satisfy the above recommendation, Thompson (2018) suggested that a multivariate logistic equation could be used to determine an appropriate proportional reduction in ABC:

$$reduction = 2\left(1+\exp\left(-\sum_{var=1}^{nvar} x_{var}\beta_{var}\right)\right)^{-1} - 1\,,$$

where **x** is a vector, each element of which is either 0 or 1, indicating whether the risk factor corresponding to that element applies in a particular assessment; and □ is a vector of non-negative coefficients.

Thompson (2018) tabulated the set of *nfac* (=32) risk factors that had been identified by the respective Groundfish Plan Team for the *nobs* (=76) cases where an ABC reduction was recommended during the period 2003-2017 (a period that predates use of the risk table). For each of the *nobs* observed reductions, he then created a vector **x** by tallying whether each of the *nfac* risk factors was mentioned in the context of that observed reduction, and then combined those into an *nfac* × *nobs* matrix **X**. Finally, he fit **β** to the *nobs* proportional reductions by a constrained least squares approach (i.e., constrained to prevent any element of **β** from becoming negative). The resulting fit gave an R^2 of 0.824, just slightly lower than the value obtained by an unconstrained least squares fit (0.826).

However, further development of this approach was suspended upon recommendation of the SSC:

- "Although it provided a valuable historical perspective, **the SSC recommends not pursuing this analysis further.**" (SSC minutes, October 2018 (emphasis original))
- "Although helpful in developing this process so far, further summary of historical ABC reductions is likely not the best avenue for development of ranges of ABC reduction appropriate for each of the three concern levels." (SSC minutes, December 2018)

Nevertheless, in view of the possibility that the SSC might be open to reconsidering the above recommendations, some initial steps toward updating Thompson's (2018) analysis were undertaken for the purpose of the present exercise. Cursory examination of the risk factors and their relationships to risk levels specified on the basis of the 23 current BSAI groundfish risk tables suggested that the difference between the numbers of positive and negative risk factors might have a substantial amount of explanatory power. The average number of positive risk factors listed in the 23 tables was 2.783, with a range of 0 – 7, while the average number of negative risk factors listed in the 23 tables was 2.261, also with a range of 0 – 7. Of the individual risk factors, the presence of a large retrospective bias appeared to be highly correlated with risk level. A simple linear regression of risk level against the negative-minus-positive difference (x_1) and the presence of a large retrospective bias (x_2) yielded the following model:

$$level = 1.328 + 0.090x_1 + 1.267x_2$$

This model gave an R^2 of 0.721, which might be considered promising for a very simple initial attempt, suggesting that further modeling efforts could be worthwhile.

Internalizing structural uncertainty in the assessment

The SSC has previously noted that there are at least two ways to address structural uncertainty in the assessment, one of which is to use the risk table (referred to as "this tool" in the following excerpt), and another of which is to use ensemble modeling:

- "Reductions from the maximum ABC are made in response to factors not included in the Tier system. Therefore, the most preferable solution to avoid invoking this tool is to find quantitative ways to include these uncertainties in the assessment analyses.... Ensemble modelling may also provide a tool for this task." (SSC minutes, December 2018)

Use of the risk table to account for structural uncertainty in the assessment might proceed according to the following algorithm:

1. Run *n* models.
2. Choose a preferred model.
3. Note that, because the *n* models imply *n* different ABCs, the preferred model does not account for structural uncertainty.
4. Raise the risk score for the assessment category accordingly.
5. After considering all four risk categories, (perhaps) recommend an *ad hoc* reduction from the maxABC implied by the preferred model.

Use of ensemble modeling to internalize structural uncertainty in the assessment might proceed according to the following algorithm:

1. Run *n* models.
2. Choose a set of model weights.
3. Create an ensemble model as the weighted average of the *n* models.
4. Recommend no reduction from the maxABC implied by the ensemble.

Note that both approaches involve subjective elements: specification of risk level in the former, and specification of model weights in the latter. However, once those have been specified, the ensemble approach is completely objective and transparent, but the risk table approach (at least as typically implemented) is neither.

Both approaches also require selection of a set of models to run. However, this selection is more critical in the ensemble approach than in the risk table approach, because only the "best" model has a direct impact on ABC in the risk table approach, but *all* models have a direct impact on ABC in the ensemble approach. Because of this, when using an ensemble approach, special care should be taken to avoid "stacking the deck." Stacking the deck occurs when the ensemble includes multiple models that might be expected to result in ABC values that satisfy some (presumably subconscious) bias on the part of the assessment scientist or other participant(s) in the assessment process. Using a factorial design to create an ensemble can help to avoid stacking the deck. Some possible factors in such a design include the following:

- Data selection, for example:
  - Choice of data sets
  - Choice of data weighting
- Parameterization, for example:
  - Choice of functional forms
  - Choice of fixed parameter values
- Model complexity, for example:

  - Number of free parameters
  - Number of constrained time-varying parameters

Double counting in the assessment category

The SSC has been clear that factors used to determine risk levels should not include those already incorporated into either the assessment model or the harvest control rules:

- "Reductions from the maximum ABC are an infrequent action prompted by extraordinary circumstances, or considerable uncertainty, in an attempt to respond to substantial unquantified risk. Importantly, adjustments from the maximum ABC are based on uncertainty and risk that is not already accounted for in the tier-system approach to reducing the maximum ABC relative to the OFL; these should not overlap." (SSC minutes, October 2018)
- "Recalling the October 2018 report on this topic, the SSC reiterated that reductions from the maximum ABC are intended to be an infrequent action to respond to substantial unquantified risk. Adjustments from the maximum ABC are used to address uncertainty and risk that is not already accounted for via the Tier system and associated harvest control rules." (SSC minutes, December 2018)

Nevertheless, a review of the current risk tables and the associated Groundfish Plan Team discussions suggests that, in practice, some ambiguities remain when it comes to the potential for "double counting" in the assessment column of the risk table. Broadly speaking, two categories of risk factors that continue to be listed in the assessment column of the risk table are:

1. Signals in the data that are being fitted by the model.
   - Example (paraphrased): "The survey biomass data show a downward trend, so the assessment risk level should go up."
2. Uncertainties in the data that are incorporated in the fitting process.
   - Example (paraphrased): "The variances associated with the survey biomass data are large, so the assessment risk level should go up."

With respect to the potential for double counting, the Groundfish Plan Teams often address risk factors such as those listed above by considering whether the information *is already used by the assessment model or harvest control rule when estimating maxABC*. By this criterion, inclusion of either of the above in determining the risk level for the assessment column would appear to constitute double counting.

However, perhaps a more relevant criterion is whether the information *suggests that the ABC should be less than the estimate of maxABC obtained by the model and harvest control rule*. By this criterion, item #1 above would still constitute double counting, but item #2 might not.

Discussion 4: Population Dynamics Risk

Daniel Goethel and Paul Spencer
(Discussion Leads & Rapporteurs)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The population category of the risk table considers whether there is additional risk to the stock based on population dynamics considerations, which are not typically included in stock assessments or the NPFMC harvest control rule. Dorn and Zador (2020) identified some of these considerations, including decreasing biomass trends, poor recent recruitments, inability of the stock to rebuild, abrupt increase or decrease in stock abundance, and other unusual changes in stock age structure or recruitment patterns. Although these guidelines represent an adequate starting point, several issues merit further consideration to improve risk table development: 1) incorporating and extending the population dynamics category to include additional factors (e.g., spatial considerations); 2) addressing risk for data-limited and non-target species; 3) evaluating "double-counting" of risk across risk table categories; and 4) development of quantitative methods to translate risk scores into ABC reductions. Here, we review the cases (i.e., species and circumstances) for which elevated population dynamics risk scores were identified, discuss alternate population dynamics factors that warrant consideration in future risk table development, and highlight broader cross-category risk table issues that require more explicit guidance.

Overall, a total of eight NPFMC managed fish species or species groups were identified as having been assigned elevated population dynamics risk table scores. The factors resulting in heightened risk scores varied widely, but were generally based on the main factors highlighted in Dorn and Zador (2020). For instance, variable and uncertain recruitment, potential changes in growth or condition due to density-dependent effects, reductions or uncertainty in biomass, erosion of age diversity, and low productivity were commonly identified as reasons for heightened scores. Of these eight species, the stock assessment author suggested reductions from the maximum permissible ABC allowed under the NPFMC harvest control rule for only two species, Alaska-wide sablefish and Eastern Bering Sea pollock. Reasons for not suggesting reductions in ABC varied, often depending on the data availability, productivity of the species, and fishery type. For instance, for many data-poor species (e.g., BSAI and GOA sharks, GOA Atka mackerel) authors suggested that ABC reductions were not likely warranted until improved data could be collected or associated stock assessment methods could be improved. In these cases, the uncertainty in the underlying data was sufficiently large to preclude any recommendation for departing from the ABC specified by the harvest control rule. The 2019 Atka Mackerel assessment characterized the population risk as "unknown", and this would be a useful designation for similar situations in the future. GOA Pacific Ocean Perch met some of the considerations specified by Dorn and Zador (i.e., "abrupt increase or decrease in stock abundance"); however, a reduction in ABC was not recommended because survey biomass estimates were at timeseries highs and the assessment was thought to underestimate biomass. For other species where harvest is taken as bycatch, lowering the ABC was viewed as having little impact on catch levels. In the NPFMC system, once the ABC is attained, directed fishing is typically prevented and incidental catch is discarded. Although population risks are recognized in recent BSAI blackspotted and rougheye rockfish and GOA Pacific cod assessments, both of these species were bycatch species (for Pacific cod the abundance was reduced such that no directed fishing was allowed for some recent years), and lowering the ABC would not affect the level of incidental catch. When ABC reductions were suggested, they tended to be dramatic (e.g., 43% for EBS pollock and 57% for sablefish). In both cases, high recruitment variability and ‘peculiarities’ (e.g., gaps or unevenness) in the age structure were cited as reasons to warrant more precautionary management approaches; both factors, but particularly spasmodic recruitment, have been demonstrated to cause issues with traditional harvest control rules leading to overly optimistic ABCs and subsequent population declines (Licandeo et al., 2020).

However, the relatively limited number of species with elevated population dynamics risk scores may be attributed to the moderately limited scope of the factors initially identified for consideration under the population dynamics category. Although downward or atypical trends and uncertainty in biomass, abundance, or recruitment are important indicators of population health that merit careful monitoring, a wide variety of alternate factors should also be considered. For instance, expansion of the population risk category to include spatial considerations is warranted. For example, the stock may expand or contract their range or change their level of spatial aggregation, which could affect their exposure to fishery effort and/or environmental conditions. For many stocks, changes in spatial distributions should be evaluated by age due to ontogenetic patterns in habitat use. Indicators of stock status are based on spatially-aggregated indices, and may mask disproportionate patterns in sub-area exploitation rates and localized depletion. These types of spatial changes in fishery targeting and availability, or stock distribution, could potentially occur abruptly, leading to interannual changes in risk evaluation. Finally, stock boundaries that are inconsistent with the spatial stock structure could increase the risk of local depletion or overfishing, and would not be indicated by the current risk table. Similarly, the risk category should also be expanded to include some useful metrics of population health based on age truncation. For example, diversity of age structure is an important indicator of population health, which can be severely altered by fishing pressure (Barnett et al., 2017), as it is undesirable for relatively long-lived stocks to have the bulk of their population concentrated into a small number of age classes (Spencer et al., 2014). Simple metrics such as the Shannon-Weiner (Shannon, 1948) index can be easily calculated and can be compared over time. Interpretation of such metrics would be enhanced by a fuller understanding of the importance to age-structure diversity for particular stocks or life-history patterns, including information on reproductive biology, recruitment, and portfolio effects.

