Independent peer review report --- 2021 stock assessment for Eastern Bering Sea Pacific cod (*Gadus macrocephalus*)

Yan Jiao

Prepared for

Center for Independent Experts

June 2021

Contents

Executive Summary

Activities

The 2021 stock assessment of the Eastern Bering Sea (EBS) Pacific cod (*Gadus macrocephalus*) was reviewed by a Center for Independent Experts (CIE) stock assessment review panel. The review panel aims to review Pacific cod stock assessment documents and to produce a summary panel report that can be used by the North Pacific Fishery Management Council (NPFMC) and other interested persons for developing management recommendations for the Pacific cod fishery. The review also aims to ensure that the stock assessment represents the best available science to date and that any deficiencies are identified and addressed. The review took place virtually on Google Meet from April 26-30, 2021. The stock assessment done by the Alaska Fisheries Science Center (AFSC) stock assessment team was presented publicly to the review panel and the validity of the data, assessment procedures, and the recommended base model or model ensemble and alternative model scenarios were discussed. All the models were processed using Stock Synthesis (SS3.30.15), and the likelihood approach was used to estimate parameters (Methot et al. 2020). The AFSC assessment team, Dr. Grant Thompson, provided all the background information, documents, and further data and model configuration explorations that were requested by the CIE review panel.

According to the Terms of Reference (TOR) for the peer review, the CIE review panel prioritized the six topics first and then identified at least one recommendation per topic to be addressed by the review. During the review process, there were some changes on the priorities and specific TORs to be discussed, and not all the TORs were adequately discussed. Based on the review, the review panel Chair and the panel members provided a summary panel report with statements on the recommended model or model ensemble for management considerations. The panel summary report also suggested future potential efforts for new tag studies and alternative model structures with specific focuses on ensemble models, movement considerations, catchability modeling, and fishery-dependent CPUE in the stock assessment.

Main review processes and findings

The recommended pre-review base model M19.12a by AFSC in the draft report is an integrated age-structured stock assessment with constant catchability, combined survey data analysis of Eastern and Northern Bering Sea, time-varying logistic selectivity and not to use the fishery CPUE. Besides this base model, there have been many other models developed in the past, and several new ones are included in the draft report (Thompson et al. 2020). All the models have a time series from 1977 to 2020 and utilized one fishery fleet, and one or two fishery-independent survey CPUEs and survey age and size compositions (EBS+ NBS (Northern Bering Sea) combined trawl surveys, or EBS and NBS separate trawl surveys) to calibrate population dynamics. There have been problems in fitting the fishery catch composition, so only fishery catch size composition was used in all the models under consideration but this was discussed in

one of the TORs. The parameters of weight-length relationships inter-annually, the maturity at length, the ageing uncertainty measurement of the standard deviation of estimated age between “reader” and “tester”, and the stock-recruitment steepness were estimated outside of the models and were fixed in the models.

There were alternative interests in the hypotheses of time-varying catchability, dome-shaped selectivity, whether the EBS and NBS surveys should be united or separated when calibrating the population trends, whether fishery-CPUE should be used, and the low CV of the survey abundance from VAST. The CIE spent lots of time on the model ensemble approach which includes both the models to be considered and model weighting. With the consideration of equal weighting the alternative hypothesis, the CIE panel suggested that besides the based model, there are another 4 models to be explored and considered in the ensemble approach. These 4 alternative models are: M19.12 (based model + time-varying catchability (q)), M20.8a (base model + allowing dome-shaped selectivity), M20.9a (base model + also using fishery CPUE), and M21.cie (base model + using estimated survey CV). The review panel suggested not to use time-varying q when using the alternative models after diagnosing the model M19.12, which showed confounding among time-varying q, time-varying selectivity parameters, and the estimated natural mortality shown through retrospective analysis. So, all the alternative models were based on the base model M19.12a, i.e., constant q. A series of explorations on alternative model assumptions and retrospective analysis were requested. Such requests helped the AFSC assessment team and CIE review panel to recommend the base model and models to be considered in the ensemble to be used for management purposes and future research recommendations.

Although there are quite some concerns and interests on the fishery age composition data, fishery CPUE analysis, and movement modeling based on the abundance survey data, there was not enough time to discuss all these topics and questions in great depth. The review panel recommended an ensemble approach that may be considered in the future assessment update for management purposes.

Given the data available and the stock assessment developed by the assessment team, I support the recommended model ensemble as the best available science and its projected biomass for management consideration.

Main recommendations

There are no disagreements on comments and recommendations between the CIE panel and me, but the recommendations represent my views. Below I include both the major recommendations that I agree with the CIE panel and extra comments and recommendations from myself.

- The ensemble included models with alternative hypotheses or data scenarios. Results from both the ensemble and each model scenario should be considered or digested when considering the stock status and management considerations. The ensemble tends to be more robust to the model selection uncertainty and

low retrospective error. Future diagnostics on its performance may be done further based on the Pacific cod data and models included. The cross-conditional model averaging approach (Thompson 2021) provided as one of the documents for CIE was not applied to the Pacific cod case, but may be considered in the future.

- The current stock assessment models all assumed time-varying selectivity for both the survey and the fishery. The current models combined all the different fisheries (trawl, hook-and-lines, pot and jig gears) as one fleet. Reasons for selectivity to change may link to fleet composition changes, differential age and/or size spatial distributions of the population or their availability to the survey gear or fishing fleets. I suggest the estimated selectivity changes over time may be diagnosed to explore its potential reasons for changes over time and whether the estimated changes are reasonable or not.

- The model weighting categories may be further discussed or to include expert opinions from the teams in charge of surveys, population ecology, and stock assessment. The model weighting categories and weights developed during the review week were based on the documents provided and the models reviewed, which may change when new concerns or new model assumptions and data are considered in the future.

- The fishery CPUE analysis can be difficult when multiple fisheries are considered together. There are two methodologies provided in the assessment reports but they are not fully comparable because the data used are not the same. The VAST model did not provide enough details and it only uses one type of fishery data (hook-and-line) and only used two months of data. I feel that a model-based approach is necessary but the rationale of only using two-month data and only use one fishery need to be addressed. The fishery has a clear monthly pattern and likely important to be considered. Also, the first two months of hook-and-line fishery may face gear saturation and it is unclear whether such factors were considered or not. The review panel felt that the model/approach needs to be further developed with details before being used. I recommend a hierarchical Bayesian model with the fishery fleet as a random effect when considering multiple fisheries in the fishery CPUE standardization. Such a model builds on widely used fishery CPUE standardization models but with hierarchies to consider the reality of multiple fleets in the data analysis.

- The estimated natural mortality is not that robust in the retrospective analysis, which causes concern of potential natural mortality changes over time. Future studies may look into it.

- Dome-shaped selectivity was used in one model scenario M20.8a. The scientists in charge of the trawl survey and previous studies do not seem to suggest it (Weinberg et al. 2016). The review panel did suggest M20.8a be added as one alternative model in the ensemble.

- During the assessment model period, 1977-2020, the environmental or climate indices did not show the unusual cold pool pattern or the fish northward shift (Stevenson and Lauth 2019) pattern before the mid-2010s because of lacking historical surveys, which makes the assessment model difficult to capture the most recent “new” movement pattern. A longer time series analysis retro the environmental or climate ocean oscillation indices back to the past may help diagnose whether the temperature and water body anomalies observed in the recent 3-5 years were observed historically. Such analysis should shed light on the future modeling of Pacific cod using environmental indices to inform movement or spatial distribution changes.

