External Independent Peer Review
Assessment of the Pacific cod stocks in the Gulf of Alaska

By

Jean-Jacques Maguire
1450 Godefroy
Québec, Qc, Canada
G1T2E4
jeanjacquesmaguire@gmail.com

For the
Center for Independent Experts

July 2018

Executive Summary

Pacific cod in the Gulf of Alaska has experienced a precipitous decline since 2015, and there is concern that the simpler assessment model rebuilt from scratch in 2016 may not adequately address the important biological complexities of the stock. This Center for Independent Experts (CIE) review of Pacific cod in the Gulf of Alaska is atypical: the preparation of a consensus summary was not included in the terms of references as this CIE review was not part of a regular assessment cycle to provide management advice. The analyst and local participants were very helpful, efficient and forthcoming. The meeting took place in a collegial and pleasant atmosphere.

Two surveys from the Alaska Department of Fish and Game (ADFG) were presented, one in the Central Region and a second in the Westward Region. I believe that the survey in the Central Region is unlikely to provide a reliable index of abundance for the much larger area of the Pacific cod assessment unit, particularly because of the small geographical area covered and the restricted depth range sampled. I do not recommend that it be used in the assessment. The Alaska Department of Fish and Game survey in the Westward Region covers a larger area. I recommend that this survey be used in the assessment and that an environmental covariate with temperature be evaluated to adjust catchability similar to what is done for the Alaska Fisheries Science Center (AFSC) longline (LL) survey. I see no disadvantages of using this survey.

The longline survey conducted by the International Pacific Halibut Commission (IPHC) has a very high sampling intensity and it shows a very close agreement with the AFSC trawl survey. The IPHC longline survey should undeniably be included in the assessment because of the high sampling intensity, since the survey is conducted every year and because the survey covers the depth distribution of Pacific cod. I see no disadvantages of using the IPHC LL survey.

I conclude that the four surveys (NMFS trawl, NMFS LL, ADFG trawl, IPHC LL) should be used in the assessment (index and compositions where available), and I recommend that their influence on results be evaluated by running an assessment with each index separately and by removing each one at a time.

It is difficult to answer categorically the question about model complexity. In principle, simpler models are preferable, but the more important question is whether there are sufficient data to support more complex models. It is comforting that the large majority of model configurations tell a similar story, particularly for the recent past, the most important for management purposes: current stock biomass is the, or close to the, lowest observed. I conclude that the simple model should be run during each assessment cycle as a baseline, but recommend that moderate complexity should be included in the base case assessment to account for specific features such as the lack of fit to size composition in 2005-2006 because there was no fishing in the B season, expected increases in M due to anomalous environmental conditions or changes in

catchability in the longline surveys related to changes in distribution due to environmental condition.

Care should be taken that data weighting does not give too much weight to a single data source that has artificially low variance. This does not seem to be a problem here, and data weighting does not imply drastically different assessment results. Considering that data weighting gave different results for the early part of the data series for 2016 models, I recommend that the effects of data weighting should be monitored, and if data weighting changes the results considerably, data weighting should only be applied if there is strong justification for it.

I conclude that the variability in selectivity in the current assessment is acceptable and justified. Random walk could be tested to evaluate how it affects the results. Selectivity of the surveys would be expected to vary less than selectivity in the fishery which may be affected by regulatory changes. Changes in catchability of the LL surveys due to changes in temperature would be difficult to disentangle from changes in selectivity.

There is no doubt that environmental conditions in the Gulf of Alaska during 2014-2016 were considerably warmer than average, earning the nickname of “endless summer”. Environmental conditions appeared to have reverted to more or less normal in 2017. The information presented in the ecosystem assessment shows that conditions were anomalous and that several species were affected, including Pacific cod. While it seems plausible that natural mortality may have increased during 2014-2016, the assessment shows that fishing mortality is estimated to have increased steadily from near zero in the late 1970 to values close to and higher than F_{OFL}. The fishing mortality on this stock could be dangerously high and it would be prudent to identify, based on the stock and recruitment scatterplot, a spawning stock biomass at which drastic management measures would be taken to prevent the stock from declining further. The reviewers were provided with three indices of the body condition of Pacific cod. None showed the large changes in 2014-2016 that would be associated with increased natural mortality. These results weaken the argument for a possible increase in natural mortality during the warm period. It is, however, difficult to accept that the very exceptional physical and biological conditions in the Gulf of Alaska during 2014-2016 did not have an effect on Pacific cod. An increase in natural mortality is one of the possible causes of an unexpected large decrease in abundance, but migration out of the sampled area or increases in unreported catches could also explain a decrease in survey abundance. I doubt that underreporting of catches or increases in discards are the reasons for the observed decline in abundance, but I recommend that these issues be investigated to eliminate them with certainty. Empirical evidence does not support the hypothesis that GOA Pacific cod migrated out of the area. I consider that the environmental information used in the assessment is handled correctly.