Additionally, data limited and/or non-target species present a conundrum for the general use of the risk table approach, which is especially apparent within the population dynamics category. As noted, many stock assessment authors have suggested that reductions in ABCs are not necessarily warranted based on increased population dynamics risk scores, because the ABC reduction would either not effectively limit the fishery (i.e., for bycatch or incidental catch species) or not enough information was available to make a well-informed decision regarding population dynamics (e.g., biomass trends for data limited species). For extremely data-limited species (e.g., NPFMC tier 6 species), the risk table approach may not be warranted or informative, because there simply is not enough information to score the various risk categories. However, for data limited species for which basic population trend (e.g., survey abundance) information is available, then a basic risk table can likely be useful, but scoring may need to be refined based on more coarse knowledge (e.g., basing population dynamics scores on biomass trends and general understanding of productivity levels rather than stock assessment based estimates of recruitment). For non-target species, novel application of risk table scores may be warranted outside of the traditional adjustment to the ABC. In addition to the potential ineffectiveness of an ABC reduction for limiting incidental bycatch, for some stocks the realized catches are substantially lower than the ABC, which would also limit the utility of lowering the ABC. There is no clear consensus on how risk table scores might improve management of bycatch and/or data-poor stocks, but potential options could include increasing implementation of spatiotemporal area or fishery closures to reduce bycatch. However, it may still be useful to suggest lower ABCs in the case of bycatch or incidentally caught species with elevated population dynamics risk scores, because it may increase awareness regarding the potential negative impacts of comparatively higher ABCs on these species even though reduced ABCs may not have any tangible impact on the realized catch.

The risk table approach was implemented by the NPFMC to identify factors that may increase the risk of overfishing, which are not directly accounted for in the stock assessment or harvest control rule. As a relatively new approach to aid management decision making, there remain a number of issues that have yet to be fully resolved. Although not necessarily unique to the population dynamics category, the lack of a quantitative framework to translate risk table scores into ABC reductions and the potential impact of double counting across risk table categories represent important factors that often influence stock assessment authors' scoring of the population dynamics category. By necessity, risk table scoring is currently subjective and varies across assessment authors as well as across review and management bodies (e.g., the Plan Teams, SSC, and Council). As a qualitative tool to identify areas of increased concern for a species, risk tables are extremely valuable. However, without a general framework to translate risk table scores into quantified ABC reductions and/or a methodology to objectively rank category scores across species, it becomes increasingly hard for authors to justify a given set of risk table scores in relation to a suggested ABC reduction. The issue of ‘double counting’ factors of concern (i.e., accounting for a single issue under separate risk table categories) can also be viewed as both a positive and negative of the risk table approach. Much of the information on population-level attributes is obtained from assessment models, which have some degree of estimation error and model misspecification; thus, it can be difficult to tease apart population risk from assessment risk. Addressing the same issue under multiple categories can lead to inflated scores across categories and might lead to a larger ABC reduction than may be necessary. Conversely, from a qualitative perspective, emphasizing important issues across categories can help highlight the most important factors that may be detrimentally impacting the population. Although a completely prescriptive approach to risk table scoring is not necessary or warranted, more complete guidance (i.e., on how to deal with cross-category factors and how to determine ABC reductions based on risk table scores) could help improve consistency across species, as well as, within assessments as inevitable stock assessment author turnover occurs.

Identifying factors that may increase the risk of overfishing is necessarily an iterative process, which is refined as new data is collected and knowledge synthesized. Necessarily, the risk table process must also be iterative and continually refined. In the case of the population dynamics category, we suggest that stock assessment authors should expand the factors considered to include spatiotemporal dynamics including expansion, contraction, and/or localized depletion, while also better emphasizing the importance of age diversity for healthy populations. More generally, expanding the risk table approach to include alternate management responses, aside from ABC reductions, when elevated scores are given is also warranted in the case of non-target species for which ABCs rarely limit harvest. Similarly, methods are needed to aid in objectively assigning ABC reductions when heightened risk table scores exist. For instance, a more formal approach to implementing alternate projections to account for demographic or recruitment uncertainty could guide bounds on ABC reductions (e.g., in the case of highly variable recruitment, an average recruitment projection could provide guidance on upper or lower bounds on ABC). As the risk table approach matures in the coming years, continued guidance to ensure consistency and objectivity within and across categories, species, authors, and management bodies will be helpful.

Discussion 5: tangible steps towards quantifying risk of external changes in ecosystem conditions

Kalei Shotwell, Bridget Ferriss, Elizabeth Siddon, Stephani Zador
(Discussion Leads & Rapporteurs)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Introduction

The ecosystem conditions category of the risk table considers whether there is additional risk to the stock based on environmental or ecosystem conditions that are not included in the main stock assessment. These conditions may include adverse trends in predators, prey, competitors, and environmental processes as reflected in environmental/ecosystem indicators, ecosystem model results, and empirical observations (e.g., survey or satellite data). Dorn and Zador (2020) state that the environmental/ecosystem considerations will usually be based on indicators that track environmental or ecosystem properties that are regarded as important to the stock because of a plausible ecological connection. They also suggest that the indicators could be species-specific or ecosystem-wide, could include direct forcing variables that have been linked to the population dynamics of the stock, or indirect indicators that inform a population process.

The ESR and the ESP reports are the two main sources of information to fill out the ecosystem category of the risk tables. The Alaska groundfish stock assessment authors first started completing the risk tables in 2018 and at that time, only five stocks conducted the risk tables. At that point, only the sablefish assessment had a completed ESP to reference. The other stock assessment authors consulted with the ESR editors regarding ecosystem conditions to fill out the risk tables. In 2019, the GOA pollock ESP was added and at least one of the ESR editors was assigned as a Point of Contact for each full stock assessment that was slated to complete a risk table. Finally, in 2020, the GOA and EBS Pacific cod ESPs were completed and the ESR editors created a standardized template (Figure 1) to help with completing the risk tables. This template consisted of four main categories (predators, prey, competitors, and environmental processes) to help organize the ecosystem-level data from the ESR and describe adverse trends in the four categories. Additionally, ESP teams now include at least one ESR representative. This helps to avoid redundancy in the contributions of the ESRs and ESPs to the risk tables for each stock.

Review of Current Risk Table Ecosystem Levels

Over the three years of conducting risk tables at the NPFMC, sixteen risk tables from nine stocks were assigned an elevated ecosystem risk table score. The elevated score was to level 2 in all cases and did not always result in a reduction. The stocks with elevated ecosystem levels and an associated reduction were usually data-rich stocks with several species-specific indicators developed for reference. Generally, there was also another category elevated (typically the population dynamics category) for stocks where an ABC reduction was recommended. This occurred in nine risk tables from four stocks (Alaska sablefish, EBS pollock, GOA pollock, and GOA Pacific cod). The exception to this was EBS Pacific cod in 2019 where only the ecosystem category was elevated, and the SSC adopted the ensemble model as a means to reduce from max ABC. Three stocks had elevated ecosystem levels that did not result in an ABC reduction (BSAI Greenland turbot, BSAI northern rockfish, and GOA arrowtooth flounder). The elevated level was related to uncertain recruitment, shifts in the cold pool extent, poor condition of the fish, and lack of forage due to heatwave conditions. Two stocks (GOA pollock and GOA dusky rockfish) had recommended ABC reductions but did not have elevated ecosystem scores, citing that the indicators were mixed but conditions were better than the previous year or that there was limited survey information (e.g., due to COVID-19). The majority of risk tables (37 of 53 or 70%) did not have elevated ecosystem levels and provided descriptions of the four categories in the template with a mix of signals.

Suggestions on New Approaches for the Ecosystem Category

The risk tables have evolved since their inception at the NPFMC in 2017; information that is now included in the ecosystem category has benefitted from two major elements, consistency and coordination. The four category template was instrumental in ensuring that the main pressures on a stock were all considered and thus enabled consistency among stocks in their ecosystem sections. The coordination among the stock assessment authors, the ESR editors, and the ESP teams has allowed for a more refined stock-specific selection of indicators for evaluation in the risk table and increased collaboration between scientists of different disciplines. In addition to the ESRs and ESPs, there are perhaps several other sources of information that could inform the ecosystem category. For instance, output from more complex multi-species or ecosystem models (e.g., closed life cycle Individual Based Models, CEATTLE, ecopath with ecosim, Atlantis, FEAST) may be helpful for identifying relevant predator/prey relationships and point to additional indicators to monitor for a given stock. Ocean modeling is now becoming more operational (e.g., ROMS, NPZ) and when combined with observations from ecosystem process studies, could increase in skill level to allow for indicator development at multiple temporal and spatial scales. This could allow further exploration of indicator importance methods and ecosystem research models that may inform the risk tables on a more quantitative level. These models could provide estimates of the direction and magnitude of the effect on the stock as well as an estimate of the unaccounted for uncertainty in the main stock assessment model. The ESP process is also evolving concurrently with the risk tables. The guideline criteria for indicator selection for an ESP and subsequent potential use within the research ecosystem model could very well be used to weight indicators for use in the risk table. This would assist in stabilizing the subjectiveness of author scoring and potentially the level of reduction if warranted.

Issues to Address in Future Risk Tables

As with the other risk table categories, there remain several unresolved issues on how to create and score the ecosystem category in the risk table for a given stock. Double counting remains an issue. As mentioned previously, multiple risk table categories may reference a single indicator (e.g., the marine heatwave) and this may cause concern that the multiple instances may artificially elevate the importance of the indicator. However, the standardized ESR framework (predator/prey/competitor/environment) should help to distinguish the different pressures of an indicator on the stock (e.g., heatwave decreases availability of prey X and increases predator Y). Consistent and disciplined use of this framework will be essential to avoiding the pitfall of double counting if we move from a qualitative to a more quantitative risk table in the future. To that end, it is subjective and difficult to determine relative importance qualitatively when there are multiple indicators within the ecosystem category. An option may be to include subject matter experts to assist with interpreting the influence of different ecosystem indicators and providing mechanistic linkages to the stock of interest. When an ESP is developed for a given stock, this would be within the scope of the associated ESP team. However, it is unclear how to accomplish this for stocks without an ESP (of which there may be many). One possible solution may be to create ESP guild teams (e.g., flatfish, rockfish) to assist with risk tables of stocks with similar life histories and ecological niches and begin development of ESPs for priority stocks. A related issue to double counting is the instance when the same mix of indicators is used to describe both level 1 and level 2 ecosystem concerns. The ESP guild approach may also help to identify stock-specific life history vulnerabilities that would cause a set of indicators to influence one stock more than another within a guild. Finally, there seems to be general agreement that the Plan Teams and SSC should not provide explicit risk scores and justifications; however, it is unclear how to record the rationale when the Plan Teams and SSC disagree with an author recommended risk table score or ABC reduction. These issues are all somewhat developed in each of the risk table categories and further guidance from the Plan Teams and SSC would be very helpful.

Discussion 6: tangible steps towards quantifying the importance of external changes in fishery performance in stock assessments

Alan Haynie, Jim Ianelli, and Stephen Kasperski
(Discussion Leads & Rapporteurs)

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Our main goals in this session were to 1) discuss fishery performance measures, 2) describe situations where they may provide valuable out-of-model insights into stock health, 3) identify research needed to better identify the relationship between performance metrics and stock health, and 4) note cases where fishery performance metrics may be relevant for the bycatch stocks rather than the target fishery stock. To this end, we constructed a pre-workshop survey on these topics. This summary reports the results from this survey and the discussion held during the workshop.