- The Pacific cod movement is complicated as it includes not only the most recently newly observed movement between EBS and NBS areas, but also spawning movement and movement into the Russian water, and the differential age/size-specific depth movement or preference (Shimada and Kimura 1994; Rand et al. 2014). I suggest a continued satellite tagging study to better understand its seasonal movement patterns and the relationship with local and large climate variations. I agree with the decision of using the EBS and NBS combined survey data as one index to calibrate population dynamic, which is supported by both the genetic studies and a satellite tagging study presented during the review (Spies et al. 2020; Nielsen et al. 2021 presentation). I also recommend an approach used in Jiao et al. (2016) for future consideration, in which spatial asynchrony was considered, and the area-specific abundance indices can be used to calibrate such process uncertainty.

- I recommend the Joint Groundfish Plan Team, NPFMC, and its SSC take into account the uncertainties listed in the findings, the extra sensitivity runs done during the review week, the model runs included in the ensemble when considering management decisions for EBS Pacific cod.

1. BACKGROUND

The 2021 stock assessment of the Eastern Bering Sea (EBS) Pacific cod (*Gadus macrocephalus*) was reviewed by a CIE stock assessment review panel. The panel is expected to review the EBS Pacific cod stock assessment and to produce a panel report that can be used by the North Pacific Fishery Management Council (NPFMC) and other interested persons for developing management recommendations for the Pacific cod fishery in EBS. The review also aims to ensure that the stock assessment represents the best available science to date and that any deficiencies are identified and addressed. The stock assessment team, led by Dr. Grant Thompson from Alaska Fisheries Science Center (AFSC), contacted the review panel and provided the draft assessment report, related documents online about 2 weeks before the review. The review took place virtually on Google Meet from April 26-30, 2021. The virtual meeting is available to the public. The review panel chair is Dr. Ingrid Spies, and the other panel members include Drs. Henrik Sparholt, Arni Magnusson and Yan Jiao (me).

The EBS Pacific cod review process was coordinated by Dr. Ingrid Spies. The stock assessment background documents were prepared and were presented at the meeting. Several scientists presented material on survey, ageing, observer program, catch accounting system, population structure studies based on genetics and tagging, new methodology development on age composition and movement studies, fishery CPUE analysis, and ecosystem and socioeconomic profiles (see tentative agenda in Appendix 2).

According to the CIE scope description, "… *Each CIE reviewer shall conduct the independent peer review in accordance with the requirements specified in the PWS, OMB guidelines, and Terms of Reference (TORs), in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus. Each reviewer should assist the Chair of the meeting with contributions to the summary report. Deliver their reports to the Government according to the specified milestones dates.*" As a review panel member, I was provided with a draft stock assessment report and web access to relevant files and documents, including 10 PPTs with recorded presentations (see Appendix 1 for a full list of documents), and participated in the Stock Assessment Review Meeting. The review process followed the tentative agenda and the TORs. Instead of reviewing the validity of the data, assessment models and procedures, and results directly, we prioritized TORs and then discussed them. However, during the discussion of these TORs, the validity of the data, assessment models, and results were discussed.

According to the TORs for the peer review, the CIE review panel prioritized the six topics first and then identified at least one recommendation per topic to be addressed by the review. During the review process, there were some changes on the priorities and specific TORs to be discussed, and not all the TORs were adequately discussed. Based on the review, the review panel Chair and the panel members provided a panel summary report with statements on the recommended model or model ensemble scenario for management considerations. The panel summary report also suggested

future potential efforts for new data collection, and alternative model structures with specific focuses on ensemble models, movement considerations, catchability modeling, and fishery-dependent CPUE in the stock assessment.

Extra documents and model runs were provided upon requests from the CIE review panel. Discussions on the appropriateness of the model assumptions, equations, model weighting in the ensemble approach, quality of the data including the data standardization or synthesis, parameterizations, model fitting were made throughout the review.

During the review meeting, the AFSC assessment team Dr, Grant Thompson, Steve Barbeaus, and Jason Conner were always available when required for further discussion, additional model exploration and clarification, and clarification of how each TOR was addressed.

As a CIE reviewer, my duty was to evaluate the stock assessments of EBS Pacific cod with respect to their TORs (in Appendix 2), and work with the CIE review panel Chair and other members to prepare a panel summary report. This report provided the findings and recommendations of the independent review that is undertaken by me following the CIE Performance Work Statement (PWS).

2. ROLE OF INDIVIDUAL REVIEWER IN THE REVIEW ACTIVITIES

My role as a CIE independent reviewer was to conduct an impartial and independent peer review in accordance with the PWS and the predefined TORs herein.

About two weeks before the review meeting, the assessment documents and supporting materials were made available to the review panel via website by Dr. Grant Thompson. I read all the documents that I received before the review.

The EBS Pacific cod 2021 peer review meeting followed the “Tentative agenda (Appendix 2)” of the CIE review. The meeting was open to the public and was organized constructively. On the morning of April 26 before the meeting, the AFSC assessment team and CIE panel met online through Google meet to introduce each other, discuss the meeting agenda and CIE review process, reporting requirements, and meeting logistics. During the meeting, all the documents were accessible online through Google drive or emails.

Presentations were given during the review according to the agenda to provide the CIE panel the background information on the following: AFSC EBS and NBS trawl survey; the observer program; the ageing method in assessing the ageing uncertainty and how it was used in the assessment model; the catch accounting syste; the population structure studies based on genetics and tagging; new methodology development on age composition and movement studies; fishery CPUE analysis; VAST models in fishery CPUE and survey abundance estimation; model development on the movement of Pacific cod based on survey data; and ecosystem and socioeconomic profiles; and

NPFMC's management control rule. The assessment team and the review panel then moved to the TOR prioritization and discussion accordingly. I was actively involved in the discussion during the presentations by 1) listening to the presentations carefully, making notes on the points that were not included or not clearly stated in the documents provided prior to the meeting; 2) asking questions for clarification on the data usage and model development; 3) making comments and providing possible alternative solutions to questions arising during the meeting; 4) discussing agreements on each model scenario and TOR with the other review panel members.

On the last day of the peer review meeting, CIE review panel chair Dr. Ingrid Spies discussed how to prepare the panel summary report together, which summarized the panel's views, requests, and conclusions; all panel members contributed to it. This review report is formatted according to my interpretation of the required format and content described in Appendix 2.

3. SUMMARY OF FINDINGS RELATIVE TO TORs

EBS Pacific cod stock is among the most commercially important stocks in the U.S. EEZ. There have been many model scenarios developed and changed over time. This review focused on the prioritized TORs and during the discussion of these TORs, the validity of the data, assessment models and results were discussed. The recommended pre-review base model M19.12a is an age-structured model with survey catchability constant over time, one fishery fleet with all the fisheries catch and size compositions combined, fishery and survey selectivity were assumed logistic and time varying, combined survey data analysis of Eastern and Northern Bering Sea is used, age and size compositions from the survey analyzed using VAST and the fishery CPUE not used. Besides this base model, there have been many other models developed in the past and several new ones included in the draft report (Thompson et al. 2020). All the models have a time series from 1977 to 2020 and utilized one fishery fleet and one or two fishery-independent survey CPUEs and survey age and size compositions (EBS+NBS trawl surveys, or EBS and NBS separate trawl surveys) to calibrate population dynamics. There have been problems in fitting the fishery catch composition, so only fishery catch size composition data were used in all the models under consideration but this was discussed in one of the TORs. There have been attempts to estimate the movement of the stock from EBS to NBS but the assessment models considered did not include such a model scenario because of limited data available from NBS surveys and from tagging studies. This was discussed in one of the TORs also.