Finally, I consider that the temperature-catchability relationship for the AFSC LL surveys is being modeled appropriate

Background

The National Marine Fisheries Service (NMFS) is mandated to conserve and protect the nation's marine living resources and manage fisheries exploiting them based upon the best scientific information available. NMFS science products, including scientific advice for fisheries management, require strictly independent scientific peer reviews. The Center for Independent Experts provides a formal external process for such independent expert reviews.

Scientific peer review of Pacific cod in the Gulf of Alaska involved three qualified experts to review the scientific information to ensure quality and credibility. These experts conducted their peer review impartially, objectively, and without conflicts of interest. Each reviewer was independent from the development of the science, without influence from any position that the agency or constituent groups may have. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

The Gulf of Alaska Pacific cod stock assessment has had a large number of alternative models over the years. In 2016, the model was rebuilt from scratch and greatly reduced in complexity from the previous model. Of particular concern is that this stock has experienced a precipitous decline since 2015, and there is concern that the simpler assessment model may not adequately address the important biological complexities of this stock in the face of climate variability. The review is to cover all aspects of the stock assessment models. The Pacific cod fisheries in the Gulf of Alaska is of great economic importance garnering $103 million ex-vessel value annually (29% of all Gulf of Alaska groundfish fisheries).

Description of the Individual Reviewer's Role in the Review Activities

I downloaded the review material on April 5, 2018 and read the review material prior to the meeting. I travelled to Seattle on April 30th, for the meeting starting on May 1. Along with my two CIE colleagues, I participated actively in the discussions and requested additional analyses during May 1st to May 3rd.

May 1 began with presentations on the fisheries data collection systems in port and at sea, as well as for those surveys used in the assessment (Alaska Fisheries Science Center trawl survey and AFSC Longline survey) and those (Alaska Department of Fish and Game, International Pacific Halibut Commission) considered to be used in the next assessment. In the afternoon, the ecosystem assessment was presented. The day closed with a review of the assessment history. The number of violations reported by at-sea observers was requested.

May 2 continued with the history of the assessment and an in-depth presentation/discussion of possible changes to the assessment model. An evolving list of requests was started, which was modified as the meeting progressed. The list below was agreed at the end of May 3. Comments in red and highlighting represents reporting/notes from the analyst.

1. For the first new model run, create a new base model by starting with 17.09.35 and switching to length-based maturity; then use that as the base model for all further model runs. Model18.09.38LM in NewModels.zip
2. Make a run without age data prior to 2007. Model18.09.40NO_AGE in NewModels.zip
3. Make a run without age data entirely. Model18.09.39NO_AGEPRE2007 in NewModels.zip
4. Make a run with block-specific ageing bias. Model18.09.41bias in NewModels.zip
5. Get sizecomp data from western ADFG survey and include that survey in a model. Model18.09.42biasSTATE in NewModels.zip
6. Update simplest model from 2016 ("16.0") with 2017 data and provide results (likelihoods and biomass trend). Model 16.0 in Models2016W2017Data.zip
7. Provide r4ss plots for all models that were included in the likelihood table, and all new models. Did this for all models in the various zip files, look in the "plot" folder under each individual model folder
8. Report "condition" in units of percentage change from the mean (priority: low) Provided three methods, Fulton's, Ianelli's (ave weight), and residuals from log linear model approach in file "Assigned_Figures.docx"
9. What priors (other than M) are influencing the Piner plot, and how? (priority: low)
10. Describe Steve's "randomization" method for testing significance of environmental effects (priority: post-meeting)
11. Compare ADFG and NMFS trawl survey sizecomps (as line plots, not bubble plots), superimposed, for years in which both took place, with a vertical axis for each. Figures provided in file "Assigned_Figures.docx"
12. Justify time blocks used in Model 17.09.35, including management history (priority: before meeting ends).

The reconvening time for May 3 was set at 11h00 to allow the analyst sufficient time to do the analyses and prepare the material for presentation. At the end of the day on May 3, the three CIE experts and other participants concluded that they had made all the requests they needed to complete their individual reports. They suggested that it would be preferable to leave time to the analyst on May 4 to complete the assignments making sure that they were done correctly and completely. The CIE experts expected that if there were more issues arising, they could be dealt with by emails. The assignments were indeed done correctly and completely and no other issue arose, therefore, there was no need for an e-mail discussion. As there was no meeting on May 4th, I travelled back home on May 4th.