Are there mechanistic linkages to stock health that are revealed by fishery performance data?

For example, are there fishery performance indicators that differ substantially from trends or conditions indicated by the assessment? One case of note was for GOA Pacific cod where fishing performance (specifically catch rates and fishery participation) dramatically declined prior to when the impacts of the marine heatwave were identified in the survey and assessment. Fish condition (skinniness given length within the fishery) was also suggested as a metric to consider. Consistency between the stock characteristics (age structure, etc.) as estimated within the assessment could be revealed by changes in product mix. For example, more large fish in the catch could reflect population structure or increased price premium for larger fish and fleet targeting in response to that price premium. The group cautioned over simple interpretations of CPUE and general effort measures.

In most fisheries, environmental conditions and management measures significantly affect fishery timing and other behaviors and may not reflect changes in the health of the stock. It is essential that stock assessment scientists consider non-stock factors that may impact fishery-dependent metrics such as CPUE, selectivity, etc., even if only presented qualitatively. Additionally, we noted that fishery data may reflect species distribution which in turn could affect how survey data should be interpreted.

Recognizing that there are data lags between the provision of management advice and the availability of fishery data, what are best practices for integrating fishery indicators with advice derived from assessment model results?

As an example, the group discussed how it appeared that the 2020 EBS pollock fishery likely caught/selected younger fish than expected based on previous years. This affected the choice of what age-specific selectivity was used for advice for the 2021 fishery. Other cases discussed were the impacts of having intermittent surveys (e.g., GOA Pacific cod when conditions were changing quickly and the management reaction could potentially have reacted sooner).

How do we define "fishery" performance risk for bycatch stocks, and how would it be affected by changes in bycatch and incidental catch?

Examples discussed included sablefish bycatch in the Bering Sea pollock fishery evaluated in the sablefish assessment. Target fishery impacts on other stocks are accounted for based on observer data collections and in-season management measures. However, should changes become apparent, impacts on specific stocks (e.g., based on Chinook salmon stock identification work) relaying such information to fishers could minimize the impact on more compromised stocks. A general discussion point was that the presence of another species bycatch may be a consideration to reduce TAC of the target species (as opposed to reducing the ABC of the target species).

Discussion 7: Frameworks for addressing scientific uncertainty: Comparing and contrasting the P* and decision-theoretic approaches

Grant G. Thompson

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Overview of the approaches

In the context of managing the BSAI and GOA groundfish fisheries, two frameworks for addressing scientific uncertainty have been discussed for a number of years: the $P*$ approach and the decision-theoretic (DT) approach.

The approach (e.g., Prager et al. 2003) consists of the following steps:

- Set a value of $P*$ between 0 and 0.5.
- Compute the cumulative distribution function of the true-but-unknown value of the overfishing level, $CDF(truOFL)$.
- Set $ABC = CDF^{-1}(P*)$.

The DT approach (e.g., Thompson 1992, 1996, and 1999; also section 3.1 of Restrepo et al. 1998) can be viewed as either a maximization problem or a minimization problem. When viewed as a maximization problem, the approach consists of the following steps (the equivalent minimization problem corresponds to the terms shown in parentheses):

- Define a utility (loss) function specifying the desirability (undesirability) of each possible outcome; for example, long-term yields.
- Weight the utility (loss) of each relevant outcome by the probability or probability density of that outcome, then sum or integrate to get the expected utility (loss).
- Fish at the rate that maximizes (minimizes) expected utility (loss).

Advantages and disadvantages of the approaches

Advantages of the $P*$ approach include the following:

- Sounds like hypothesis testing, so is a natural choice for advocates of hypothesis testing.
- Much more widely known than the DT approach.
    - In fact, until the most recent revision of the NS1 guidelines, it was the only approach officially allowed.
- Computationally simpler than the DT approach (integrating a function versus maximizing the integral of a product).

- Resulting ABC is always less than OFL.

Advantages of the DT approach include the following:

- Rooted in Bayesian theory, so is a natural choice for advocates of Bayesian methods.
- Considers all relevant outcomes.
- Provides an estimate of the optimal catch.

Disadvantages of the P^* approach include the following:

- Considers only one possible outcome ($ABC > truOFL$), regardless of the amount of overestimating or underestimating.
- Does not provide an estimate of the optimal catch.
- As with a value in hypothesis testing, difficult to justify P^* value.
- Choice of model/data can have major impacts on the form of the CDF.

Disadvantages of the DT approach include the following:

- Computationally more complicated than P^* approach.
- Requires specifying a utility (loss) function.
    - However, unlike the value of P^*, the utility (loss) function can be estimated from experimental data.
- In some (perhaps rare?) situations, can result in ABC > OFL.
- Choice of model/data can have major impacts on the form of the PDF.

Possible hybrid approaches

The problem of choosing an approach does not have to be an either/or situation, as some hybrid options are also possible:

- Choose the DT approach unless the resulting ABC exceeds the OFL, in which case default to the P^* approach.
- Choose the minimum of the ABCs resulting from the two approaches.
- Use P^* approach for ABC, DT for a TAC option (if less than ABC).

Current state of the discussion

During development of Amendments 96/87 (implemented November 2010), it became apparent that some issues related to the treatment of ACLs in the National Standard Guidelines were too complicated to address fully. Trailing amendments were anticipated for some issues, such as the buffer between ABC and OFL. A discussion paper was therefore developed in the spring of 2011. This was reviewed by the SSC in June 2011 and by the Joint Groundfish Plan Teams in September 2013, with follow-up comments provided by the SSC in October 2013. The recommendations from the Teams and SSC were as follow:

- SSC recommendations (June 2011)
    - "The SSC recommends a deliberative approach to improving the treatment of uncertainty in the groundfish FMPs and encourages the author and/or other analysts to further develop the document to:

        - "explore the advantages and disadvantages of the DT and P^* approaches using more realistic scenarios, and
        - "determine how the approaches would be applied across different tiers...."
    - "This will require continued research on developing appropriate models for understanding the interactions between fisheries in response to changes in harvest policy."

- Joint Team recommendations (September 2013)
    - "The Teams did not recommend a preferred alternative for this issue, but did recommend that any future analysis of the DT approach [should] consider a variety of utility functions.
    - "It was noted that AFSC economist Chang Sueng has done some work in this regard.
    - "Furthermore, the Teams recommended that analysis of all options should evaluate risk for a range of years and species."
- SSC recommendations (October 2013)
    - "In their September 2013 meeting, the Joint Plan Teams provided new advice..., which the SSC supports."
    - "The SSC encourages further development of these analyses over a reasonable time frame."

Discussion 8: Frameworks for addressing scientific uncertainty: A joint probability approach for linking the risk table to ABC reductions

Grant G. Thompson

Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

Overview

An approach for implementing the risk table, termed the “joint probability” approach, is described here. The joint probability approach augments the current approach in a way that ties the risk table directly to:

- The need for a reduction from *maxABC*.
- The appropriate amount of reduction (if any).

The joint probability approach is completely consistent with the current features of the risk table. Although the number of columns, or risk “categories” (*ncat*), and the number of rows, or risk “levels” (*nlev*), is entirely flexible, for the example considered here *ncat* and *nlev* will both be set equal to 4, corresponding to the structure of the current risk table:

- *ncat*=4 categories (assessment, population dynamics, environmental/ecosystem, fishery performance).
- *nlev*=4 levels (1, 2, 3, 4), with definitions as currently given in the risk table template.

From a broad, overview perspective, the steps involved in the current and joint probability approaches to implementing the risk table can be summarized as described below.

For the current approach:

1. The assessment author uses a set of subjective methods to arrive at levels for the categories in the risk table.
2. The assessment author uses a *second* set of subjective methods to determine whether an ABC reduction is necessary.
3. If step 2 results in an affirmative determination, the assessment author uses a *third* set of subjective methods to determine the size of the reduction.

For the joint probability approach:

1. The assessment author uses a set of subjective methods to arrive at “scores” for the categories in the risk table (similar or identical to the current approach; see below).
2. The need for an ABC reduction is determined statistically.

3. If step 2 results in an affirmative determination, the size of the reduction is determined statistically.

A quantifiable interpretation of "concern," with an example

The language used in the risk table template suggests that the currency of the risk table is "concern," but this term is left undefined. In the joint probability approach, "concern" is interpreted in terms of the probability that *maxABC* exceeds the true-but-unknown overfishing level (*truOFL*, as distinguished from the overfishing level specified on the basis of the assessment model point estimate, *OFL*). In the joint probability approach, an ABC reduction is necessary if the probability of *maxABC* being greater than *truOFL* exceeds 50%.

Two types of probability need to be considered:

- Probabilities of overfishing that are internal to the model.
  - These are routinely quantified.
- Probabilities of overfishing that are external to the model.
  - These are associated with the factors identified under the categories in the risk table, and are *not* routinely quantified.

Figure 5 provides an example of the cumulative probability of exceeding *truOFL*. The cumulative probability (i.e., the cumulative distribution function of *truOFL*, *CDF*(*truOFL*)) that is internal to the model is shown by the blue curve (this example is based on a lognormal distribution with μ = ln(100,000) and $\sigma = 0.2$). The point estimate of *OFL* obtained from the assessment model is 100,000 t (solid red line), corresponding to a cumulative probability of 50% (dashed red line). The point estimate of *maxABC* obtained from the assessment model is about 90,000 (solid green line), corresponding to a cumulative probability of about 30%. Thus, if ABC were set equal to *maxABC*, the "internal" probability of overfishing would be about 30%. The question addressed by the joint probability approach is whether consideration of the external factors would be expected to bridge the gap between the 30% probability associated with *maxABC* and the 50% probability threshold.

Viewing the problem in terms of joint probabilities

The joint probability of overfishing, $Pjnt$, can be viewed in terms of the internal probability of overfishing, $Pint$, and the external probability of overfishing, $Pext$, as follows:

$$Pjnt = 1 - (1 - Pint)(1 - Pext)\,.$$

Because there are $ncat$ categories in the risk table, $Pext$ itself is also a joint probability, and depends on the probabilities associated with the $ncat$ individual categories, $Pext.ind_j$, as follows:

$$Pext = 1 - \prod_{j=1}^{ncat} \left(1 - Pext.ind_j\right).$$

From this perspective, the past practice of ignoring all categories other than the one with the highest level is incorrect, as it implicitly re-sets each of the other $Pext.ind_j$ to a value of 0.

What is needed is a way to move from the information already contained in the risk table categories to a set of probabilities. Both the current and joint probability approaches begin by requiring authors to specify a value of $level$ for each category. The joint probability approach expands on this by allowing authors to specify an (optional) *intralevel* value for each category, with a range of 0 to 1.

A continuous $score$ is then defined for each category j as:

$$score_j = \frac{(level_j - 1 + intralevel_j)}{nlev},$$

with a range of 0 to 1.

If an author prefers not to specify an *intralevel* value for each category, a default value (e.g., 0 or 0.5) could be assumed instead.

The next step is to convert each $score$ into an individual external probability as

$$Pext.ind_j = Pmax \cdot score_j{}^{\alpha},$$

where $Pmax = 1 - 2^{-\frac{1}{ncat}}$ and α is a parameter (choosing a value for α is addressed below).