According to the TORs, "*The TORs were compiled from recommendations submitted by the Groundfish Plan Team for the Bering Sea and Aleutian Islands, the Scientific and Statistical Committee, and Alistair Dunn (a consultant contracted by the Freezer Longline Coalition). These were organized into six general topics, with three specific recommendations per topic. After reading the background materials and receiving the initial set of presentations during the review, the reviewers will prioritize the six topics and identify at least one recommendation per topic to be addressed by the review. The*

reviewers will then address as many of the topics (and the identified recommendation(s)), in priority order, as time allows." The CIE review panel prioritized the six topics first and then identified at least one recommendation per topic to be addressed by the review. During the review process, there were some changes on the priorities and specific TORs to be discussed, and not all the TORs were adequately discussed.

Below I provide the summary of findings for the EBS Pacific cod stock assessment review, in which the weaknesses and strengths are described in accordance with the TORs prioritized and adopted by the review panel.

3.1. *Ensemble modeling (originally labeled as "Topic 2")*

This TOR was addressed adequately in general, although further exploration and improvement are suggested.

Recommendation 2a:
Evaluate the use of ensemble modeling in the NPFMC management system, and specifically whether the structural uncertainty and historical challenges in identifying a robust base model make Pacific cod a good application for ensemble modeling.

Under the discussion for Topic 2b, the review panel agreed that the ensemble approach can be used to deal with situations with structural uncertainty, alternative hypothesis on key parameters, and the application of alternative datasets. The review panel suggested to investigate and compare the model goodness of fit (negative loglikelihood and # of parameters used) and retrospective error for each model considered in the ensemble, and then compared the performance of the ensemble with the base model only, by looking at their estimated key values of management interests and their retrospective values. The key parameters and values compared include: SSB, R, F, F_{msy}, SSB_{msy}, F/F_{msy}, SSB/SSB_{msy}, and recommended ABC when each approach was used (Table 1). The weights used in the last column of Table 1 are 0.254, 0.222, 0.164, 0.106 and 0.254, not fully the same as the ones used in Table 2 in which the weights were further modified after diagnosing the fit of model M21.cie. Because of time limitations, the CIE review panel did not ask for an update of this table, but it would only influence the last column slightly. For an updated stock assessment report, tables 1 and 2 should be updated.

The CIE review panel and the assessment team considered an ensemble model averaging criteria and average emphasis used previously (Table 2.20 in the draft assessment report), and further modified it based on the new ensemble coming out from the Topic 2b below. The new table of model weight computation is now shown as Table 2.

The review panel balances the criteria needed to be considered in general and the criteria important for Pacific cod. For the criteria when all the model weights are the same, the weight emphasis is treated as 0 to avoid resulting in equal weights for all models in the ensemble (See the criteria that are highlighted in red). The current weighting is based on the rank of the three reviewers. For future usage of this approach, it should be reasonable for AFSC and SSC members to participate in the weighting surveys.

The review panel and the assessment team also discussed the cross-conditional model averaging approach developed by Thompson (2021). Reviewers generally demonstrated interest in this approach. However, it is not possible to try this approach based on the EBS Pacific cod case in a short time. So, this approach was not further discussed in great detail during the review.

Recommendation 2b:
Develop the models to include in an ensemble.

Topic 2b was discussed before Topic 2a based on the reality that it is important to discuss which models to be included in an ensemble first before considering the weighting strategy. The review panel considered the alternative hypothesis on the parameters of catchability and selectivity, whether or not to use fishery CPUE, and the CV of VAST survey abundance. The review panel also considered how to avoid overweighting one type of hypothesis. For example, the original factor design table on models to be considered used the time-varying q when further considering including fishery CPUE, or dome-shaped selectivity. The review panel felt that by doing so, 4 models considered time-varying q but only one model used constant q, so the hypothesis of time-varying q is automatically overweighted. Unless there is strong evidence of time-varying q, such a situation is suggested to be avoided. The review panel suggested not to use time-varying q when using the alternative models after diagnosing the model M19.12, which showed confounding among time-varying q, time-varying selectivity parameters, and the estimated natural mortality shown through retrospective analysis. To avoid overweighting time-varying q, four new models were requested based on the base model M19.12a which has a constant q (see Table 2).

Among the four new models, they are M19.12 (based model + time-varying catchability), M20.8a (base model + allowing dome-shaped survey selectivity), M20.9a (base model + also using fishery CPUE), and M21.cie (base model + using estimated survey CV). The review panel requested a series of explorations on alternative model assumptions and retrospective analysis before the five models were selected in the ensemble. Such requests and quick responses from the AFSC assessment team helped the assessment team and CIE review panel to recommend the base model and models to be considered in the ensemble to be used for management purposes and future research recommendations.

For model M20.9a the fishery CPUE used in the assessment is from the VAST analysis. The review panel has concerns about the VAST fishery CPUE analysis which used only two months of hook-and-line fishery data. The analysis is not fully comparable with the catch-weighted CPUE approach which used all the gear types and all the months of each gear type. Discussion on the TOR of fishery CPUE below addressed this issue. For model M21.cie which was suggested after diagnosing the model fitting to the survey abundance in M19.12a. The review panel feels that the CV (=0.06) of VAST estimated survey abundance is much lower than known fisheries surveys. One panel member then suggested this model scenario to estimate the CV of survey abundance instead of using the estimated CV from VAST.

Given the data available and the stock assessment developed by the assessment team, I support the recommended model ensemble as the best available science and its projected biomass for management consideration. Extra comments on the model scenarios included or data used were included here also.

Recommendation 2c:
Consider whether to apply the sloping harvest control rule before or after ensemble averaging of SSB and other reference points.

Topic 2c was discussed after Topics 2b and 2a. Dr. Thompson gave a very good presentation on the “before” and “after” ensemble averaging of parameters or estimates of interests. The review panel discussed the “before” and “after” approaches based on the differences of the approaches, the complexity of the algorithms, and the possibility of applying the approach using the existing SS3 functions.

The “before” approach averages SSB and reference points from each model in the ensemble, so easy to perform. The “after” approach requires generating a “new” model with averaged parameters, so computationally can be more complicated. I generally support the “after” approach. When Bayesian approaches are used, the computation of the “after” approach can be done by resampling the posterior runs of the parameters including the estimated F, and project into the next year etc., so not impossible. Bootstrap algorithm can be used to reach the goal also. After discussion with the assessment team, such an approach is currently not available in the SS3 existing functions but may be considered outside of the SS3 computation.

The assessment team lead, Dr. Thompson, provided results using the “before” and “after” approaches (Table 3). However, in the “after” approach, unlike in the "before" approach, F_{2021} is assumed as a constant rather than a function of internally estimated parameters, and so has zero standard deviation. The estimated standard deviation of the ensemble ABC in the "after" approach is smaller because of this. If computation outside of SS3 is not available in a short time, the average uncertainty of F_{2021} from each model may be considered as an

approximate measurement of F_{2021} in the "after" approach when computing ABC in the future.

