Assessment Methods Workshop Report

Ensemble Modeling and ABC adjustments

The NPFMC BSAI Groundfish Plan Team convened a public workshop June 27-28, 2018 on ensemble modeling and approaches to making ABC adjustments with the following objectives:

- to review ensemble stock assessment modeling and evaluate how it fits in the NPFMC system;
- to discuss considerations for potentially reducing an ABC from the maximum to account for observations and uncertainties not included in the assessment model or Tier system, and
- to produce recommendations and a report to be considered by the September Joint Groundfish Plan Team.

The workshop was divided into two main sections. Ensemble modeling and Determining ABC.

Workshop participants (in person in Seattle or via Webex for all or part) included: **AFSC:** Steve Barbeaux, Alan Haynie, Grant Thompson, Jim Ianelli, Sandra Lowe, Ron Felthoven, Tom Wildebuer, Dana Hanselman, Marin Dorn, Paul Spencer, Chris Rooper, Stan Kotwicki, Andy Kingham, Mike Dalton, Kari Fenske, Cindy Tribuzio, Kalei Shotwell, Elizabeth Siddon, Bob Foy, Steve Kasperski, Farron Wallace, Dan Lew, Jennifer Cahalan. **IPHC:** Allan Hicks, **NWFSC:** Kelli Johnson. **NMFS HQ**: Rick Methot. **NMFS AKRO**: Obren Davis, Jason Gasper, **UW/JISAO:** Nick Bond**. WDFW:** Lisa Hilliard, Theresa Tsou. **Oregon DFW:** Allison Whitman**. ADF&G:** Bob Clark, Katie Palof, Sara Miller, Jane Sullivan, Shareef Siddeek, Jocelyn Runnebaum. **NPFMC:** Diana Stram, Jim Armstrong**. Non-agency public:** Craig Cross (Aleutian Spray Fisheries). Chad See (FLC), Ruth Christianson (UCB), Susan Robinson (Ocean Peace), Paul Wilkins, Richard Timmell (ALF)

The agenda for the workshop is in attachment 1. All presentations are available here for those with a NOAA email and others by requesting access. Because the full presentations are available online, this report provides only a very brief summary of them, and instead focuses primarily on the discussions held during the workshop, followed by the recommendations for consideration by the full Joint Groundfish Plan Team in September.

Part 1 Ensemble Modeling

Brief description of ensemble modeling and model averaging.

Grant Thompson provided general information on ensemble models and introduced information related to the evaluation of the pros and cons of ensemble modeling as compared with single stock assessments for use in recommending catch quantities. This led to discussion of the topics to be considered in ensemble modeling approaches, some of which were covered later in the workshop but identified here.

The group discussed some perceived benefits of model averaging, including the potential for better estimation of summary statistics and true uncertainty. The group discussed some procedures used on the west coast, such as substituting a meta-analytic estimate of between-assessment uncertainty for the estimated within-assessment uncertainty whenever the former exceeds the latter. There was discussion of different issues to be resolved including the difference between a sensitivity run and model being considered for use in setting harvest specifications. The group also questioned what the role of management strategy evaluations (MSEs) might be in ensemble modeling.

The challenges of clearly communicating results was noted as a key discussion point in considerations moving forward. Specifically, concern over difficulty in communicating ensemble model results and practices (relative to current approaches which are already complex) was a common concern. Less transparency and public engagement are a likely cost of moving towards more complex methods. A potentially more constructive application of multiple models may be to use them as operating models that can then be used in MSEs or other simulations to test simpler, more transparent catch control rules. The group discussed the issue of whether current models are actually more transparent or less complex than ensemble models. The Council has expressed interest in developing such approaches and workshops such as this and measures to improve communication may help future application of ensemble models for management.

Can ensemble models improve biological realism relative to the current approach? For example, would they likely provide a means for management measures to be more reactive to changing environmental conditions? Alternatively, including different models that are less likely to represent reality may add uncertainty and reduce the effectiveness of management measures when models are averaged.

The dividing line between statistics and machine learning.

Grant Thompson presented an overview of machine learning as it relates to ensemble modeling. Some specific issues were noted as to how machine learning differs from what is currently done in tuning stock assessments. Machine learning appears to go beyond model averaging. Three examples of how this improves performance were mentioned: bagging (decreasing variance), boosting (decreasing bias), and stacking (improving predictions).