The analyst and local participants were very helpful, efficient and forthcoming. The meeting took place in a collegial and pleasant atmosphere.

This CIE review of Pacific cod in the Gulf of Alaska is atypical: the preparation of a consensus summary was not included in the terms of references as this CIE review was not part of a regular assessment cycle to provide management advice.

Summary of Findings

During the presentations, it was noted that the TAC seemed not to be caught in some years. Page 190 of the December 2017 Assessment Report provides an interesting description of the TAC setting process: "*To understand the relationships between ABC, TAC, and catch for the period since 1997, it is important to understand that a substantial fishery for Pacific cod has been conducted during these years inside State of Alaska waters, mostly in the Western and Central Regulatory Areas. To accommodate the State-managed fishery, the Federal TAC was set well below ABC (15-25% lower) in each of those years. Thus, although total (Federal plus State) catch has exceeded the Federal TAC in all but three years since 1997, this is basically an artifact of the bi-jurisdictional nature of the fishery and is not evidence of overfishing as this would require exceeding OFL. At no time since the separate State waters fishery began in 1997 has total catch exceeded ABC, and total catch has never exceeded OFL*". The caption to table 2.3 of the assessment report also says something to that effect but not as clearly. It is an interesting example of the cooperative spirit with which fisheries management is implemented in this area that the Federal TAC is set lower to take into account expected catches in non-federal waters.

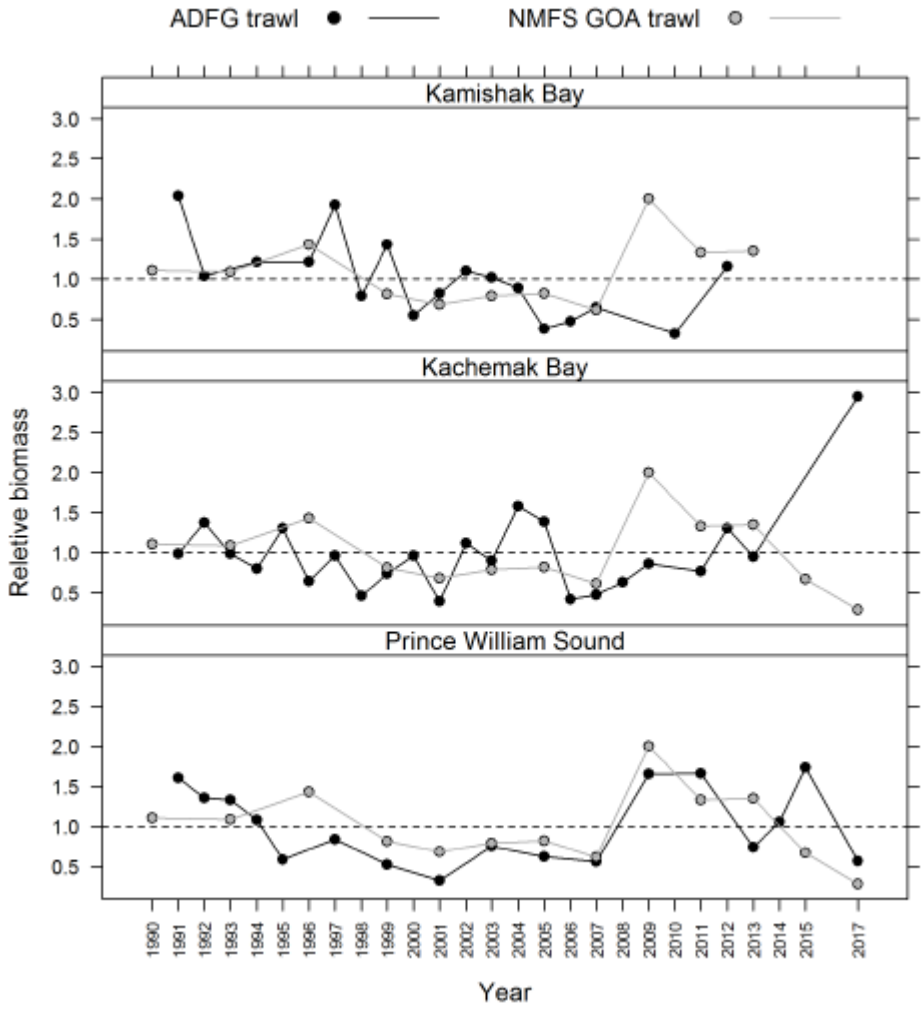
Evaluate and provide recommendations on data used in the assessment models. In particular:

a. What are the benefits vs disadvantages of including data from the ADFG large-mesh trawl and the IPHC longline surveys in the assessment?