3.2. *Movement (originally labeled as "Topic 1")*

This TOR was addressed adequately in general although further exploration and documentation are suggested.

Recommendation 1a:
Comment on avenues for incorporating spatial dynamics and movement.

Topic 1a was discussed first by the CIE review panel. The review panel found the satellite tag experiment and model developed very informative and may be further continued if possible. The satellite tagging study and the genetic study suggested that the EBS and NBS Pacific cod is appropriate to be managed as one stock but there is a seasonal movement of the individuals and how the movement rate may change given environmental factors or age groups needs further studies (Rand et al. 2014; Spies et al. 2020; Nielsen et al. 2021 presentation). The studies also suggest that the EBS cod may move to the Russian water and further communication or data exchange with the Russian fisheries management agency should help future studies on movement or changes in the spatial distribution of Pacific cod.

The movement study based on the survey data is not in detail, so it is difficult to evaluate or comment on. There is only a presentation provided for the advection-diffusion-taxis (ADT) modeling but only one slide on the taxis rate modeling provided and no explanations on movement rate and advection rate modeling. I think such a study seems creative but not clear on what kind of environmental factors considered, the need for tagging data, and whether it is appropriate to Pacific cod specifically.

Because the satellite tagging study was only for about one year, and years with NBS survey are limited also, the data available for the potential to incorporate movement is limited, a simulation study to look into the influence of movement on the stock assessment, how the model ensemble or base model without considering movement may perform should help in a short time before further tagging data available. I would also recommend an approach developed in Jiao et al. (2016), in which the spatial asynchrony was considered, and the area-specific population abundance indices were used to calibrate it.

Recommendation 1b:
Consider how to inform the dynamics of movement or abundance between the Northern Bering Sea and the Eastern Bering Sea, specifically from additional experiments and analyses, data analyses that include these assumptions (i.e., VAST), and how these can best be used within the different models as indices of abundance.

This was addressed as an outgrowth of Topic 1a. The review panel was not able to evaluate the VAST or the ADT modeling approach because of lacking details on the model developed, data used, and results. See findings from Topic 1a.

Recommendation 1c:
Develop movement models.

This was addressed as an outgrowth of Topic 1a. The review panel questioned whether we really need movement models for a stock whose distribution is covered almost entirely by the EBS and NBS bottom trawl surveys. The panel felt understanding the degree to which the stock ranges into Russian waters is of great concern also. Overall, the panel recommends further tagging studies.

3.3. ***Fishery CPUE (originally labeled as "Topic 4")***

This TOR was discussed based on Topic 4c. Some suggestions were provided below.

Recommendation 4a:
Discuss standardization of fishery CPUE using alternative statistical methods, including a discussion of historical changes in the fishery that may affect the relationship of the index to abundance.

CIE review panel feels that this topic is almost the same as 4c, so mainly focused on the discussion on 4c.

Recommendation 4b:
Develop a fishery CPUE index.

CIE review panel all agreed that the development of an appropriate index is important, but it cannot be accomplished during this meeting. I agree with this recommendation.

Recommendation 4c:
Consider how best to further analyze CPUE, including development of spatio-temporal analyses of fleet specific CPUE indices that may help inform the model or supplement the trawl survey biomass indices.

The fishery CPUE analysis can be difficult when multiple fisheries are considered together. There are two methodologies provided in the assessment reports but they are not fully comparable because the data used are not the same. The VAST model did not provide enough details and it only uses one type of fishery data (hook-and-line) and only used two months of data (January and February). I feel that a model-based approach is necessary but the rationale of only using two-month data and only use one fishery need to be addressed. The fishery has a

clear monthly pattern and likely important to be considered. Also, the first two months (January and February) of hook-and-line fishery may face gear saturation and it is unclear whether such factors were considered or not. The review panel feels that the model and approach are promising but further details are needed.

I would recommend a hierarchical Bayesian model or a mixed effect model to be considered. In such approaches, the fishery fleet can be considered as random effect when considering multiple fisheries in the fishery CPUE standardization. Such kind of models build on widely used fishery CPUE standardization models but with hierarchies or random effects to consider the reality of multiple fleets in the data analysis.

The assessment team indicated the data quality issue with some fisheries lacking effort data. I feel fleet-specific fishery CPUE analysis for the fishery with better quality data, such as with larger spatial coverage, credible logbook records with spatial-temporal information, etc., is reasonable. The specific gear selectivity needs to be further considered since the current stock assessment models used one fleet that combines all the fishery types.

3.4. *Age data (originally labeled as "Topic 3")*

This TOR was not discussed in great depth. Some suggestions were provided below.

Recommendation 3a:
Attempt to resolve problems with using fishery age compositions.

This topic is not fully discussed because of time limitations. The review panel was only able to discuss it based on previous stock assessment results. Dr. Thompson provided a table with the loglikelihood and Mohn's rho estimations when both fishery size composition and age composition were used versus when only fishery size composition was used. The Mohn's rho is much higher when fishery age composition was used. I am wondering whether the number differences of age bin and size bin matters when using age composition and size compositions. For example, there are only 12 age groups but the size bin used in the model is 1cm, which implied that there are lots more size composition data to be fitted than that of age composition. The future diagnostics may include checking the fitting to age and size compositions to see which year and age groups that the model did not fit well and where the retrospective error mainly caused by; checking the model performance when using larger size bins to match the number of age groups in the age composition data. Overall, the assessment team has explored lots of model scenarios in the past (see appendix 2.3 of the draft assessment report).

Recommendation 3b:
Consider how best to include the fisheries age and size composition data, including consideration of fleet specific age composition data in the model.

Not selected for discussion.

Recommendation 3c:
Investigate whether a change in growth contributed to the ageing bias fit for 2008 and onward in the complex models as ageing bias and growth may be confounded.

This topic is not fully discussed because of time limitations. Some review panel members suggested that this may be diagnosed step by step. For example, one model scenario can be to turn off ageing bias and see what happens; another scenario can use the externally estimated growth with ageing uncertainty and see what happens.

3.5. ***Compositional data (originally labeled as "Topic 5")***

This TOR was not discussed in great depth. Some suggestions were provided below.

Recommendation 5a:
Consider methods (e.g., bootstrapping) to estimate uncertainty and variance in the composition data, with the results then used to estimate initial sample sizes for each season, fleet, combination for input into the assessment model.

Not selected for discussion.

Recommendation 5b:
Review methods to scale the composition data and include consideration of methods that scale observer samples to the catch by vessel, location, and time of event.

Not selected for discussion.

Recommendation 5c:
Consider analyses of the size- and age- composition data to identify if there are specific locations or time periods when a recruitment signal may be apparent to assist in informing the assessment model of the strength of recent recruitment.

This topic is not fully discussed because of time limitations. The review panel felt that this topic is meaningful and very useful for the assessment. The study may start from the age composition or size composition data to look into the overlap cross cohorts in the earlier age groups. The analysis may also look into the age or size groups with low or zero selectivity by the fishing gears but selected by the survey gears. Because the assessment models all used time-varying selectivity, it

may confound with the cohort signals to be estimated. An external analysis with plots such as bubble plots, etc., is encouraged for future external analysis.

3.6. ***Other (originally labeled as "Topic 6")***

This TOR was not discussed in great depth. Some suggestions were provided below.