One main difference lies in the intent of machine learning being to predict the future population while stock assessment attempts to provide a mechanistic explanation of how past conditions give rise to current conditions. Machine learning may work for predicting quantities such as survey biomass, but with stock assessment our primary goal is to predict the correct value of next year's ABC. However, we lack previous observations of what the "correct" ABC, which would be required to train the model. Our assessment models are designed to reflect what is occurring in nature whereas machine learning methods are more "black box," providing a good estimate regardless of how plausible the underlying structure is.

Examples of ensemble modeling in fisheries stock assessment: the American experience

Allan Hicks provided an overview of the use of ensemble modeling in the IPHC as well as some issues and discussions on the Bering Sea cod model from 2017. He presented a historical overview of IPHC modeling and issues being addressed in chasing the "perfect" assessment model. In their experience, a single assessment model cannot capture all of the issues necessary to capture the full uncertainty; thus the IPHC transitioned to an ensemble modeling approach and moved from catch advice to risk advice in doing so. Four models are included in the ensemble, based on treatments of spatial and historical data. The distributions of spawning biomass resulting from the four models differ. The justification for use of the ensemble is that each model has good and bad fits, and pros and cons in singular use. The outputs from the models are equally weighted, and account for the overall uncertainty in each. With their ensemble model they can better characterize integrated results and probabilities (e.g., reference points). Presentations of point estimates are avoided in favor of risk probabilities.

Allan also provided an overview of the 2017 BS Pacific cod assessment. Six different models were presented. The range of resulting ABCs was a substantial change from 2017, which formed the basis of much of the Team discussion and industry concerns at that time. The Team selected a best model, not an average across models (model averaging was presented in an appendix with model weighting and a review of factors for comparing across models). The Team then recommended reducing the ABC given concerns with model results, desirable features of other models, and potential movement of cod. The SSC selected the same model but did not recommend reducing the ABC because, in their words, there was "not unequivocal information justifying a further reduction."

There was considerable workshop discussion regarding how the ABC was adjusted by the Team, which essentially used model averaging to select a lower ABC. Concerns were raised that there should have been more discussion at the time on the identification of and development of plausible models. The choice of factors and weights associated with this weighting would benefit from a more complete discussion, and presentation in the minutes, of the logic of model selection.

Examples of ensemble modeling in fisheries stock assessment: the ICES experience

Kelli Johnson provided an overview of progress on an ICES project that uses a form of ensemble modeling. Here ensemble modeling is being used primarily in an MSE framework with six planned case studies. This project is still under development and discussion was held of some of the weighting schemes being considered as well as what key aspects should be considered, including the differences in 2- to 4-year projections as well as treatment of recruitment deviations. She noted that an ICES workshop is being held in Portugal in August 2018 and more discussion of this study will be held then.

Examples of ensemble modeling in other disciplines

The SSC requested that the workshop involve "feedback from other fields that use ensemble modeling, such as weather." Nick Bond provided an overview of ensemble modeling examples being used in weather and climate predictions. He described the development and rationale for the use of ensemble modeling in weather prediction and noted that the ensemble mean is more accurate on average. He described the treatment of uncertainty both in initial conditions as well as in model structure and numerical methods.

There are a variety of ways they view ensembles and their output and this is critical to what information is intended to communicate to the public. Some examples he showed included spaghetti diagrams which provide at a glance the consistency and robustness of results. Other examples included forecasts of SST from different models, showing individual model results, and a national multi-model ensemble averaged together. They also show the probability of exceeding some amount; e.g., plumes as output and showing the same model run with different initial conditions.

The group discussed the applicability of these methods to stock assessment. One key complication was noted in that, in the case of weather, the quantities being predicted are directly observable with minimal measurement error, whereas in the case of fisheries stock assessment, this is not the case. The group discussed the difficulty in communicating probability. Nick noted that, although people often think that they understand probability, in reality they typically do not. One suggestion for application to stock assessment would be for short-term forecasts of catch, looking at a range of initial conditions.