By asking for comments only on the ADFG and IPHC surveys, this ToR implies that the AFSC longline survey is now part of the assessment. This is a good addition, more surveys cannot be bad, unless there are clear reasons, e.g. very small area covered unlikely to be representative of the assessment unit.

Two surveys from the Alaska Department of Fish and Game were presented, one in the Central Region and a second in the Westward Region. Both were fixed stations large mesh trawl surveys originally designed to survey tanner crab.

The survey in the Central Region covers inshore areas in Prince William sound, Kachemak bay and Kamishak bay. While the relative biomass trends, particularly in the Prince William sound (Figure below), are similar to the NMFS trawl survey in the Gulf of Alaska,

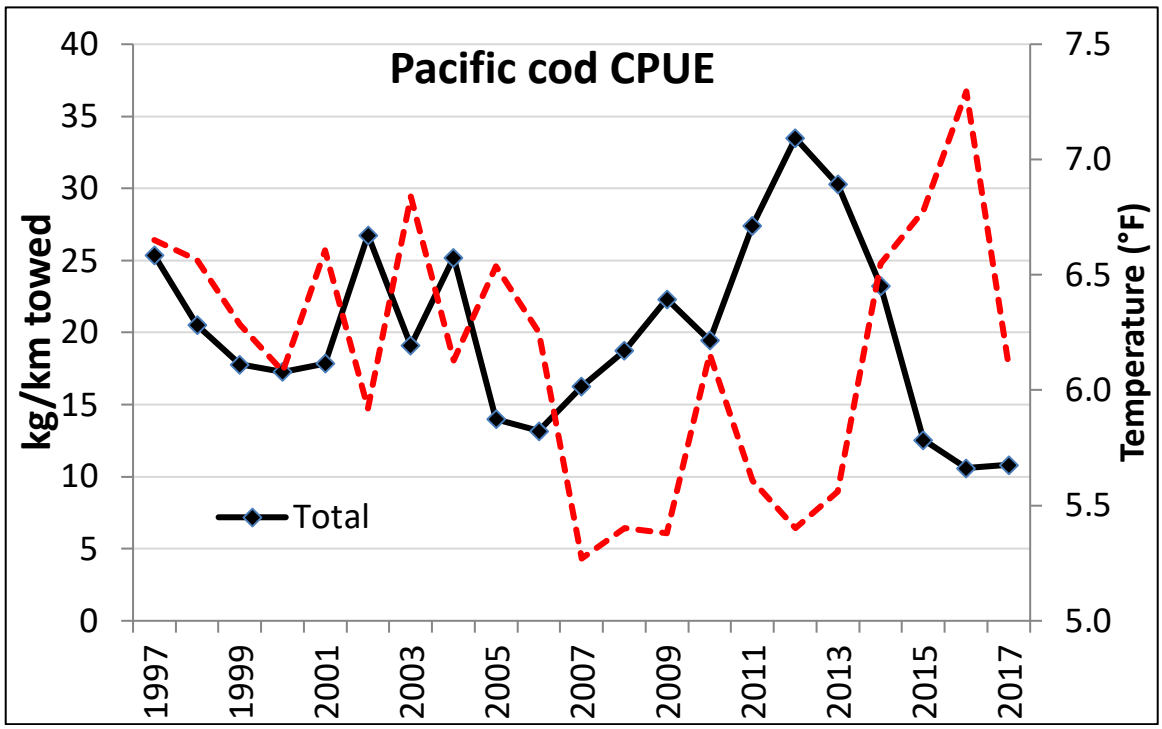


the area surveyed is small and covers exclusively inshore waters. While the similarity in trends with the AFSC trawl survey is interesting, the survey in the Central Region is unlikely to provide a reliable index of abundance for the much larger area of the Pacific cod assessment unit, particularly because of the small geographical area covered and the restricted depth range sampled. I do not recommend that it be used in the assessment as its possibly spurious good correlation with the AFSC trawl survey could mislead the assessment.