The coefficient $Pmax$ is needed in order to:

- Keep the external probability of overfishing from expanding in the event that more categories are added in the future.
- Keep the ABC associated with $Pjnt = 0.5$ positive.

Given the above, only two more steps are necessary. The first is to solve for $Pabc$, which is the value of $Pint$ that sets $Pjnt = 0.5$, viz.:

$$Pabc = \frac{(1 - 2Pext)}{\big(2(1 - Pext)\big)}.$$

Finally, ABC is set as follows:

- If $Pabc \geq Pint$, then set $ABC = maxABC$.
- If $Pabc < Pint$, then set $ABC = CDF^{-1}(Pabc)$.

Finishing the example

For the example illustrated in Figure 5, recall that the internal probability of exceeding the true-but-unknown OFL was given by the model at a value of about 30% (more precisely, 29.9%). Suppose that *level*$_j$ and *intralevel*$_j$ were set by the assessment author at values of 2 and 0.5, respectively, for all j; and that the value of α was set by the SSC or Council at a value of 0.2. Given these quantities, the determination of whether an ABC reduction is appropriate, and the size of the reduction in the event of an affirmative determination, proceeds formulaically as follows:

- $score_j = \frac{(2+0.5-1)}{nlev} = 0.375$ for all j
- $Pext.ind_j = \left(1 - 2^{-\frac{1}{ncat}}\right) 0.375^{0.2} = 0.131$ for all j
- $Pext = 1 - \prod_{j=1}^{ncat} \quad (1 - 0.131) = 0.429$
- $Pjnt = 1 - (1 - 0.299)(1 - 0.429) = 0.600 > 0.5$
- $Pabc = \frac{(1-2\cdot 0.429)}{(2(1-0.429))} = 0.124 < Pint$

- $ABC = CDF^{-1}(0.124) = 79{,}400$ (a 12% reduction from $maxABC$)

The entire set of computations can be done in an Excel spreadsheet of only 10 kb in size.

Figure 6 adds the above results to Figure 5. The blue curve and the red and green lines are the same as those shown in Figure 5. The dashed magenta line corresponds to $Pjnt$ (=0.600). Although not listed in the above set of calculations, the dashed magenta line crosses the internal cumulative distribution function at a value that might be termed the "effective" $maxABC$ (approximately 105,000), indicated by the solid magenta line. The dashed purple line corresponds to $Pabc$ (=0.124), which crosses the internal cumulative distribution function at the final ABC value of about 79,400 (solid purple line).

Choosing a value of α

As noted above, the joint probability approach requires a method for converting the value of each $score_j$ into a probability. Although many other methods can be imagined, the method for doing so suggested here involves raising each $score_j$ to a power α. Presumably, the value of α would be set by the SSC or Council. Setting the value of α should involve an understanding of how it relates to quantities such as $Pabc$ and the reduction (if any) from $maxABC$, which will also require information about the range of $Pint$ values suggested by the assessment models. One way to simplify the task is to assume that the value of $score_j$ is constant across j (although information about the range of $Pint$ values suggested by the assessment models will still ultimately be necessary).

Figure 7 shows $Pabc$ as a function of $score_j$ for various values of α in the special case where $score_j$ is the same for all categories. The dashed black lines demarcate the four values of *level* in the risk table. Recalling that an ABC reduction is necessary only if $Pint > Pabc$, Figure 7 can be used to gain some understanding of the likely frequency of ABC reductions under different values of α. For example, suppose that an assessment author were to set $level_j$=1 and $intralevel_j$=0.5 for all j, thus giving $score_j$=0.125 for all j. Then, if the Council or SSC were to set α at a value of 0.2, an ABC reduction would be required whenever $Pint$ exceeded a value of about 0.221. At the other extreme, if the Council or SSC were to set α at a value of 1.0, an ABC reduction would not be required unless $Pint$ exceeded a value of about 0.458.

Figure 8 shows the reduction from $maxABC$ as a function of $score_j$ for various combinations of $Pint$ and lognormal σ in the special case where α=0.2 (as in Figure 2) and where $score_j$ is the same for all categories. If the range of σ values shown here (0.3 to 0.6) is sufficiently broad, then, if $score_j$=0.125 (the midpoint of level 1), reductions will fall within a range of about 0.143−0.266 if $Pint$=0.4, and within a range of about 0.071−0.136 if $Pint$=0.3; but no reductions will be necessary if $Pint$=0.2 or $Pint$=0.1. If $score_j$=0.875 (the midpoint of level 4), reductions will fall within a range of about 0.208−0.662 for all parameter combinations shown.

The α value of 0.2 used to develop Figure 6 was estimated by the method of least squares, using as data the levels specified in all risk tables completed to date and the associated reductions (including reductions of zero), setting $intralevel_j$=0 for all j in all assessments, and calculating the lognormal σ for each $truOFL$ distribution by assuming that the OFL and $maxABC$ values from each of these assessments corresponded to the geometric and harmonic means of the distribution, respectively.

Concluding thoughts: potential non-independence of events

Note that the equations for computing joint probability listed above assume that the events are independent. This may not be entirely accurate, but it should at least be a reasonable starting point. Based on the most recent risk tables for assessments of BSAI groundfish, the specified risk levels for the various categories tend to be positively correlated in practice (whether they should be positively correlated *in principle* may be another matter). If the risk levels are positively correlated, then the value of $Pext$ obtained by assuming that the $Pext.ind_j$ are independent will be biased upward, meaning that the value of $Pjnt$ will likewise be biased upward. Given the inherent subjectivity of the *level* determinations that are necessary in both the current and joint probability approaches, such bias could reasonably be assumed to be of little concern, comparatively speaking. If, however, such bias is nevertheless deemed a serious concern, the approach could be modified by treating the specified reduction (if any) as an *upper bound* on the appropriate reduction rather than the final value. Although this would result in a situation somewhat similar to the status quo, in which the assessment author (or Groundfish Plan Team or SSC) would be left with the problem of how to set the appropriate reduction within the resulting range, at least the range would be bounded far more reasonably than at present, in which the admissible range is essentially 0% to 100%.

References

Barnett, L.A.K., T.A. Branch, R.A. Ranasinghe, and T.E. Essington. 2017. Old-growth fish become scarce under fishing. Current Biology. 27: 2843-2848.

Clark, W. G. 1991. Groundfish Exploitation Rates Based on Life History Parameters. Canadian Journal of Fisheries and Aquatic Sciences 48: 734-750.

Clark, W. G. 1993. The Effect of Recruitment Variability on the Choice of a Target Level of Spawning Biomass Per Recruit." pp. 233-246. In Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. G. Kruse, D. M. Eggers, R. J. Marasco, C. Pautzke and T. Quinn II. (eds). Alaska Sea Grant College. Program, P. O. Box 755040, Fairbanks AK.

DiCosimo, J., Method, R.D., and O.A, Ormseth, O.A., 2010. Use of annual catch limits to avoid stock depletion in the Bering Sea and Aleutian Islands management area (Northeast Pacific). ICES Journal of Marine Science 67:1861-1865.

Dorn, M.W. and S.G. Zador. 2020. A risk table to address concerns external to stock assessments when developing fisheries harvest recommendations. Ecosystem Health and sustainability. 6(1):1-11.

Federal Register. 2009. Magnuson-Stevens Act Provisions; Annual Catch Limits; National Standard Guidelines; final rule. vol. 74, no. 11, January 16, p. 3178–3213. GPO, Washington, D.C.

Goodman, D., Mangel, M., Parkes, G., Quinn, T., Restrepo, V., Smith, T., and Stokes, K. 2002. Scientific Review of the Harvest Strategy Currently Used in the BSAI and GOA Groundfish Fisheries Management Plans. http://www.alaskafisheries.noaa.gov/npfmc/misc_pub/f40review1102.pdf, 138 pp.

Holsman, K., J. Samhouri, G. Cook, E. Hazen, E. Olsen, M. Dillard, S. Kasperski, et al. 2017. An Ecosystem-based Approach to Marine Risk Assessment. Ecosystem Health and Sustainability 3 (1): e01256. doi:10.1002/ehs2.1256.

Licandeo, R., D.E. Duplisea, C. Senary, J.R. Marentette, and M.K. McAllister. 2020. Management strategies for spasmodic stocks: a Canadian Atlantic redfish fishery case study. Canadian Journal of Fisheries and Aquatic Sciences, 77: 684-702.

Prager, M. H., C. E. Porch, K. W. Shertzer, and J. F. Caddy. 2003. Targets and limits for management of fisheries: a simple probability-based approach. *North American Journal of Fisheries Management* 23:349-361. https://doi.org/10.1577/1548-8675(2003)023<0349:TALFMO>2.0.CO;2

Privitera-Johnson, K. M. and Punt, A. E. Leveraging scientific uncertainty in fisheries management for estimating among-assessment variation in overfishing limits. ICES Journal of Marine Science, 77: 515–526.

Ralston, S., Punt, A. E., Hamel, O. S., Devore, J. D., and Conser, R. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fishery Bulletin, 109: 217–231.

Restrepo, V. R., G. G. Thompson, P. M. Mace, W. L. Gabriel, L. L. Low, A. D. MacCall, R. D. Methot, J. E. Powers, B. L. Taylor, P. R. Wade, and J. F. Witzig. 1998. *Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act*. NOAA Tech. Memo. NMFS-F/SPO-31. National Marine Fisheries Service, NOAA. 1315 East-West Highway, Silver Spring, MD 20910. 54 p.

Shannon, C.E. 1948. A mathematical theory of communication. Bell System Technical Journal, 27, 379–423.

Shertzer, K. W., M. H. Prager, and E. H. Williams. 2008. A probability-based approach to setting annual catch levels. Fish. Bull. 106:225–232.

Shotwell, S.K. 2018. Update on the Ecosystem and Socioeconomic Profile (ESP). Unpublished Report to Joint Groundfish and Crab Plan Teams, September 2018. 11 p. Available online at: https://meetings.npfmc.org/CommentReview/DownloadFile?p=d467ccc4-a136-4d63-b445-fdc6d0fe4629.pdf&fileName=ESP_Update_PT-0918_Shotwell.pdf

Shotwell, S.K. 2020. Update on the Ecosystem and Socio-economic Profile (ESP) in the Alaska groundfish and crab fishery management plans. Unpublished Report to Joint Groundfish, September 2020. 18 p. Available online at: https://meetings.npfmc.org/CommentReview/DownloadFile?p=8f5233fb-3b62-4571-9b49-8bb7ce675916.pdf&fileName=ESP_Shotwell.pdf

Shotwell, S.K., 2021. Collated minutes on risk tables from the NPFMC Groundfish Plan Teams, Scientific and Statistical Committee, Advisory Panel, and Council. Unpublished Report for SSC Risk Table Workshop, February 2021. 39 p.

Spencer, P.D., S.B.M. Kraak, and E.A. Trippel. 2014. The influence of maternal effects in larval survival on fishery reference points for two life-history patterns. Canadian Journal of Fisheries and Aquatic Sciences. 71:151-161.

Thompson, G. G. 1992. A Bayesian approach to management advice when stock-recruitment parameters are uncertain. *Fish. Bull.* 90:561-573.

Thompson, G. G. 1996. Application of the Kalman Filter to a stochastic differential equation model of population dynamics. *In* D. J. Fletcher, L. Kavalieris, and B. F. J. Manly (eds.), *Statistics in Ecology and Environmental Monitoring 2: Decision making and risk assessment in biology*, p. 181-203. Otago Conference Series 6, University of Otago Press, Dunedin, New Zealand.