Recommendation 6a:
Consider incorporation of dome-shaped survey selectivity.

The CIE review panel found that existing studies based on field data do not suggest that a dome-shaped survey selectivity is largely possible (Weinberg et al. 2016). To address this hypothesis, a model (M20.8a) with dome-shaped survey selectivity was suggested to be included in this year's models of the ensemble or simply as a sensitivity run.

Recommendation 6b:
Consider the diagnostic plots of fits and residuals (including normalised or Pearson residuals) for the age and size composition data and make recommendations on how the model fits may be improved.

The CIE review panel found this topic useful but did not discuss it because of time limitations.

Recommendation 6c:
Consider inclusion of other survey information (e.g., the IPHC and sablefish surveys).

This topic was not selected for discussion.

Although not listed in the original set of recommendations for the "other" category, the review panel suggested that consideration of density dependence in a variety of life history processes may be important in assessment models.

4. SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS RELATIVE TO TORs

The CIE review panel recommended the ensemble model after the discussion on the TORs and comparing all the model scenarios explored. Given the data available and the models developed and the concerns on the catchability, movement, dome-shaped survey selectivity, and fishery CPUE, I support the developed model ensemble as the best available science. Future updates on the models after fishery CPUE better analyzed is expected in one of the models in the ensemble.

My conclusions and recommendations are consistent with those from the CIE Panel. There is no obvious disagreement between the CIE panel and me on comments and recommendations. I here reorganize my recommendations based on each TORs without specifically listing the sub-TORs.

Recommendations for ensemble modelling

The ensemble included models with alternative hypotheses or data scenarios. Results from both the ensemble and each model scenario should be considered or digested when considering the stock status and management considerations. The ensemble tends to be more robust to the model selection uncertainty and low retrospective error. Future diagnostics on its performance may be done further based on the Pacific cod data. The cross-conditional model averaging approach (Thompson et al. 2021) provided as one of the documents for CIE was not applied to the Pacific cod case, but it may be considered in the future.

The current stock assessment models all assumed time-varying selectivity for both the survey and the fishery. The current models combined all the different fisheries (trawl, hook-and-lines, pot, and jig gears) as one fleet. Reasons for selectivity to change may link to fleet composition changes, differential age and/or size spatial distributions of the population or their availability to the survey gear and fishing fleets. I suggest the estimated selectivity changes over time may be diagnosed to explore its potential reasons for changes over time and whether the estimated changes are reasonable or not.

The model weighting categories may be further discussed or to include expert opinions from the teams in charge of surveys, population ecology, and stock assessment. The model weighting categories and weights developed during the review week were based on the documents provided and the models reviewed, which may change when new concerns or new model assumptions and data are considered in the future.

Recommendations for movement

During the assessment model period, 1977-2020, the environmental or climate indices did not show the unusual cold pool pattern or the fish northward shift pattern before, which makes the assessment model difficult to capture the most recent “new” movement pattern. A longer time series analysis retro the environmental or climate ocean oscillation indices back to the past may help diagnose whether the temperature and water body anomalies observed in the recent 3-5 years were observed historically. Such analysis should shed light on the future modeling of Pacific cod using environmental indices to inform movement or spatial distribution changes.

I suggest continued satellite tagging study to better understand its seasonal movement patterns and the relationship with local and large climate variations. I agree with the decision of using the EBS and NBS combined survey data as one index to calibrate population dynamic, which is supported by both the genetic studies and a satellite tagging study presented during the review (Spies et al. 2020; Nielsen et al. 2021 presentation). I also recommend an approach used in Jiao et al. (2016) for future consideration, in which the population dynamics process uncertainty considered the spatial asynchrony among areas, and the area-specific population abundance indices were used to calibrate it.

Recommendations for fishery CPUE

The fishery CPUE analysis can be difficult when multiple fisheries are considered together. I feel that a model-based approach is necessary but the rationale of only using 2-month data and only use one fishery need to be addressed when VAST is used. The fishery has a clear monthly pattern and likely important to be considered. Also, gear saturation may be considered when only use the January and February data. The model and approach seem promising but further details are needed. I also recommend a hierarchical Bayesian model with the fishery fleet as a random effect when considering multiple fisheries in the fishery CPUE standardization. Such a model builds on widely used fishery CPUE standardization models but with hierarchies to consider the reality of multiple fleets in the data analysis.

Fleet-specific fishery CPUE analysis for the fleet with better quality data is reasonable and may be developed and compared among fleet indices and with survey abundance index. The specific gear selectivity needs to be further considered since the current stock assessment models used one fleet that combines all the fishery types.

Recommendations for age data

This topic is not fully discussed because of time limitations. The retrospective pattern is much more obvious with much higher Mohn's rho when fishery age composition was used. I would recommend future diagnostics: checking the fitting to age and size compositions to see which year and age groups that the model did not fit well and where the retrospective error mainly caused by; checking the model performance when using larger size bins to match the number of age groups in the age composition data.

Recommendations for composition data

The review panel feels the topic "*Consider analyses of the size- and age-composition data to identify if there are specific locations or time periods when a recruitment signal may be apparent to assist in informing the assessment model of the strength of recent recruitment*" meaningful and very useful for the future

assessment. Because the assessment models all used time-varying selectivity, it may confound with the cohort signals to be estimated. An external analysis with specific focuses on the earlier age groups and using plots such as bubble plots etc., is encouraged for future external analysis.

Recommendations for other topics

Dome-shaped selectivity was used in one model scenario M20.8a. However, during the review, the staff in charge of the trawl survey and previous studies do not seem to suggest it. The review panel feels it is reasonable to add a model scenario as a sensitivity run and included in the ensemble.

5. Comments on the NMFS review process

I find the CIE review process effective, clear, and meaningful. This specific review done for Eastern Bring Sea Pacific cod was well organized both in the conduct of the meeting and in presentations of the assessment. The AFSC assessment team has been very patient and cooperative in dealing with requests and interpretation of the meaning or goal of the TORs. The extra requests on model scenarios and retrospective analysis or model goodness of fit diagnostics likely made them working overnight during the review. I originally found some documents not fully ready for review at the beginning such as the VAST analysis and the movement model analysis based on the survey data because of lacking details on the data and model for Pacific cod, but feel okay to evaluate it based on the general ideas of the models instead of the model details and results after the interpretation of the assessment team on the TORs. I have no further recommendations about the review process.

6. Acknowledgements

I would like to thank all the AFSC Stock Assessment Team members contributing to the meeting for their informative presentations on the background information of the stock assessments of EBS Pacific cod and for their hard work and willingness to provide helpful responses to the review panel's questions and requests. I also would like to thank the review panel Chair Dr. Ingrid Spies who coordinated the review process. Special thanks also go to other members of the review panel, Drs. Henrik Sparholt and Arni Magnusson for their respectful and productive discussions on the assessments and TORs.

7. References

Jiao, Y., O'Reilly, R., Smith, E., and Orth, D. 2016. Integrating spatial synchrony/asynchrony of population distribution into stock assessment models: a spatial hierarchical Bayesian statistical catch-at-age approach. ICES Journal of Marine Science, 73:1725-1738.

Methot, R.D., Wetzel, C.R., Taylor, I., and Doering, K. 2020. Stock Synthesis user manual. Model Version 3.30.15 (March 2020). NOAA Fisheries, Seattle, WA. 223 p.