The Alaska Department of Fish and Game survey in the Westward Region covers a larger area, some offshore areas have been surveyed in some years, but it too is mostly restricted to inshore waters. The density index agrees broadly with the NMFS trawl survey. The Alaska Department of Fish and Game survey for the Westward Region has

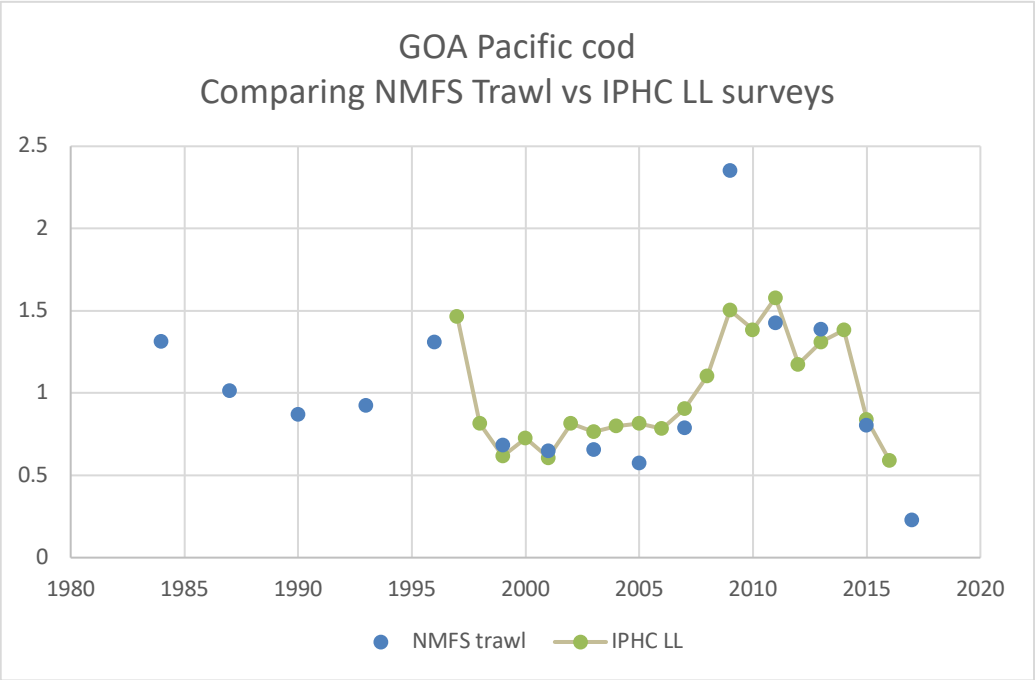
been conducted every year since 1988, with some changes in the area surveyed. Estimates of length composition are also available.

The presenter identified a negative relationship in the ADFG trawl survey between Pacific cod abundance and temperature (Figure below). There were expectations that the ADFG trawl survey, because of its near shore coverage, might provide an index for Pacific cod recruitment, but that does not seem to be the case, possibly because of the large mesh used. I recommend that this survey be used in the assessment and that an environmental covariate with temperature be evaluated to adjust catchability similar to what is done for the AFSC LL survey. I see no disadvantages of using this survey.