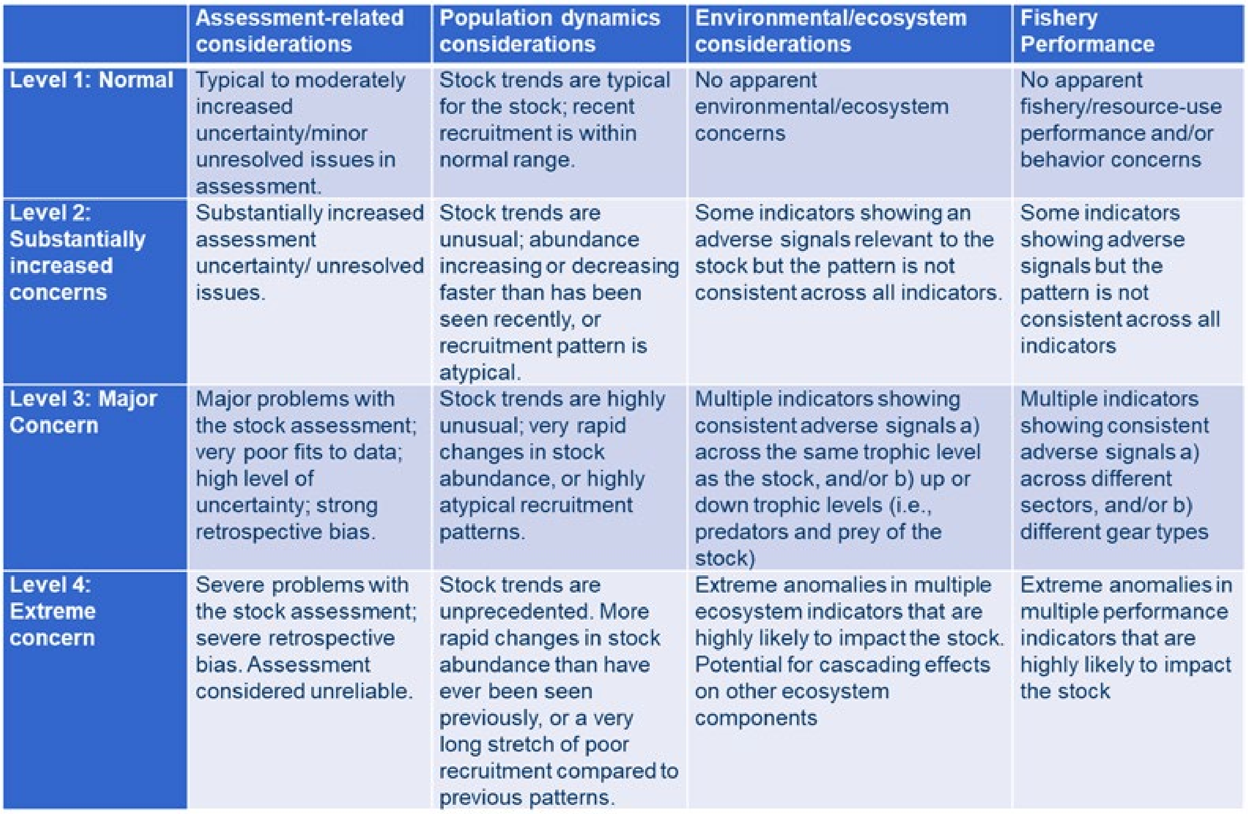
Thompson, G. G. 1999. Optimizing harvest control rules in the presence of natural variability and parameter uncertainty. *In* V. R. Restrepo (ed.), *Proceedings of the 5th National NMFS Stock Assessment*

Workshop: Providing scientific advice to implement the precautionary approach under the Magnuson-Stevens Fishery Conservation and Management Act, p. 124-145. NOAA Tech. Memo. NMFS-F/SPO-40. National Marine Fisheries Service, NOAA. 1315 East-West Highway, Silver Spring, MD 20910.Thompson, G. G. 2018. Analysis of past BSAI and GOA Groundfish Plan Team recommendations to set ABC below the maximum permissible level. Document submitted to the September 2018 meeting of the Joint Groundfish Plan Teams. https://meetings.npfmc.org/CommentReview/DownloadFile?p=f5899836-d32c-45a7-b565-93d1d44dc10c.pdf&fileName=Reverse-engineering%20previous%20ABC%20reductions%203.pdf

Wetzel, C. R. and O. Hamel. 2019. Accounting for increased uncertainty in setting precautionary harvest limits from past assessments. Document prepared for Pacific Fishery Management Council March 2019 meeting.

Tables

Table 1. Risk classification table for assessment, population dynamics, and environmental/ecosystem considerations.



	Assessment-related considerations	Population dynamics considerations	Environmental/ecosystem considerations	Fishery Performance
Level 1: Normal	Typical to moderately increased uncertainty/minor unresolved issues in assessment.	Stock trends are typical for the stock; recent recruitment is within normal range.	No apparent environmental/ecosystem concerns	No apparent fishery/resource-use performance and/or behavior concerns
Level 2: Substantially increased concerns	Substantially increased assessment uncertainty/ unresolved issues.	Stock trends are unusual; abundance increasing or decreasing faster than has been seen recently, or recruitment pattern is atypical.	Some indicators showing an adverse signals relevant to the stock but the pattern is not consistent across all indicators.	Some indicators showing adverse signals but the pattern is not consistent across all indicators
Level 3: Major Concern	Major problems with the stock assessment; very poor fits to data; high level of uncertainty; strong retrospective bias.	Stock trends are highly unusual; very rapid changes in stock abundance, or highly atypical recruitment patterns.	Multiple indicators showing consistent adverse signals a) across the same trophic level as the stock, and/or b) up or down trophic levels (i.e., predators and prey of the stock)	Multiple indicators showing consistent adverse signals a) across different sectors, and/or b) different gear types
Level 4: Extreme concern	Severe problems with the stock assessment; severe retrospective bias. Assessment considered unreliable.	Stock trends are unprecedented. More rapid changes in stock abundance than have ever been seen previously, or a very long stretch of poor recruitment compared to previous patterns.	Extreme anomalies in multiple ecosystem indicators that are highly likely to impact the stock. Potential for cascading effects on other ecosystem components	Extreme anomalies in multiple performance indicators that are highly likely to impact the stock

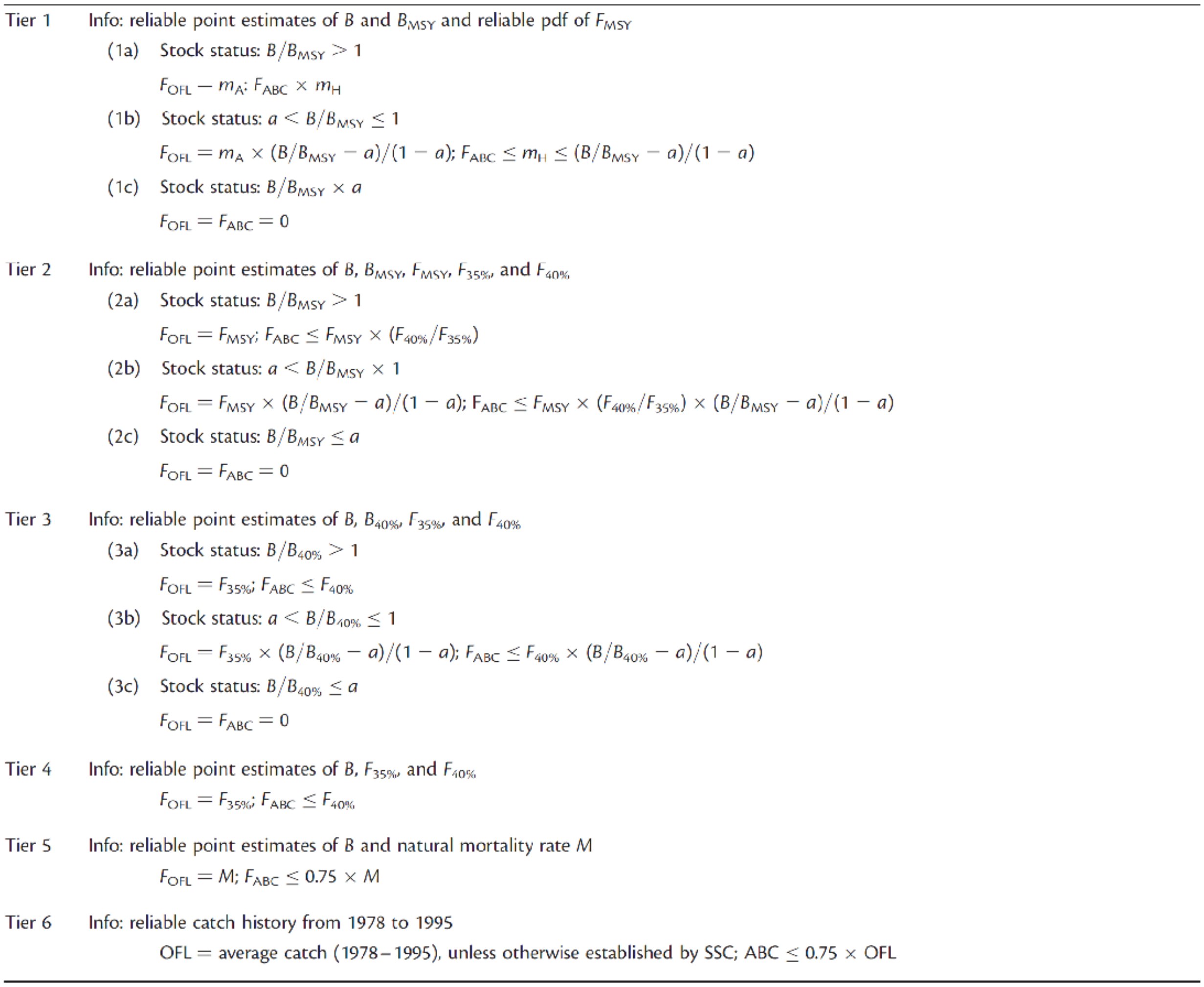
Table 2: Summary table of risk table scores for Alaska groundfish from 2018 to 2020. Note that the fishery performance category was added in 2019.