Lessons from the 1998 NRC study

The SSC suggested that, "it may also be desirable to obtain one or more datasets with known information to explore during the workshop. The datasets developed by the National Research Council in 1998, and the results of models fitted to those datasets could be used to compare single vs. multi-model approaches." Grant Thompson provided a review of the 1998 NRC study. The NRC study was focused only on point estimates, as statistical PDFs were not included in the report. The study provided indications that simulation approaches could be used to test which individual models tend to perform better than others. The ensemble average did not always out-perform the single best model; but, absent such simulation testing, it is usually be difficult to know *a priori* which model is the single best one. One limitation of the study was that the various models were not all tested against all data sets and did not all report the same quantities, making comparison difficult.

Review of the 2017 SSC ensemble modeling workshop

Allan Hicks provided an overview of the 2017 SSC workshop and a review of the resulting minutes. The remainder of this section pertains to discussion held during the 2018 Team workshop. The group discussed issues associated with how long it takes to run the number of models used in the other examples of ensemble modeling as well as additional discussion of the utility of approaches for ensemble modeling as compared to an MSE. Some participants commented on the difficulty of communicating risk analysis results to managers in a way that facilitates making harvest specifications, noting that harvest specifications are largely a policy decision. Participants also suggested that ensemble results are, in a sense, already embedded into the MCMC posterior distribution. The focus now should be on adding structural uncertainty by use of ensemble modeling.

Questions were raised with respect to how some modeling efforts in the BSAI dealt with alternative ranges of years for estimating stock-recruitment (SR) parameters. One example from flatfish uses three time series, resulting in range of curves. This could provide a good example of how to account for uncertainty in SR relationships, or how to expand treatment of uncertainty in general. It was suggested that ensemble approaches could potentially be used to explore a range of SPR values, leading to a range of proxy MSY estimates.

Review of the 2018 NSAW on ensemble modeling

Dana Hanselman provided an overview of the recent ensemble modeling component of the 2018 NSAW. With respect to acceptance by the public, NSAW participants envisioned two contrasting outcomes, one in which the models in the ensemble mostly tend to produce similar results, thereby strengthening public confidence, and another in which the models mostly tend to produce *dis*similar results, thereby weakening public confidence. The IPHC approach is the only one in use for management in the US although US scientists are involved internationally in multiple studies and a few US regions are attempting to experiment with ensemble modeling presently.

The remainder of this section pertains to discussion held during the 2018 Team workshop. Some participants felt that ensemble modeling could be useful when we have two very different viewpoints of the stock. Alternatively, the ensemble could be inappropriate if there is a sudden change in the ecosystem, as with GOA P. cod, and if some of the models in the ensemble are incapable of responding to such a change. Alternatively, the BSAI Team's decision to focus on only two (fairly similar) models out of six in the case of EBS P. cod may have resulted in an insufficient response to a sudden change in biomass.

Some participants suggested that it would be better for the ensemble to consist of a fixed set of models for at least a few years, rather than attempting to review a wide band of models each year. The IPHC ensemble has been mostly static for the 3 years following its initial development, but next year a full assessment will be conducted, and other models will be evaluated. Concerns were raised that moving to this approach from a single stock assessment model may be an irreversible decision. Assessment is an annual or semi-annual exercise and concerns were expressed that an ensemble approach may lock us into a static range of alternatives. There was discussion of how to choose models such that one is not simply selecting a symmetrical range that averages out to the same outcome as the single best model.

Choosing models in an ensemble

Grant Thompson provided an overview of considerations in selecting models for an ensemble based upon a list developed during the BSAI Plan Team review in November 2017. For discussion purposes he provided some preliminary recommendations for these considerations. Some discussion topics included: 1) why you would exclude models that are less plausible but still reasonable and 2) whether the purpose of a certain model is testing hypotheses or addressing structural uncertainty. Discussion considered what constitutes sensitivity in this context and the difference between testing robustness versus sensitivity, demonstrating axes of uncertainty, and factors that can really affect results. There was a suggestion to use retrospective analysis to help inform weighting.

Models useful for management should be able to provide a basis for setting ABC and OFL as well as status determination. Some suggested that models should also be able to provide advice that explicitly addresses shifts in fishery or ecosystem dynamics.

There was discussion of linking the ESR to assessments.

Other discussion questions included: Who should choose the models; should the author make the first cut? Should this be done through external (e.g., CIE) review or Plan Team review? One suggestion was to go through the assessment and, for each decision point in selecting model parameters or characteristics, evaluate whether it is a major decision point and, if so, provide a range of plausible alternatives. It was also noted that thorough application of this approach would inevitably result in an immense number of models, unless “major” is defined very narrowly.