Rand, K. M., P. Munro, S. K. Neidetcher, and D. G. Nichol. 2014. Observations of seasonal movement from a single tag release group of Pacific cod in the eastern Bering Sea. Marine and Coastal Fisheries, 6:287-296.

Shimada, A. M., and D. K. Kimura. 1994. Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. Fishery Bulletin, 92:800-816.

Spies, I., Gruenthal, K.M., Drinan, D.P., Hollowed, A.B., Stevenson, D.E., Tarpey, C.M. and Hauser, L., 2020. Genetic evidence of a northward range expansion in the eastern Bering Sea stock of Pacific cod. Evolutionary applications, 13:362-375.

Stevenson, D.E., and Lauth, R.R. 2019. Bottom trawl surveys in the northern Bering Sea indicate recent shifts in the distribution of marine species. Polar Biology, 42:407–421

Thompson, G. 2021 Cross-conditional model averaging: A potential tool for improving stock assessment estimates. Draft manuscript.

Thompson, G., Conner, J., Shotwell, K., Fissel, B., Hurst, T., Laurel, B., Rogers, L., Siddon, E. 2020. Assessment of the Pacific cod stock in the Eastern Bering Sea. (DRAFT). National Oceanic and Atmospheric Administration National Marine Fisheries Service. 344 p.

Weinberg, K.L., Yeung, C., Somerton, D.A., Thompson, G.G. and Ressler, P.H., 2016. Is the survey selectivity curve for Pacific cod (*Gadus macrocephalus*) dome-shaped? Direct evidence from trawl studies. Fishery Bulletin, 114:60-370.

Table 1: Mohn's Rho of key parameters compared and their values for each of the models in the ensemble. The weights used in the last column are 0.254, 0.222, 0.164, 0.106 and 0.254, not fully the same as the ones used in Table 2 in which the weights were further modified after diagnosing the fit of model M21.cie. (prepared by the assessment team).

	M19.12a	M19.12	M20.8a	M20.9a	M21.cie	Ensemble with equal weight	Ensemble with weights
SSB	-0.019	-0.017	0.053	0.106	-0.029	0.028	0.011
R	-0.132	-0.148	-0.276	0.008	-0.157	-0.136	-0.146
F	0.008	0.009	-0.037	-0.090	0.043	-0.007	0.005
SSB/SSB_{msy}	0.016	0.028	-0.044	0.153	0.008	0.038	0.025

Table 2: Factors in model scenarios considered for EBS Pacific cod, model weighting criteria considered and the computation of model weights. The current weighting is based on the rank of the three reviewers.

Factor 1: Allow *Q* to vary?		no	yes	no	no	no
Factor 2: Allow domed selex?		no	no	yes	no	no
Factor 3: Use fishery CPUE?		no	no	no	yes	no
Factor 4: Estimate survey CV?		no	no	no	no	yes
Criterion	Emph.	19.12a	19.12	20.8a	20.9a	21.cie
General plausibility of the model	3	2	1	0.6667	1	1.3333
Acceptable retrospective bias	3	2	2	1.3333	1	2
Uses properly vetted data	3	2	2	2	0	2
Acceptable residual patterns	3	2	2	2	2	1
Comparable complexity	2	2	1	1	2	2
Fits consistent with variances	2	1	2	1	0	2
Dev sigmas estimated appropriately	0					
Incremental changes	0					
Objective criterion for sample sizes	0					
Change in ageing criteria addressed	0					
Density dependence (other than R) addressed	0					
Regime shifts addressed	0					
Average emphasis:		0.9375	0.8438	0.6875	0.5000	0.8438
Model weight (Ensemble CIE):		0.2459	0.2213	0.1803	0.1311	0.2213

Table 3: Results from running each model with F_{2021} set by applying the HCR with model-specific parameters (Table 3a) and by applying the HCR with "average" parameters (Table 3b). The CV of ABC from 3a and 3b are 38% and 22% separately. In 3b, the standard deviation of F is treated as 0, so the overall CV of ABC is smaller.

3a

	Models	19_12a	19_12	20_8a	20_9a	21_cie	Ensemble
Quantity	Weight:	0.2459	0.2213	0.1803	0.1311	0.2213	1.0000
B_{2021}	Mean:	228219	210551	285785	212363	143142	213781
	Sdev:	18820	23753	30200	17851	33557	52541
$B_{40\%}$	Mean:	260965	265460	310114	264200	259533	270930
	Sdev:	6135	7727	14599	5624	9505	20680
$B/B_{40\%}$	Mean:	0.8651	0.8651	0.9226	0.7941	0.5475	**0.7959**
	Sdev:	0.0791	0.0791	0.0919	0.0727	0.1228	0.1651
$F_{40\%}$	Mean:	0.3494	0.3285	0.2856	0.3524	0.3254	**0.3284**
	Sdev:	0.0175	0.0185	0.0164	0.0178	0.0200	0.0291
F_{2021}	Mean:	0.3029	0.2570	0.2620	0.2796	0.1718	0.2533
	Sdev:	0.0396	0.0421	0.0366	0.0378	0.0469	0.0622
ABC	Mean:	123453	99310	128966	107922	47195	**100190**
	Sdev:	22621	24099	26092	20771	22096	**38127**

3b (Results from running each model with F_{2021} set by applying the HCR with "average" parameters)

	$B/B_{40\%}$:	**0.7959**	$F_{40\%}$:	**0.3284**	F_{2021}:	**0.2578**	
	Model:	19_12a	19_12	20_8a	20_9a	21_cie	Ensemble
Quantity	Weight:	0.2540	0.2222	0.1640	0.1058	0.2540	1.0000
F_{2021}	Mean:	**0.2546**	**0.2546**	**0.2546**	**0.2546**	**0.2546**	**0.2546**
	Sdev:	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ABC	Mean:	105716	98469	125643	99132	67891	**98471**
	Sdev:	8070	10480	11469	7808	15254	**21831**

Appendix 1: Bibliography of materials provided for review

Stock assessment report for review:

Thompson, G., Conner, J., Shotwell, K., Fissel, B., Hurst, T., Laurel, B., Rogers, L., Siddon, E. 2020. Assessment of the Pacific cod stock in the Eastern Bering Sea. (DRAFT). National Oceanic and Atmospheric Administration National Marine Fisheries Service. 344 p.

Background material provided for* EBS Pacific cod stock assessment review*:

Thompson, G. 2021 Cross-conditional model averaging: A potential tool for improving stock assessment estimates. Draft manuscript.

Spies, I., Gruenthal, K.M., Drinan, D.P., Hollowed, A.B., Stevenson, D.E., Tarpey, C.M. and Hauser, L., 2020. Genetic evidence of a northward range expansion in the eastern Bering Sea stock of Pacific cod. Evolutionary applications, 13(2), pp.362-375.