Stock	Assessment related			Population Dynamics			Environment / Ecosystem			Fishery Performance		Proposed Reduction		
	2018	2019	2020	2018	2019	2020	2018	2019	2020	2019	2020	2018	2019	2020
Sablefish	2	2	3	4	3	3	2	2	2	3	3	45%	57%	57%
EBS pollock	1	1	1	2	2	1	2	2	2	2	2	30%	43%	30%
Bogoslof pollock			1			1			1		1			0%
AI pollock			1			1			1		1			0%
EBS Pacific Cod		1	1		1	1		2	2	1	1		*	0%
AI Pacific cod		1	1		1	1		2	2	1	1		**	0%
BSAI Yellowfin sole		1	1		1	1		1	1	1	1		0%	0%
Alaska Plaice		1			1			1		1			0%	
BSAI Greenland turbot			1			1			2		1			0%
BSAI Arrowtooth			1			1			1		1			0%
BSAI Kamchatka			1			1			1		1			0%
BSAI Northern rock sole			2			1			1		1			0%
BSAI Flathead			1			1			1		1			0%
BSAI Other Flatfish			1			1			1		1			0%
BSAI POP			2			1			1		1			0%
BSAI Blackspotted/ RE			3			2			1		2			0%
BSAI Northern Rockfish		2			1			2		1			0%	
BSAI Shortraker			1			1			1		1			0%
BSAI Other Rockfish			2			1			1		1			0%
BSAI Atka Mackerel	1	1	1	1	1	1	1	1	1	1	1	0%	0%	0%
BSAI Skates			1			1			1		1			0%
BSAI Sharks			2			2			1		1			0%

BSAI Octopus			1			1			1		1			0%
GOA pollock	2	2	1	2	1	1	2	1	1	1	1	15%	10%	0%
GOA Pacific cod	2	2	2	***	2	2	2	2	1	1	1	13.6%	****	0%
GOA Northern Rockfish			1			1			1		1			0%
GOA Arrowtooth		1			1			2		1			0%	
GOA Deepwater Flatfish		2			1			1		1			0%	
GOA POP		2	2		2	2		1	1	1	1		0%	0%
GOA Northern Rockfish			1			1			1		1			0%
GOA Dusky Rockfish			2			1			1		1			24%
GOA Rougheye/BS		1			1			1		1			0%	
GOA Thornyheads			1			1			1		1			0%
GOA Other Rockfish		1			1			1		1			0%	
GOA Shortraker		1			1			1		1			0%	
GOA Atka Mackerel		1			*****			1		1			0%	
GOA Skate		1			1			1		1			0%	
GOA Sharks			2			2			1		1			0%
GOA Octopus		1			1			1		1			0%	

*Authors did not provide a recommendation and deferred to the SSC. The SSC adopted the ensemble to lower the ABC.
**Authors did not provide a recommendation and deferred to the SSC. The SSC did not recommend a reduction since the stock was at Tier 5.
***Author recommended a level 4 for population dynamics in 2018 and the SSC downgraded that to a level 2.
****Authors did not provide a recommendation and deferred to the SSC. The SSC set the 2021 ABC the same as the 2020 ABC.
*****Author stated "Unknown" for this category

Table 3. Description of the groundfish tier system used by NPFMC since 1999 for defining fishing-mortality rate related to the overfishing level (FOFL) and the acceptable biological catch (FABC) based on the type of information available (From DiCosimo et al. 1991).



Tier 1 Info: reliable point estimates of B and B_{MSY} and reliable pdf of F_{MSY}

(1a) Stock status: $B/B_{MSY} > 1$

$F_{OFL} - m_A$; $F_{ABC} \times m_H$

(1b) Stock status: $a < B/B_{MSY} \leq 1$

$F_{OFL} = m_A \times (B/B_{MSY} - a)/(1 - a)$; $F_{ABC} \leq m_H \leq (B/B_{MSY} - a)/(1 - a)$

(1c) Stock status: $B/B_{MSY} \times a$

$F_{OFL} = F_{ABC} = 0$

Tier 2 Info: reliable point estimates of B, B_{MSY}, F_{MSY}, $F_{35\%}$, and $F_{40\%}$

(2a) Stock status: $B/B_{MSY} > 1$

$F_{OFL} = F_{MSY}$; $F_{ABC} \leq F_{MSY} \times (F_{40\%}/F_{35\%})$

(2b) Stock status: $a < B/B_{MSY} \times 1$

$F_{OFL} = F_{MSY} \times (B/B_{MSY} - a)/(1 - a)$; $F_{ABC} \leq F_{MSY} \times (F_{40\%}/F_{35\%}) \times (B/B_{MSY} - a)/(1 - a)$

(2c) Stock status: $B/B_{MSY} \leq a$

$F_{OFL} = F_{ABC} = 0$

Tier 3 Info: reliable point estimates of B, $B_{40\%}$, $F_{35\%}$, and $F_{40\%}$

(3a) Stock status: $B/B_{40\%} > 1$

$F_{OFL} = F_{35\%}$; $F_{ABC} \leq F_{40\%}$

(3b) Stock status: $a < B/B_{40\%} \leq 1$

$F_{OFL} = F_{35\%} \times (B/B_{40\%} - a)/(1 - a)$; $F_{ABC} \leq F_{40\%} \times (B/B_{40\%} - a)/(1 - a)$

(3c) Stock status: $B/B_{40\%} \leq a$

$F_{OFL} = F_{ABC} = 0$

Tier 4 Info: reliable point estimates of B, $F_{35\%}$, and $F_{40\%}$

$F_{OFL} = F_{35\%}$; $F_{ABC} \leq F_{40\%}$

Tier 5 Info: reliable point estimates of B and natural mortality rate M

$F_{OFL} = M$; $F_{ABC} \leq 0.75 \times M$

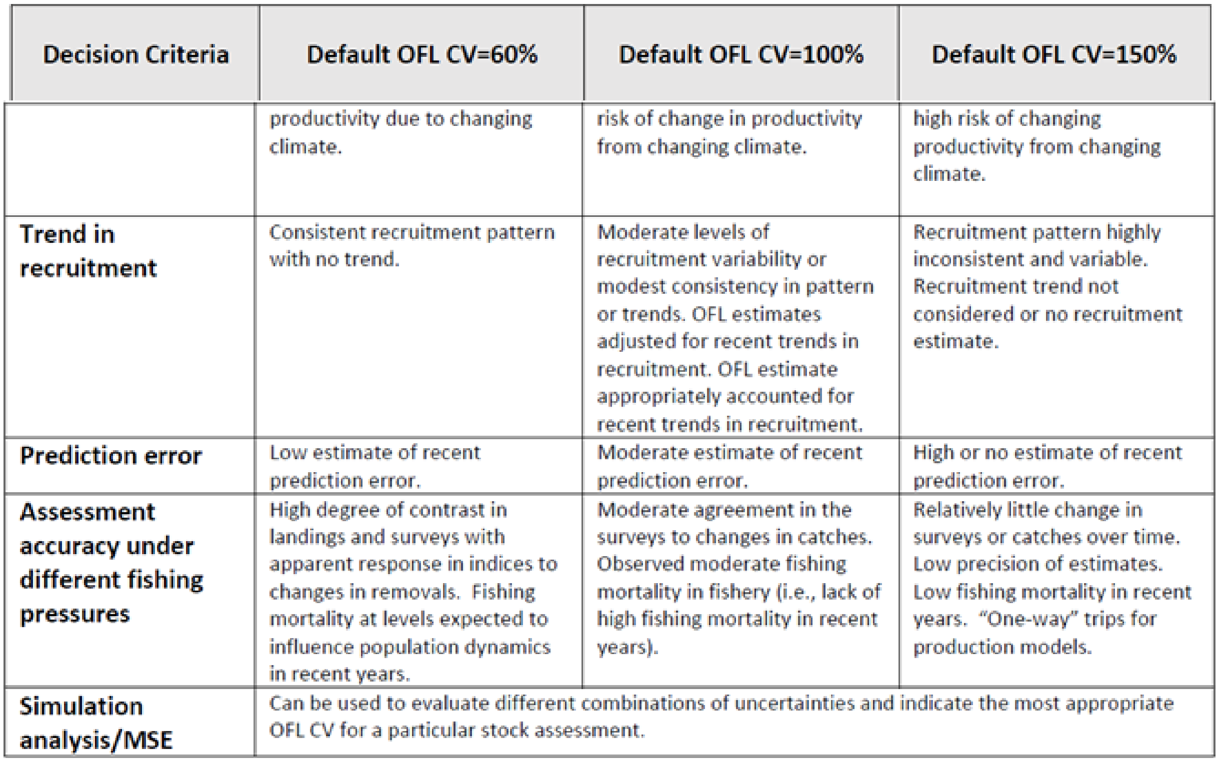
Tier 6 Info: reliable catch history from 1978 to 1995

OFL = average catch (1978–1995), unless otherwise established by SSC; ABC $\leq 0.75 \times$ OFL

Table 4. Mid-Atlantic Fishery Management Council framework table for assessment 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.	Single model structure with many parameter sensitivities explored. Moderate agreement among different model runs indicating low 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.	Moderate retrospective patterns.	No retrospective analysis or severe retrospective patterns.
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 empirical estimates or simpler analyses.	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 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	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 productivity and habitat quality are variable, with mixed productivity and uncertainty signals among comparable species in the region. Climate vulnerability analysis suggests moderate	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 suggests that ecosystem productivity and habitat quality are variable and degrading. Comparable species in the region have high uncertainty in short term predictions. Climate vulnerability analysis suggests

Table 4. Continued. Mid-Atlantic Fishery Management Council framework table for assessment 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%
	productivity due to changing climate.	risk of change in productivity from changing climate.	high risk of changing productivity from changing climate.
Trend in recruitment	Consistent recruitment pattern with no trend.	Moderate levels of recruitment variability or modest consistency in pattern or trends. OFL estimates adjusted for recent trends in recruitment. OFL estimate appropriately accounted for recent trends in recruitment.	Recruitment pattern highly inconsistent and variable. Recruitment 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 in recent years.	Moderate agreement in the surveys to changes in catches. 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.		

Table 5. Counts of "positive" risk factors that were listed more than once in current BSAI risk tables.

Rationale	Count
Small retrospective bias	10
Good fits to data overall	9
Results from analysis of missing a survey indicate no major problems	8
Annual surveys through 2019	4
No convergence issues	3
Good availability of age data	2
Mis-ageing is not a concern	2
New data had little impact	2
Recruitment estimates are consistent with the data	2
Retrospective bias is improved relative to previous assessments	2

Table 6 . Counts of "negative" risk factors that were listed more than once in current BSAI risk tables.

Rationale	Count
Lack of 2020 summer trawl surveys	10
Large retrospective bias	6
Data conflicts exist	4
Lack of EBS slope surveys since 2016	4
Tight prior distributions cause uncertainty to be underestimated	3
Alternative models show disparate results	2
Shortcomings in harvest control rule	2
Shortcomings in structure of Tier 5 RE model	2
Strong residual patterns	2
Survey biomass estimates are relatively imprecise	2

Figures

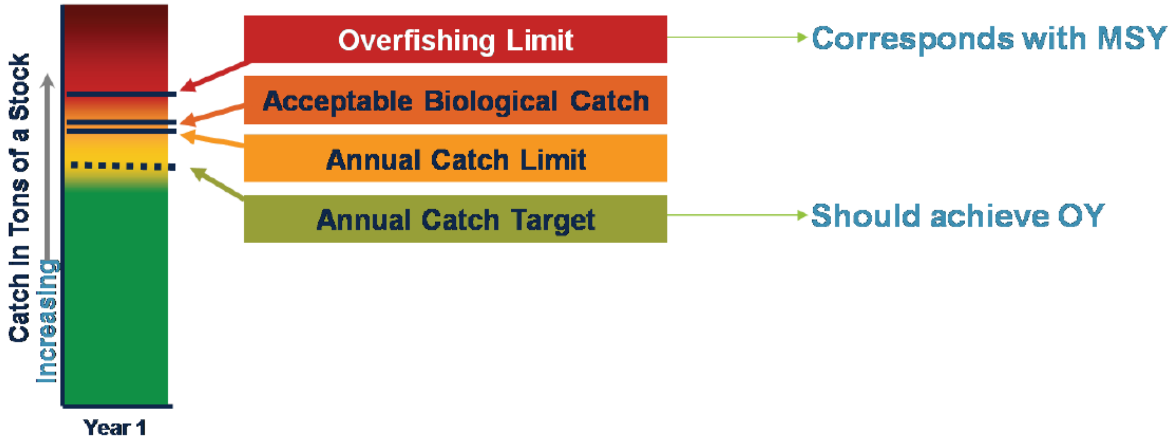


Figure 1. Fishing limits and targets defined by the National Standard 1 Guidelines (2009).