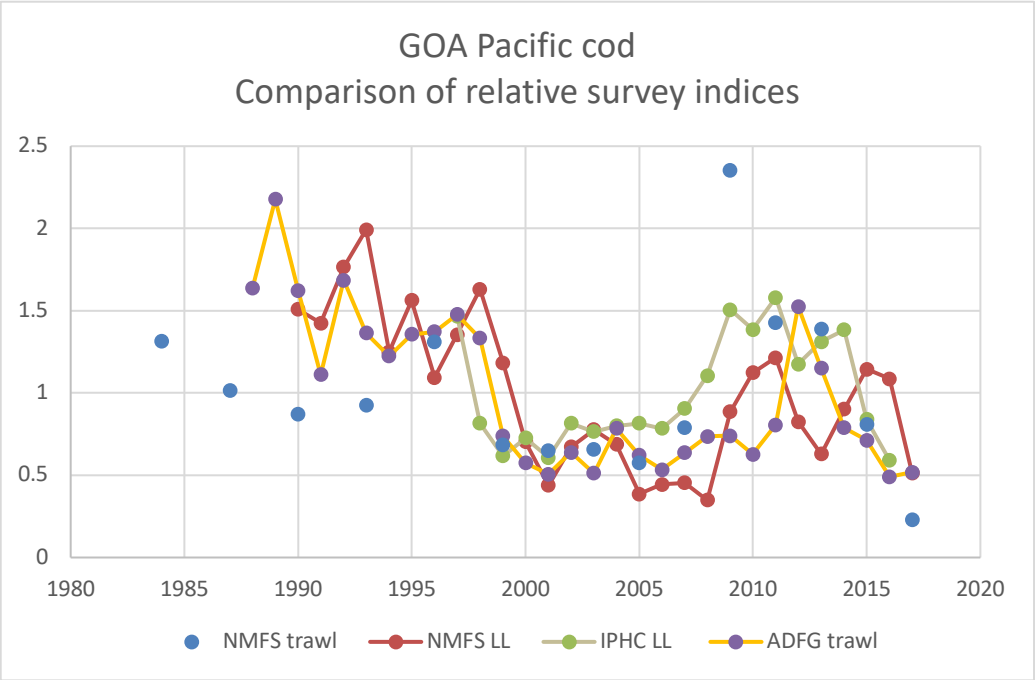


The longline survey conducted by the International Pacific Halibut Commission has a very high sampling intensity but limited Pacific cod size information. More size composition for Pacific cod is expected to be collected starting with the 2018 survey. The absence of size information is a concern, but the very close agreement with the AFSC trawl survey (Figure below) suggests that the IPHC LL survey sees Pacific cod of sizes similar to those in the AFSC trawl survey. This could be further evaluated by estimating the expected size composition in the IPHC longline survey based on the hook and bait sizes used compared with the AFSC LL survey and with the general understanding of Pacific cod selection on hooks. Past survey results should be used immediately in the next assessment, and if and when sufficient Pacific cod size composition become available, they too should be used in the assessment. The IPHC longline survey should undeniably be included in the assessment because of the high sampling intensity, because the survey is conducted every year and because the survey

covers the depth distribution of Pacific cod. I see no disadvantages of using the IPHC LL survey.



The ADFG trawl survey, the NMFS LL survey and the IPHC LL survey are compared with the NMFS trawl survey in the figure below (data copied from Appendix 2.3 Stock Synthesis Files for Model 17.09.35 from the 2017 assessment document). The NMFS trawl survey was conducted every 3 years from 1984 to 1999, and every second year starting in 2001. The other surveys are conducted every year. The 2017 value for the IPHC was not available when the assessment was prepared.



The NMFS LL survey and the ADFG trawl survey are well correlated for the whole series (r = 0.71) and particularly well for the earlier part of the series from 1990 to 2008 (r = 0.87). The two longline surveys show similar declining trends for the first few years of the IPHC survey, but the IPHC survey does not decline as much as the NMFS survey in the early 2000s and begins to increase earlier. Over the time series, the NMFS and IPHC LL surveys are poorly correlated (r = 0.24).

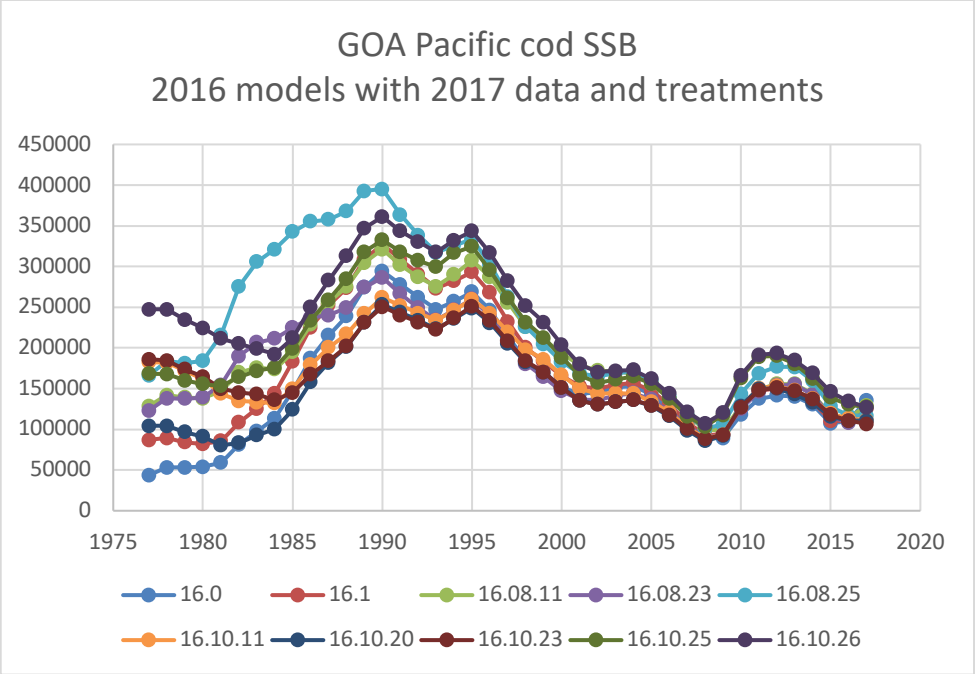
I conclude that the four surveys (NMFS trawl, NMFS LL, ADFG trawl, IPHC LL) should be used in the assessment (index and compositions where available) and I recommend that their influence on results be evaluated by running an assessment with each index separately and by removing each one at a time. The high sampling intensity and broad coverage of the IPHC survey could be particularly helpful, noting that longline gear can sample a different part of the population on untrawlable grounds and in deep water. In addition, using more stock size indices might provide more stable results. The fit to the trawl survey is not very good in any of the configurations examined, while the fit to the AFSC LL survey is better. This is a further reason to include more surveys in the modeling.