It was suggested that the Plan Teams draft a set of guidelines for ensemble models, to be reviewed by the SSC and then given to assessment authors. Questions were raised as to whether we are limiting the ensemble to tweaking Stock Synthesis (or whatever software the assessment currently uses) or are we considering an ensemble which includes alternative modeling frameworks such as FEAST, CEATTLE, or others? There was concern expressed regarding whether assessment authors have the expertise to develop some of these alternative models.

There was a suggestion to compare single species models with multispecies models (e.g., CEATTLE) in the ensemble. This would imply that the benefits of including multi-species dynamics outweigh the inherent single-species simplifications inherent in the latter. CEATTLE can be run in both single species and multispecies modes for comparison. It may also be useful to take the value of M estimated by the multispecies model and use it in the single species assessment model for purposes of comparison (as with GOA P.cod). One question raised was how can ecosystem models inform the various hypotheses going into the range of models for the ensemble?

Combining models and assigning weights

Grant Thompson provided some examples of, and led a discussion on, combining models and assigning weights to models. The group discussed how best to weight models, noting that ignoring candidate models implies zero weight. Suggestions included ad hoc, AIC, and Bayesian model weighting, along with the method presented in the 2017 EBS P. cod assessment. Note that the IPHC assessment is able to use only ad hoc model weighting, because data and data weights vary across models. One suggestion for computing objective model weights was to base them on cross validation of important data components (e.g., survey index or age composition predictions relative to out-of-sample observations).

Calculating statistics and uncertainty

The SSC suggested that "some discussion about characterizing uncertainty, say with confidence intervals, using ensemble models should be considered." Grant Thompson presented results of a study in which a parametric estimator of the PDF associated with an arbitrary model quantity was compared to a correctly weighted average of the individual PDFs from the models in the ensemble. He found that the parametric estimator tended to give better performance. Workshop discussion of these results focused on hypothetical cases in which the true (population) distribution is strongly multi-modal, in which case the sample average would be expected to out-perform the parametric (unimodal) estimator if the models in the ensemble were chosen randomly from the population of all possible models; and also on the questions of how often such cases are likely to arise in nature and how large the ensemble would have to be in order to have confidence that the population distribution is truly multi-modal. Grant also summarized the discussion from last November's joint Team meeting in which the Teams responded to an SSC request for a discussion of whether the mean or the median is a more useful measure of central tendency in the context of model averaging. The Teams' recommendation in November was that "the choice of central tendency measure depends on the task at hand and the approach taken and that, in the context of model averaging, the choices involved in assembling the suite of models are likely more important than the choice of central tendency measure." No response has been received from the SSC.

Pros and cons of implementation in NPFMC system

The SSC suggested that the workshop should "discuss whether a Plan amendment would be necessary to utilize an ensemble of models." Grant Thompson and Jim Ianelli led a discussion on the pros and cons of implementation in the NPFMC system. Grant reviewed the language in the BSAI FMP and the SAFE chapter guidelines and concluded that, if an ensemble model is viewed as a model in its own right, neither text would have to be revised, although some clarifications may be helpful. Jim provided an overview of Southern Bluefin Tuna assessment and use of MSE to inform their management procedure. While Tier 1 in the NPFMC system directly incorporates uncertainty in the ABC there was concern for direct inclusion of uncertainty into Tier 3 assessments. Concerns were expressed that this may be technically challenging for assessment authors as well as the Plan Team and the SSC. It may be preferable to focus efforts instead on providing a firm rationale for model averaging.

Communicating and using results

Allan Hicks presented an overview of communicating ensemble modeling results from the perspective of the IPHC. The group discussed the communication of both quantitative and qualitative results, noting that the qualitative discussion is also important for characterizing model results. Perhaps focus groups could be convened for the purpose of looking at example results and considering how best to communicate them. A more general workshop on communication of science might also be helpful.

Workload and logistics for assessment authors

Grant Thompson provided an overview of issues potentially associated with ensemble modeling. How much more work would ensemble modeling be than just presenting the individual models? Can we quantify why ensemble modeling is causing additional work? It was noted that the initial steps to create the ensemble involves significant work but then potentially less work moving forward if the ensemble remains constant over time. Frontloading on document structure (e.g., using R markdown to make updating documents faster) could be useful to save time in long run.