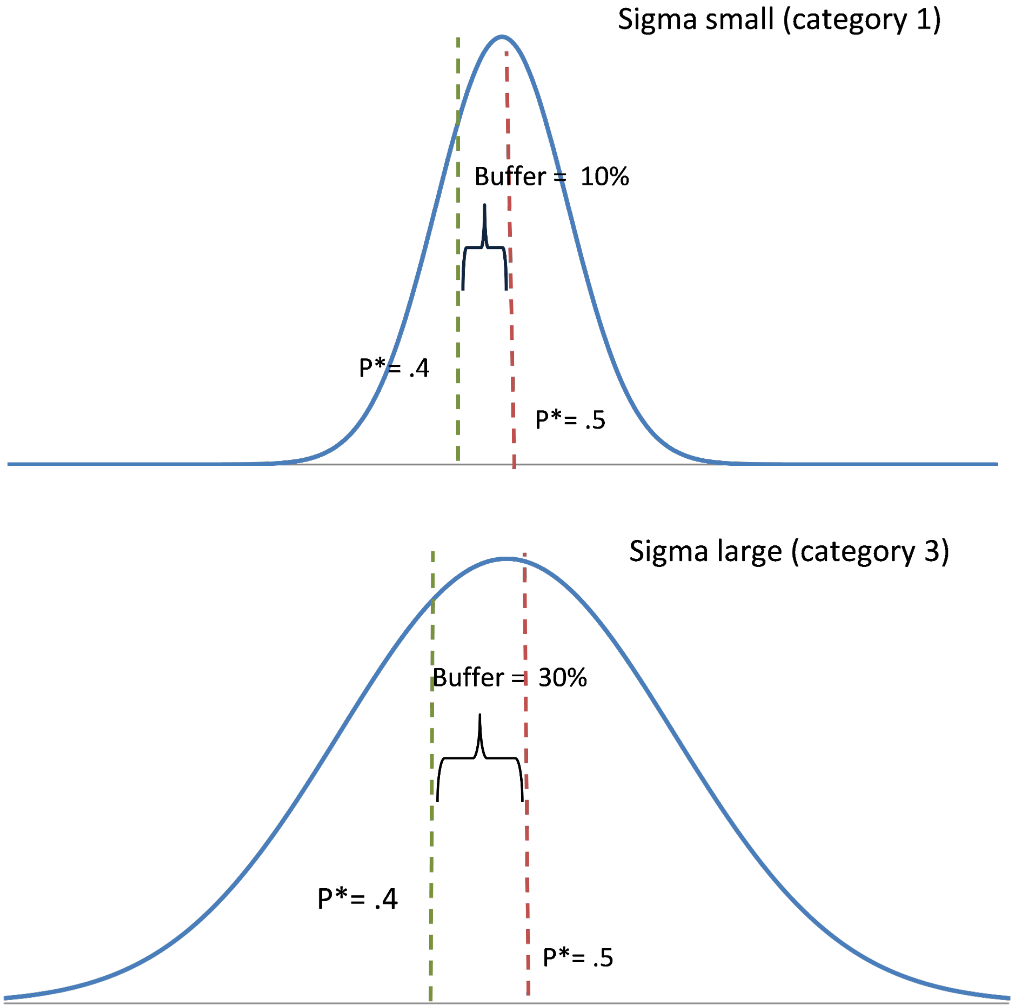


Figure 2. Illustration of the P-star approach.

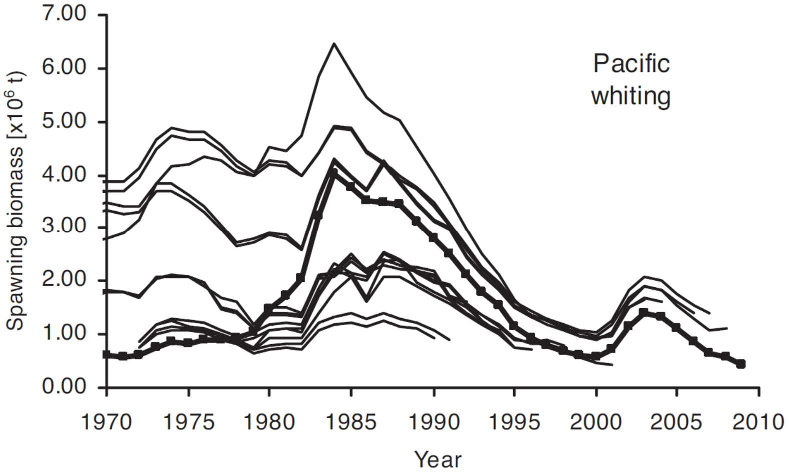


Figure 3. Example of between assessment variation in estimated spawning biomass (from Ralston et al. 2011).

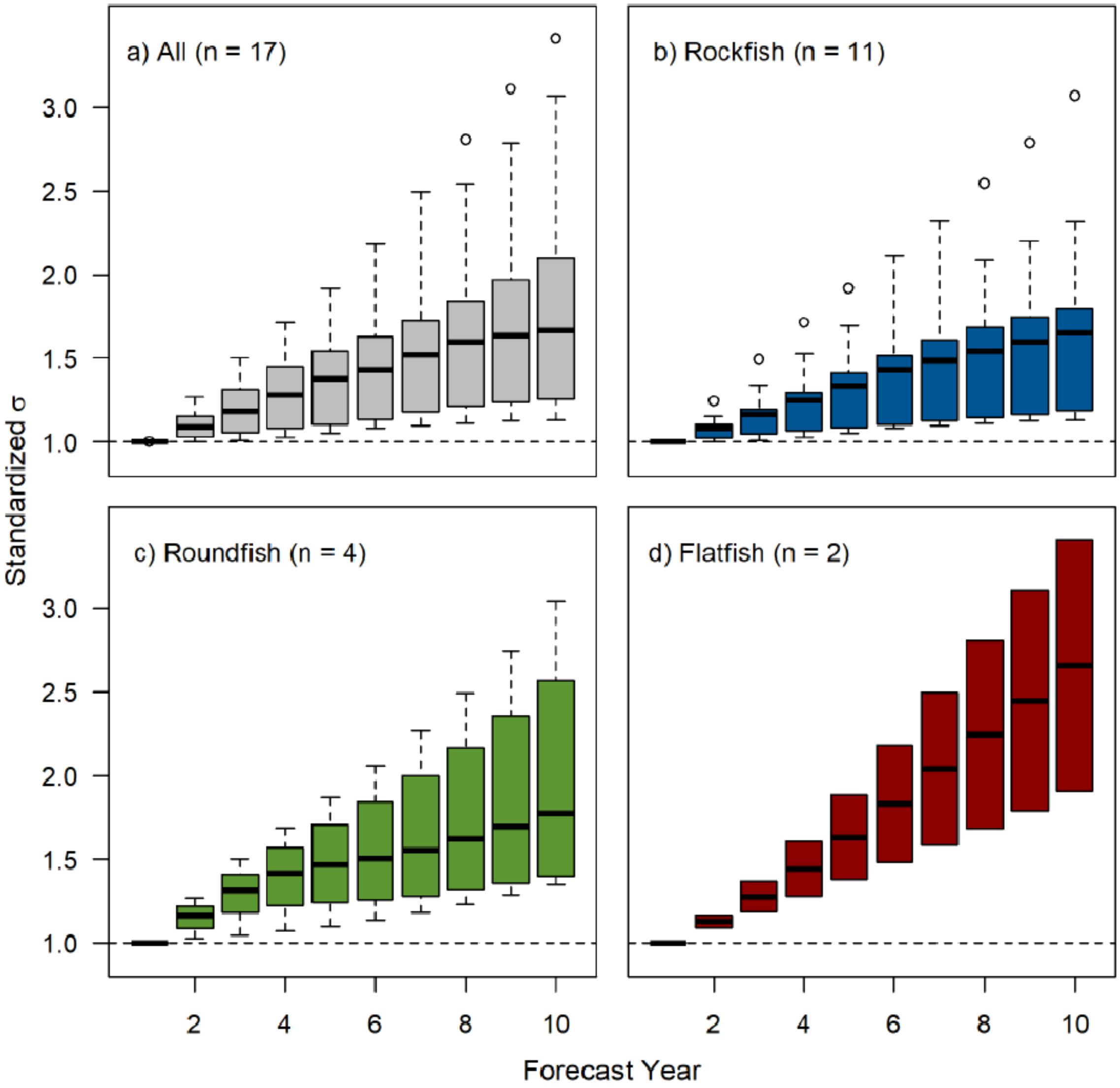


Figure 4. Change in sigma reflecting increased uncertainty during the projection period grouped by life history for assessed West Coast stocks. The number of species in each life history grouping is shown in each figure. (From Wetzel and Hamel 2019).

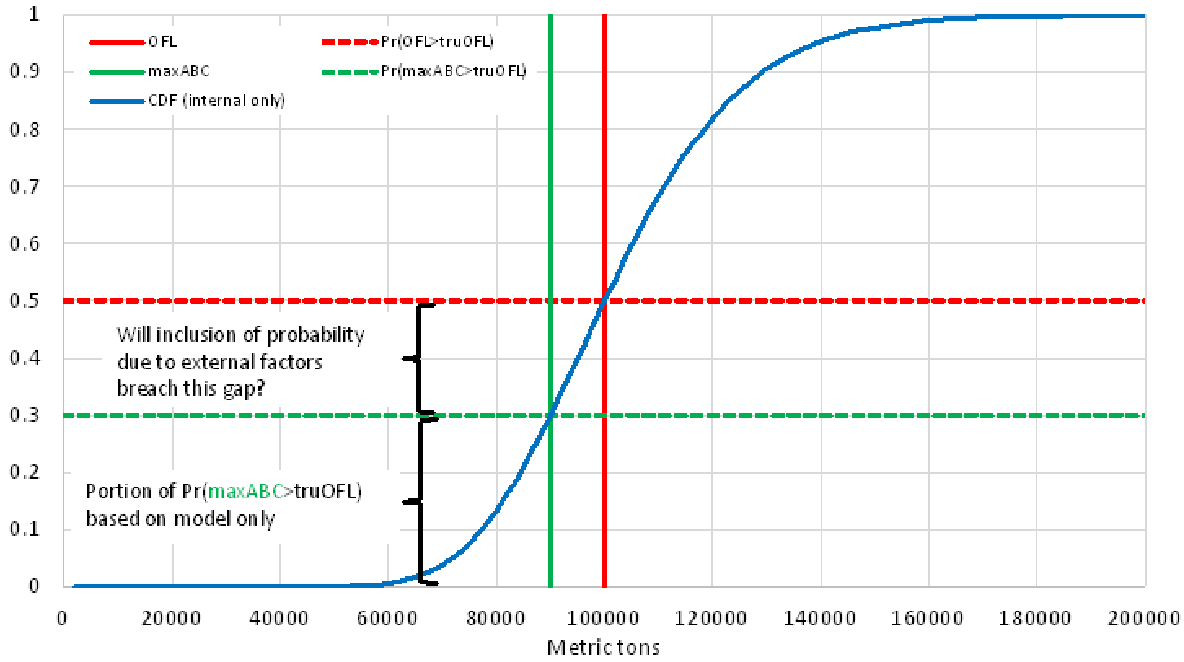


Figure 5 . Example cumulative distribution function of *truOFL* (lognormal with μ=100,000 and σ=0.2).