Evaluate and provide recommendations on model structure, assumptions, and estimation procedures. In particular:

a. How would you evaluate the appropriate level of complexity in the stock assessment model given that we have historically used simple and more complex models to manage this stock?

It is difficult to answer categorically the question about model complexity. In principle, simpler models are preferable, but the more important question is whether there are sufficient data to support more complex models. Three broad sets of models were reviewed: i) updates of the various 2016 model configurations with 2017 data and data treatments, ii) 2017 models continuing the evolution from the last 2016 model configurations, and iii) new models prepared for the CIE review.

Model 16.0 is the simpler model mentioned in the ToR. It uses a single stock size index (the AFSC trawl survey), age-composition and conditional length at age from the trawl survey, length compositions for the trawl fishery, the longline fishery, the pot fishery and the trawl survey. Natural mortality is fixed at M = 0.38, catchability of the AFSC trawl survey is fixed at Q = 1.0 and selectivity is asymptotic for all fleets / indices except for the pot fishery. Under these conditions, the spawning stock biomass (SSB) is low for a few years after the start of the fishery in 1977 around 50000t until 1983 (SSB from all 2016 models are shown in the Figure below), increases rapidly to 300000t in 1990 before declining irregularly to slightly less than 100000t in 2009. SSB subsequently increased slightly to near 150000t in 2011-2013, followed by declines in the following three years to about 100000t in 2015-2016 with a small uptick in 2017 difficult to see in the Figure.



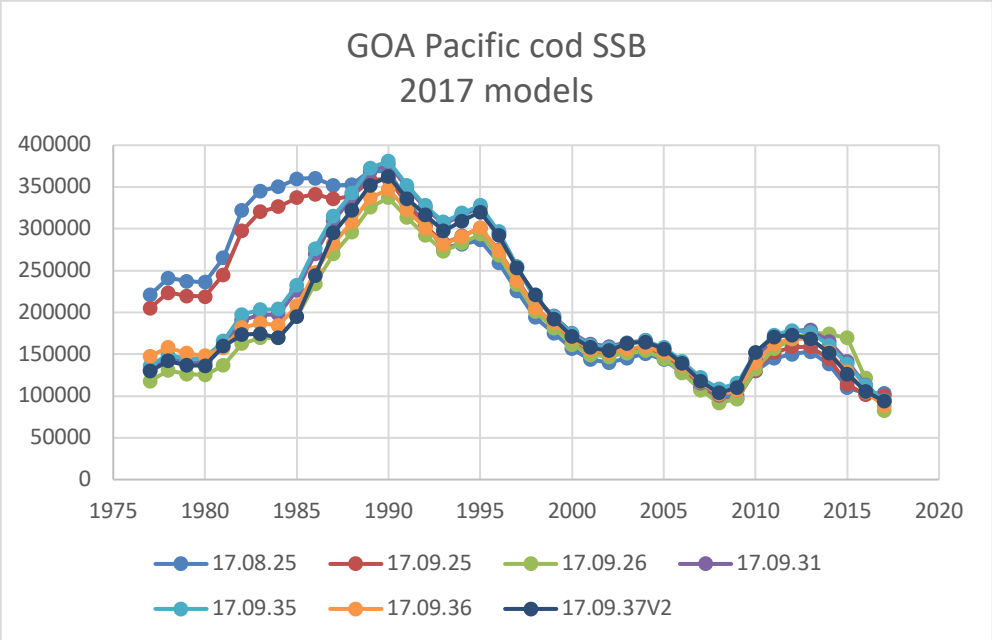
Changes were introduced one-at-a-time to evaluate their effect. Adding the AFSC longline survey (model 16.1) keeps basically the same trends with slightly higher SSB at the start of the series, but very similar absolute values from the early 2000s onwards. Allowing annually varying fishery selectivity to be dome-shaped (model 16.08.11) also increases SSB at the beginning of the series, the maximum SSB is close to that from

model 16.1, but, again, SSB from 2005 onwards are very close to the other two models. Using blocks for fishery and survey selectivity (model 16.08.23) results in early SSB close to model 16.08.11 where annually varying fishery selectivity is allowed to be dome-shaped, SSB in the middle of the series are close to the simple model, and terminal (2017) SSB is lower than in previous models. Fitting natural mortality and survey catchability (model 16.08.25) increases SSB for the whole series, reaching a maximum of 400000t in 1990 with similar trend afterwards and similar terminal value slightly higher than 115000t.