The group emphasized the need for authors' discretion in pursuing ensemble modeling.

Identifying assessments amenable to ensemble modeling

The SSC recommended that the workshop "attempt to identify one or more stock assessments to further test the application of ensemble modelling for presentation to the Groundfish Plan Teams in September 2018." The group held an open discussion on identifying assessments amendable to ensemble modeling. Some factors were noted as potential considerations for developing an ensemble modeling approach for a given stock including: model stability issues, models provide different results, data have inherent signals that are not easily fit/explained by single model, and importance of the assessment to the fishery (e.g., catch relative to ABC). It was suggested that ensemble modeling should address a specific need and not be a default for all stocks. There are two potential applications: for use in management and for use in research. Both would have utility in the NPFMC system. One could help estimate the robustness of a selected model, while the other could indicate that none of the models are 100% preferred and that an ensemble approach would be preferable in order to mitigate selection of a single 'wrong' model.

The group noted that this report and recommendations should weigh the pros and cons of use in NPFMC system.

Part 1 Ensemble Modeling wrap-up

Pros/Cons

The working group identified the following benefits of ensemble models and model averaging.

- An ensemble of models includes various plausible hypotheses which characterize the structural uncertainty in the stock assessment,
- When the individual models making up the ensemble are a random sample of possible models from the population of models, the predictions (point estimates) will be better.
- ABC recommendations may be more stable, once an ensemble is implemented, because reactions of a single model will be dampened by other models.
- Some models in an ensemble may be better predictors of the stock status and appropriate catch levels under some states of the population (e.g., high abundance vs. low abundance) than others, so allowing model weights to vary over time should provide better predictions.

The working group identified the following disadvantages to ensemble modelling.

- Vetting and reviewing multiple models will add workload to the analyst and reviewers.
- Less transparency may result due to a more complex assessment.
- Choosing and weighting models may be difficult and there is currently no defined process or guidance for doing this.

How to consider within NPFMC system

The current system appears to allow for an ensemble model for providing advice for ABC/OFL specifications. Ensemble models may be most appropriate as a strategic evaluation tool (e.g., as a research model to evaluate the current model and Tier system). It also could provide better point estimates (as has been born out in the literature) and test the robustness of a simpler, more transparent model used for management.

Recommendations

1. Assuming that some sort of model averaging is involved, an ensemble model should be treated the same as any other model (i.e., an ensemble is a "model" and should be treated as such in reference to the existing language in the FMP and SAFE report guidelines).

2. Continue efforts on ensemble modeling, including approaches that could be used in this year's assessment cycle.
3. Resolve the following critical issues::
    1. Choosing and justifying members of the ensemble model
    2. Choosing among a number of available weighting schemes
    3. Justify the benefits of the added complexity resulting from moving to an ensemble model
4. Identify criteria for stocks amenable to ensemble modeling (e.g., fully-exploited, high model result variability).
5. BS P cod and Northern rocksole and/or YFS assessments should move forward with ensemble modeling options in the upcoming assessment cycle
6. Ensemble modeling seems appropriate for consideration in some NPFMC assessments but not necessarily for all assessments.
7. For example, a good use of an ensemble model (at high levels of inclusion and complexity) would be to test current assessment methods and harvest control rules. This would help with a) supporting a simple model for management purposes by showing that it compares favorably with the ensemble and b) improving transparency and alleviating review and model selection process at the Plan Team/SSC meetings.
8. Candidate stocks for an ensemble model should be chosen judiciously because it will add significant workload to both assessment authors and reviewers.
9. The process may need to be modified to allow for adequate review of model selections and weighting schemes (e.g, a CIE review may be required or additional Plan Team meeting for model selection)
10. Selection of models for the ensemble should be made no later than the September/October time frame and preferably earlier.
11. If the SSC wishes to entertain ensemble models, they may need to devote more time for model review (e.g., during the February meeting).

Part 2 Determining ABC

Review how maxABC and ABC are determined in the NPFMC system

Diana Stram provided an overview of how maxABC and ABC are determined in the NPFMC Tier system for groundfish and for BSAI crab under the federal fishery management plans.

Examples of reductions from maxABC in the past

Dana Hanselman provided an overview of instances where ABC was set below the maximum permissible for groundfish. The ABC has been reduced below the maxABC on multiple instances. The Team and SSC both have the ability to recommend such reductions.