Presentations pre-recorded and posted on the website before the review:

1. Conner, J. AFSC Survey Estimates of Pacific Cod Abundance.
2. Stone, K and Anderl, D. Age Determination of Eastern Bering Sea Pacific cod at the Alaska Fisheries Science Center.
3. Kraski, J. North Pacific Observer Program, Alaska Fisheries Science Center Fisheries Monitoring and Analysis Division.
4. Furuness, M. Pacific cod Catch Estimation and Pacific cod Management in the Bering Sea - 2021.
5. Nielsen, J. et al. Seasonal movement and environmental conditions of satellite-tagged northern Bering Sea Pacific cod.
6. Correa, G.M. Impacts of temporal and spatial variability in somatic growth on fish stock assessment models.
7. Thorson, J. et al. High-resolution movement rates and habitat utilization from environmental variables, tags, fishery catch-and-effort, and resource surveys using advection-diffusion-taxis modelling.
8. Thorson, J. Standardizing fishery-dependent CPUE data using VAST.
9. Shotwell, K. Ecosystem and Socioeconomic Profile (ESP), EBS Pacific Cod.
10. Thompson, G. Assessment background as context for the Terms of Reference.

Additional presentations during the review:

Merrigan, G. Additional information on Bering Sea p-cod fisheries from the Freezer-Longline Coalition (FLC = Catcher-processor hook-and-line vessels).

O'Leary, C. Estimating spatiotemporal availability of transboundary fishes to fishery-independent surveys.

Additional materials provided during the review:

O'Leary, C., Kotwicki, S., Hoff, G., Thorson, J., Kulik, V., Ianelli, J., Lauth, R., Nichol, D., Conner, J., Punt, A. Estimating spatiotemporal availability of transboundary fishes to fishery-independent surveys. (DRAFT NOT FOR CIRCULATION).

Kimberly M. Rand, Susanne F. McDermott, David. R Bryan, Julie K. Nielsen, Ingrid B. Spies, Steve Barbeaux, and Grant Thompson. Comparison of fishery and survey length distributions of Alaskan Pacific cod (*Gadus macrocephalus*): is there a mismatch? (DRAFT NOT FOR CIRCULATION).

Appendix 2: Statement of Work

Performance Work Statement (PWS)

National Oceanic and Atmospheric Administration (NOAA)

National Marine Fisheries Service (NMFS)

Center for Independent Experts (CIE) Program

External Independent Peer Review

Virtual Panel Review of the Stock Assessment for Pacific Cod in the Eastern Bering Sea

April 26-30, 2021

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).

Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The fishery for Pacific cod in the Eastern Bering Sea is among the most commercially important in the U.S. EEZ. Recent developments of note include a substantial northward migration of the stock, to waters outside the area that has been surveyed annually by the NMFS Alaska Fisheries Science Center (AFSC) since 1982. Efforts at modeling this movement have been hampered by the scarcity of both survey data from the northern region and tagging data in general. Conflicts between fishery age composition data and the other data used in the assessment models also pose problems for the assessment. Ensemble modeling has been advocated as a potential solution to the problem of

structural uncertainty in the assessment models, but attempts to date have been mostly unsuccessful.

The goal of this review will be to ensure that the stock assessment represents the best available science to date and that any deficiencies are identified and addressed. The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (TORs) of the peer review are listed in **Annex 2**. Lastly, the tentative agenda of the virtual panel review meeting is attached in **Annex 3**.

Requirements
NMFS requires three (3) reviewers to conduct an impartial and independent peer review in accordance with the PWS, OMB guidelines, and the TORs below. The reviewers shall have a working knowledge of, and recent experience in, the following areas:

- The Stock Synthesis modeling framework;
- Movement (migration) models;
- Ensemble modeling (model averaging); and
- Federal fisheries science requirements under the Magnuson-Stevens Fishery Conservation and Management Act.

The chair, who is in addition to the three reviewers, will be provided by the AFSC; although the chair will be participating in this review, the chair's participation is not covered by this contract.

Tasks for Reviewers

1) Two weeks before the peer review, the NMFS Project Contact will send by electronic mail or make available at an FTP site to the CIE reviewer all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review.

2) Additionally, two weeks prior to the peer review, the CIE reviewers will participate in a test to confirm that they have the necessary technical (hardware, software, etc.) capabilities to participate in the virtual panel in advance of the review meeting. The AFSC NMFS Project Contact will provide the information for the arrangements for this test.

3) Attend and participate in the virtual panel review meeting. The meeting will consist of presentations by NMFS scientists, review of model runs conducted during the course of the evening, and discussion among the reviewers, assessment scientists, other scientists involved in the assessment or management process, and members of the public.

4) After the virtual panel review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.

5) Each reviewer should assist the Chair of the meeting with contributions to the summary report.

6) Deliver their reports to the Government according to the specified milestones dates.

Place of Performance
The place of performance will be held remotely, via Google Meets video conferencing.

Period of Performance
The period of performance shall be from the time of award through June 30, 2021. The CIE reviewers' duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables
The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	Deliverables and Milestones
Within two weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
April 26-30, 2021	Virtual panel review meeting
Approximately 3 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each TOR as specified; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
No travel is necessary, as this meeting is being held remotely.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact:
Grant Thompson
grant.thompson@noaa.gov

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Performance Work Statement
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

The Terms of Reference were compiled from recommendations submitted by the Groundfish Plan Team for the Bering Sea and Aleutian Islands, the Scientific and Statistical Committee, and Alistair Dunn (a consultant contracted by the Freezer Longline Coalition). These were organized into six general topics, with three specific recommendations per topic. After reading the background materials and receiving the initial set of presentations during the review, the reviewers will prioritize the six topics and identify at least one recommendation per topic to be addressed by the review. The reviewers will then address as many of the topics (and the identified recommendation(s)), in priority order, as time allows.

Topic 1: Movement

Recommendation 1a:
Comment on avenues for incorporating spatial dynamics and movement.

Recommendation 1b:
Consider how to inform the dynamics of movement or abundance between the Northern Bering Sea and the Eastern Bering Sea, specifically from additional experiments and analyses, data analyses that include these assumptions (i.e., VAST), and how these can best be used within the different models as indices of abundance.

Recommendation 1c:
Develop movement models.

Topic 2: Ensemble modeling

Recommendation 2a:
Evaluate the use of ensemble modeling in the NPFMC management system, and specifically whether the structural uncertainty and historical challenges in identifying a robust base model make Pacific cod a good application for ensemble modeling.

Recommendation 2b:
Develop the models to include in an ensemble.

Recommendation 2c:
Consider whether to apply the sloping harvest control rule before or after ensemble averaging of SSB and other reference points.

Topic 3: Age data

Recommendation 3a:
Attempt to resolve problems with using fishery age compositions.

Recommendation 3b:
Consider how best to include the fisheries age and size composition data, including consideration of fleet specific age composition data in the model.

Recommendation 3c:
Investigate whether a change in growth contributed to the ageing bias fit for 2008 and onward in the complex models as ageing bias and growth may be confounded.

Topic 4: Fishery CPUE

Recommendation 4a:

Discuss standardization of fishery CPUE using alternative statistical methods, including a discussion of historical changes in the fishery that may affect the relationship of the index to abundance.

Recommendation 4b:
Develop a fishery CPUE index.

Recommendation 4c:
Consider how best to further analyze CPUE, including development of spatio-temporal analyses of fleet specific CPUE indices that may help inform the model or supplement the trawl survey biomass indices.

Topic 5: Compositional data

Recommendation 5a:
Consider methods (e.g., bootstrapping) to estimate uncertainty and variance in the composition data, with the results then used to estimate initial sample sizes for each season, fleet, combination for input into the assessment model.

Recommendation 5b:
Review methods to scale the composition data and include consideration of methods that scale observer samples to the catch by vessel, location, and time of event.