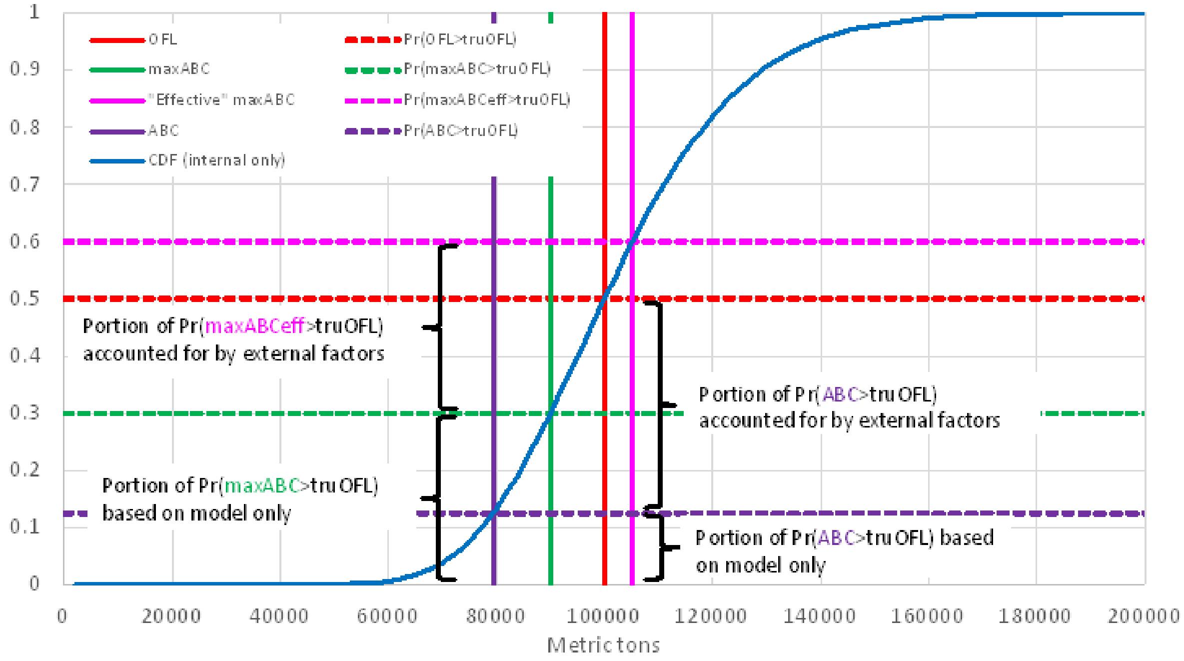


Figure 6 . Application of the joint probability approach to the example shown in Figure 1, with α=0.2 and *score*=0.375 for all categories.

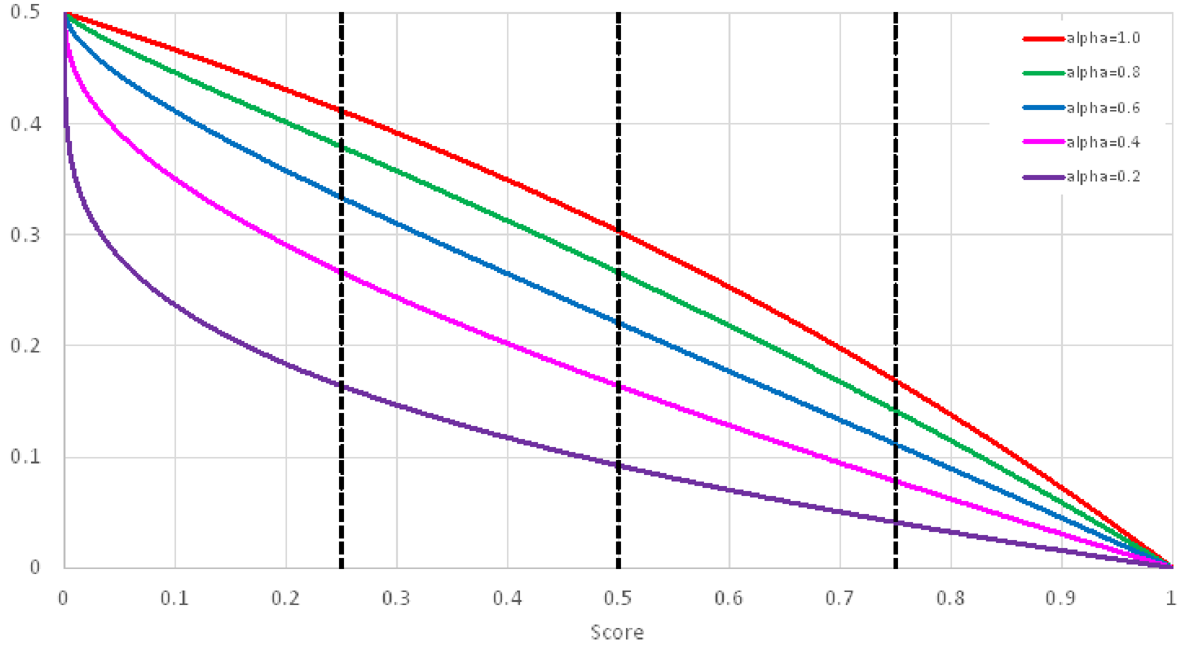


Figure 7 . *Pabc* as a function of *score* for various values of *α* for the special case where *score* is the same for all categories.

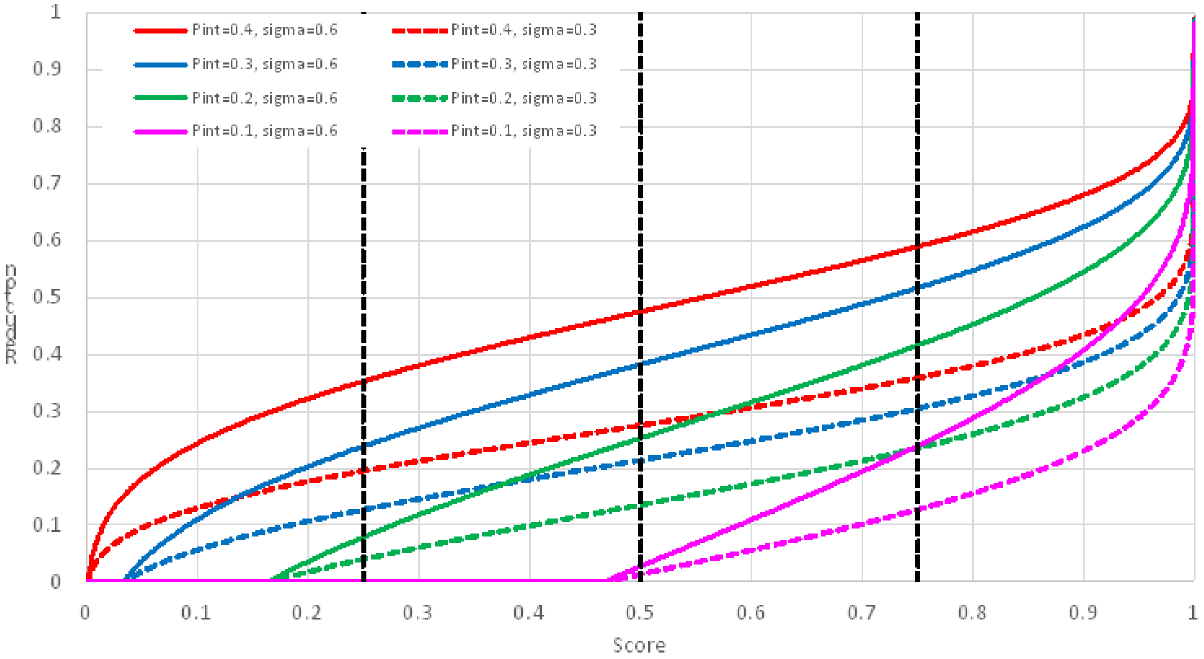


Figure 8 . Reduction from *maxABC* as a function of *score* for various combinations of *Pint* and lognormal *σ* for the special case where *α*=0.2 (as in Figure 2) and *score* is the same for all categories.

Appendix 1.

Agenda

Feb 5, 8am - 12:00pm AST

8:00 - 8:10 AM AKT Introduction and workshop goals
8:10 - 8:30 (20 min) Summary of case studies for risk table adjustments
(Shotwell, Zador and Dorn)

- Brief historical overview of risk tables (timeline and purpose) (Table 1)
- Synthesis of stock responses 2018-2020 (Tables 2 and 3)
- Lessons learned (e.g., transparency, evaluation of consistency, documentation of when there is no concern, acknowledgement of novel observations)

8:30 - 8:50 (20 min) Issues, challenges and concerns
Group Discussion (Anne Hollowed facilitator)

- Challenges for species complexes
- Challenges for data limited stocks
- Interpreting response for non-target stocks
- Should we continue to produce risk tables for all (or any) full assessments
- Challenges with time constraints - Should Plan Teams and SSC review all of them, or only when a reduction is recommended?

8:50 - 9:40 (50 min) Breakout Session 1
Discussion of tangible steps towards quantifying the importance of external changes in fishery performance in stock assessments
(Haynie, Ianelli, and Kasperski)

- Are there mechanistic linkages to stock health that are revealed by fishery performance data?
- Recognizing Data lags and interpreting trends (e.g., fishery selectivity changes in most recent years, predictions in future years).
- How do we define “fishery” performance risk for bycatch stocks, and how would it be affected by changes in bycatch and incidental catch?

Discussion of tangible steps toward quantifying the importance of assessment risk (Thompson)

- Data selection
- Parameterization
- Trade-offs in model complexity
- Ensembles

8:50 - 9:40 (50 min) Breakout Session 2
Discussion of tangible steps towards quantifying the risk of external changes in population conditions
(Spencer and Goethel)

- Importance of age diversity?
- Importance of recruitment uncertainty?
- Importance of growth uncertainty?
- Importance of maturation uncertainty?
- Interaction between perceived stock status and population risk category.

Discussion of tangible steps towards quantifying risk of external changes in ecosystem conditions

(Shotwell, Ferriss, Siddon, Zador)

- Mechanistic linkages quantifying risk of ecosystem process (the four factors).
- Pathway for moving from recognition of ecosystem anomalies to qualitative projection of risk of overfishing.

9:40 - 9:55 Break

9:55 - 10:25 30 min Plenary discussion of key findings from breakout groups

The following session will include ~ 30 min of introduction to the topic followed by a ~15 min open discussion between stock assessment authors, PTs and SSC

10:25 - 11:10 (45 min) Frameworks for addressing scientific uncertainty

(Dorn and Thompson)

- What sources of scientific uncertainty are already incorporated in the existing buffer between ABC and OFL? Do these differ from the Risk Table?
- P* approaches for crab and PFMC Decision theoretic approaches
- A probabilistic approach for linking the risk table to ABC reductions
- Full feedback MSE
- Scoring - pros and cons of overall scores?
- Should "increased" concern be evaluated relative to: (1) previous assessments of the same stock/complex or (2) typical assessments with the same tier or (3) typical assessments across all tiers, conditions under which elevated risk levels should result in reduction from maxABC.

11:10 - 11:40 (30 min) Public Testimony relevant to workshop topics

11:40 - 12:00 (20 min) SSC discussion