Diana Stram provided examples from BSAI crab and related rationale for reducing stocks below ABC since 2011. The group discussed issues of internal consistency and potentially considering a framework for reducing below maxABC.

How can ensemble modeling inform maxABC and ABC

Allan Hicks provided an overview of considerations in ensemble modeling that can help to inform appropriate reductions from the maximum permissible. The group discussed use of the ensemble to inform a value below maxABC or if it is better to just move to an ensemble approach to calculate maxABC from the ensemble. As an interim step it was suggested to bring side-by-side approaches of single models and an ensemble model. Issues were raised with respect to consistency as compared with flexibility.

One approach suggested was as follows: 1) estimate the uncertainty in OFL from the single "best" model, 2) compute the P* that corresponds to the estimated uncertainty and point estimate of maxABC, then 3) apply that P* to the uncertainty estimated by the ensemble model to obtain the recommended ABC.

Participants noted that if the ensemble does not include ecosystem models then it is not actually capturing full uncertainty. The list of example factors in the FMPs that might warrant setting ABC below maxABC (data uncertainty, recruitment variability, and declining population trend) could be used to help set up an ensemble that captures the range of uncertainties.

Other methods of accounting for uncertainty when determining ABC

Grant Thompson and Dana Hanselman led a discussion on other methods of accounting for uncertainty when determining ABC. There was discussion of the trade-offs between the P* and decision-theoretic approaches. There was some discussion of how EBS Pacific cod might have been managed differently if a P* approach had been in place.

Potential tools/metrics for guiding reductions

Alan Haynie led a discussion on potential tools and metrics for guiding reductions from maxABC. He reviewed current ACL carry-over provisions used in other regions. There was a proposal to add a section in the SAFE report that looks at whether circumstances warrant explicit considerations for management action. This could be an addition to the introduction chapter or in the individual assessment. This could include a section on pros and cons of catching the maxABC for the stock. On the other hand, it was also noted that the SAFE chapter guidelines already require each chapter to include "discussion of information and rationale, if any, that might warrant setting ABC below the maximum permissible level." Such information could also be coordinated with the stock-specific ESR considerations.

The role of ecosystem or socio-economic considerations in reductions from maxABC

Alan Haynie led a discussion on the role of ecosystem and socio-economic considerations in reducing below maxABC. What types of economic or socio-economic considerations should be included in ABC setting versus TAC setting? Some suggestions included adding a section in each assessment under "Harvest Recommendations" describing whether economic and socio-economic considerations are, or should be, accounted for in ABC or TAC setting; and consideration of some inter-annual flexibility in the maxABC control rule with exploration of how much such flexibility would matter in different stocks. Conceptually, ABC should move toward OY from MSY and thus could be more informed by socio-economic considerations. While it was noted that some of these factors may be more appropriate for the TAC-setting process, it was suggested that there are other factors, such as when lower catch of smaller fish now will result in higher values for larger fish in the future (e.g., sablefish), that should be included a discussion of possible reasons to reduce the ABC from maxABC.

Wrap-up discussion

Alternative approaches to recommending ABC<maxABC include the following:

1. Maximum flexibility—keep current case-by-case procedure
2. Less flexibility—choose one of the following:
    a. Develop a set of rules based on rigorous scientific analysis; or
    b. Develop a set of rules based on scientists' best collective judgment, which may stop short of rigorous scientific analysis; or

c. Develop a list of things to consider without specific rules as to how they are applied.

The group discussed formulating rules for reducing from maxABC as compared to a more flexible approach. A committee (chaired by Martin Dorn) was formed to develop rules/defaults for reducing the ABC from the maxABC, with the expectation that the committee would provide a report at the September Joint Team meeting. Some recommendations for this committee are contained in the following "Recommendations" section.

Recommendations

The SSC recommended "identification of clear and transparent rules for defining the specific criteria to be used when adjusting the recommended ABC. Stock assessment uncertainty relative to levels upon which the Tier system was constructed, atypical data availability or usage (e.g., reliance on only catch-per-unit-effort vs. a survey index), ecosystem considerations, and other factors are potential candidates."