Recommendation 5c:
Consider analyses of the size- and age- composition data to identify if there are specific locations or time periods when a recruitment signal may be apparent to assist in informing the assessment model of the strength of recent recruitment.

Topic 6: Other

Recommendation 6a:
Consider incorporation of dome-shaped survey selectivity.

Recommendation 6b:
Consider the diagnostic plots of fits and residuals (including normalised or Pearson residuals) for the age and size composition data and make recommendations on how the model fits may be improved.

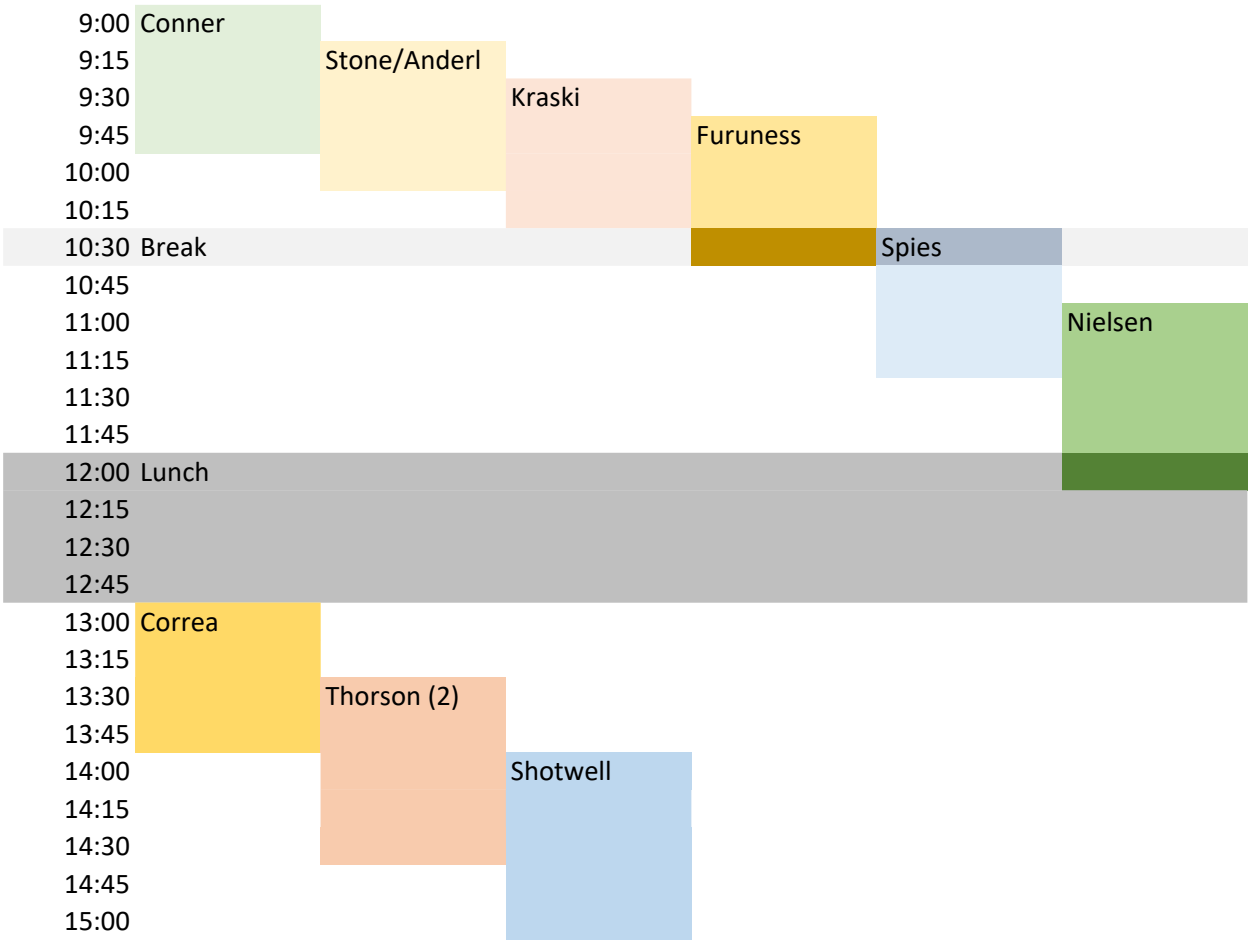
Recommendation 6c:
Consider inclusion of other survey information (e.g., the IPHC and sablefish surveys).

Annex 3: Tentative Agenda

Here is the tentative schedule (agenda) for the meeting:

- Monday
    - Take up questions and answers related to pre-recorded presentations #1-9 in the above list, plus the genetics paper by Ingrid Spies linked in the "Documents" section of the meeting website.
    - See the table below for *a rough idea* of the envisioned schedule for this day (overlapping time slots are intended to suggest that the beginning and ending times are only approximate).
- Tuesday
    - Attend to any unfinished business from Monday.
    - Take up questions/answers related to the final pre-recorded presentation (#10).
    - Prioritize the topics listed in the ToR and select one or more specific recommendations from each topic.
    - Begin working through the ToR, in priority order.
        - This may involve homework assignments for the assessment team.
- Wednesday and Thursday
    - Review results of the preceding day's homework assignments, if any.
    - Continue working through the ToR, in priority order.
        - This may involve additional homework assignments for the assessment team.
        - Note that it may not be possible to address all of the ToR in the time available.
- Friday
    - Review results of the preceding day's homework assignments, if any.
    - Write reports

Table of approximate time slots for questions and answers on pre-recorded presentations during Monday's session (note that presentation #10 will be taken up on Tuesday).


9:00
9:15
9:30
9:45
10:00
10:15
10:30
10:45
11:00
11:15
11:30
11:45
12:00
12:15
12:30
12:45
13:00
13:15
13:30
13:45
14:00
14:15
14:30
14:45
15:00
Conner
Stone/Anderl
Kraski
Furuness
Break
Spies
Nielsen
Lunch
Correa
Thorson (2)
Shotwell

Appendix 3: Panel membership or other pertinent information from the peer review meeting

CIE review panel:

Ingrid Spies, Chair, (NOAA Fisheries, AFSC),
Henrik Sparholt, (CIE)
Arni Magnusson, (CIE)
Yan Jiao, (CIE)

Stock Assessment Team:

Grant Thompson, (NOAA Fisheries, AFSC)
Steve Barbeaux, (NOAA Fisheries, AFSC)
Jason Conner, (NOAA Fisheries, AFSC)
Kalei Shotwell, (NOAA Fisheries, AFSC)

Other participants and their affiliations:

Chad See, (Freezer Longline Coalition)
Craig Kastelle, (NOAA Fisheries, AFSC)
Delsa Anderl, (NOAA Fisheries, AFSC)
Gerry Merrigan, (Freezer Longline Coalition)
Giancarlo Correa, (University of Oregon)
Joel Kraski, (NOAA Fisheries, AFSC)
Julie Neilsen, (University of Alaska Fairbanks)
Kali Stone, (NOAA Fisheries, AFSC)
Mary Furuness, (NOAA Fisheries)
Suzanne Mcdermott, (NOAA Fisheries, AFSC)
Thomas Helser, (NOAA Fisheries, AFSC)
Tim Loher, (International Pacific Halibut Commission)

Abbreviations:

CIE – Center for Independent Experts
NOAA - National Oceanic and Atmospheric Administration
AFSC – Alaska Fisheries Science Center (NMFS/NOAA)