The workshop's recommendations are as follow:

1. Include a section in the Introduction to the SAFE report outlining extraordinary circumstances and major uncertainties which should feature discussion of:
   - Who will be impacted by choosing an ABC below the maxABC?
   - What are the current hypotheses related to how this extraordinary circumstance has impacted the stock and what are the current research priorities?
   - What data can be collected to evaluate these hypotheses?
2. Any reductions of ABC should be transparent and clearly described.
3. Clarify, with the SSC, the issue of the extremely high bar set for reducing the ABC for EBS P. cod.
4. The committee charged with developing rules for setting ABC<maxABC (chaired by Martin Dorn) should consider:
   a. Elements to include (e.g., ecosystem indicators, uncertainty in data, trend in stock status, missing surveys)
   b. Specific reductions (defining % reduction)
   c. Setting ABC by using the maxABC a Tier other than that used for setting OFL
   d. Using a different model than the chosen assessment model to justify a reduction (may be an alternative single-species model, an ensemble model, or a multispecies model)
5. The Joint Teams should recommend that AFSC task staff to continue to work on P* and decision theory approaches to develop uncertainty-based buffers. Two ideas are:
   e. Update the previous analysis using survey uncertainty to define the uncertainty to consider in a P* approach
   f. Determine the P* implied by a single "best model" approach and determine how different the buffer would be when using that P* with an ensemble approach.
6. Biologists, economists, and other social scientists should spend more time together discussing how socioeconomic factors are relevant to stock assessment and how changes in abundance, size, and distribution impact fishers, communities, and consumers.
7. Further investigate the impact of TAC reductions from ABC for different species. For example, when pollock ABC is lower, this can lead to a significant increase in flatfish species TAC and catch.

Attachment 1. BSAI Groundfish Plan Team (convenors) Assessment Methods Workshop Agenda

TIME	TOPIC	LEAD
	Wednesday June 27	
9:00 AM	1. INTRODUCTIONS	Co-chairs
9:15 AM	2. PURPOSE OF WORKSHOP	Co-chairs
	3. ENSEMBLE MODELING	
9:30 AM	Brief descriptions of ensemble modeling and model averaging	Co-chairs
10:00 AM	The dividing line between statistics and machine learning	Thompson
10:15 AM	3.3. Examples of ensemble modeling in fisheries stock assessment: the American experience	Hicks
10:45 AM	BREAK	
11:00 AM	3.4. Examples of ensemble modeling in fisheries stock assessment: the ICES experience	Johnson
11:30 AM	3.5. Examples of ensemble modeling in other disciplines	Bond
12:00 PM	LUNCH	
1:00 PM	3.6. Lessons from the 1998 NRC study	Thompson
1:15 PM	3.7. Review the 2017 SSC ensemble modeling workshop	Hicks
1:45 PM	3.8. Review the NSAW on ensemble modeling	Hanselman
2:15 PM	3.9. Choosing models in an ensemble	Thompson
3:00 PM	BREAK	
3:15 PM	3.10. Combining models and assigning weights	Thompson
4:00 PM	3.11. Calculating statistics and uncertainty	Thompson
4:30 PM	3.12. Pros and cons of implementation in NPFMC system	Thompson/Ianelli
	Thursday June 28	
9:00 AM	3.13. Communicating and using results	Co-chairs
9:15 AM	3.14. Workload and logistics for assessment authors	Thompson
9:30 AM	Identifying assessments amenable to ensemble modeling	Co-chairs
09:45 AM	Summarize, discuss, and make preliminary recommendations	Co-chairs
10:15 AM	BREAK	
	4. DETERMINING ABC	
10:30 AM	Review how maxABC and ABC are determined in NPFMC system	Stram
10:45 AM	4.2. Examples of reductions from maxABC in the past	Hanselman
11:15 AM	4.3. How can ensemble modeling inform maxABC and ABC	Hicks
12:00 PM	LUNCH	
1:00 PM	4.4. Other methods of accounting for uncertainty when determining ABC	Thompson/Hanselman
1:15 PM	4.5. Potential tools/metrics for guiding reductions	Haynie
1:45 PM	The role of ecosystem or socio-economic considerations in reductions from maxABC	Haynie
3:00 PM	BREAK	
3:15 PM	Summarize, discuss, and make preliminary recommendations	Co-chairs
	5. REPORT	
4:00 PM	Draft final recommendations to bring to the Plan Teams	Co-chairs
4:45 PM	5.2. Items for the agenda of the September Plan Team meeting and other preparations	Co-chairs
5:00 PM	ADJOURN	