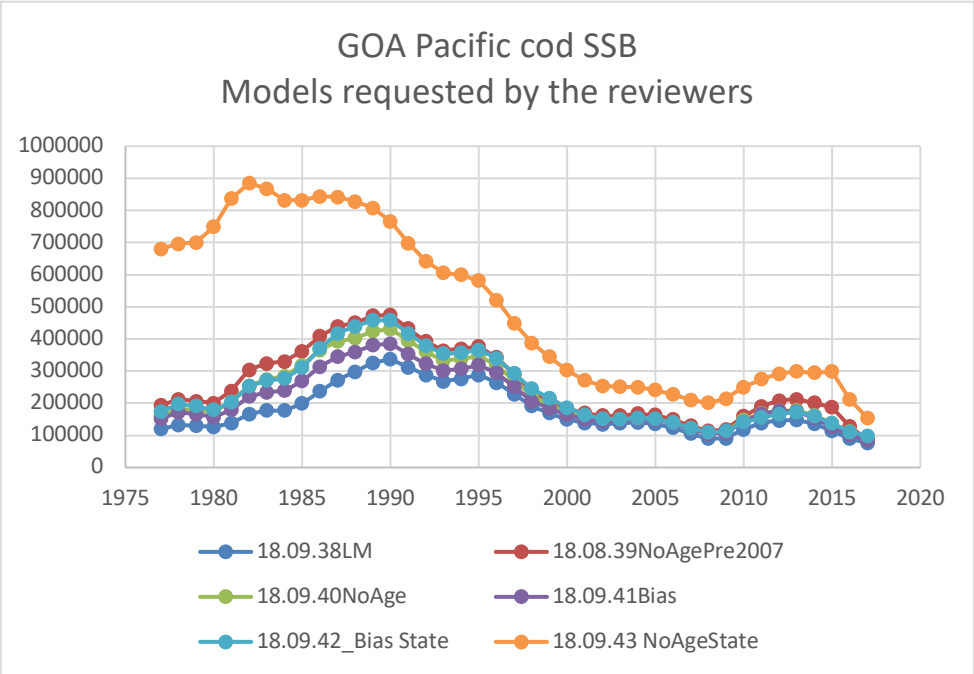
Tuning using the Francis TA1.8 method changes the biomass estimates early in the time series, but results in similar trends afterwards. Allowing annually varying fishery selectivity to be dome-shaped with fixed M and Q (model 16.08.11) and tuning the weights using the Francis TA1.8 method (model 16.10.11) results in a different trend early in the time series, SSB starts around 175000t in 1977, declines steadily to less than 150000t in 1984, but subsequent trends and absolute values are similar to other runs, except that SSB declines in every year during 2012 to 2016. Applying tuning to the simple model using both AFSC surveys (model 16.1) and fixed M and Q (model 16.10.20) results in higher SSB at the beginning of the series (1977-1980), lower SSB for most of the series (1981-2008), and similar values for recent years. With blocks of selectivity (16.08.23), tuning (16.10.23) shows larger SSB differences early in the time series but very similar values during 1995-2017. Fitting M and Q (16.08.25) with tuning (16.10.25) also produces smaller SSBs for the early part of the series, but very similar ones during 1995-2017. Model 16.10.26 is the same as Model 16.10.25 except for block on M for 2015-2016 and selectivity for fisheries allowed to vary annually with CV 0.2 tuned using the Francis TA1.8 method. The SSB trends are the same as other “tuned” 2016 models, but it has the highest absolute values among those.

The 2017 models continued the evolution from configurations in 2016 with the two AFSC surveys and length compositions from the trawl, longline and pot fisheries. Similar to the 2016 models, changes were introduced one-at-a-time. Model 17.09.25 allows dome-shaped selectivity by blocks for all fisheries and surveys, and estimates M and Q. Model 17.09.26 uses blocks for M and Q, model 17.09.31 adds a covariate with temperature to estimate the LL survey Q, model 17.09.35 blocks selectivity for the LL and trawl fishery for 2005-2006, model 17.09.36 is tuning the weights using the Francis TA1.8 method. Interestingly, it does not show the declining biomasses at the beginning of the time series that the 2016 models showed. Model 17.09.37V2 includes four survey indices (AFSC trawl and LL surveys, ADFG trawl survey, IPHC LL survey). Results are similar to other model configurations. SSB estimates are very similar for all models for the entire period, except models 17.08.25 and 17.09.25 which have higher estimates for

1977 to 1990.



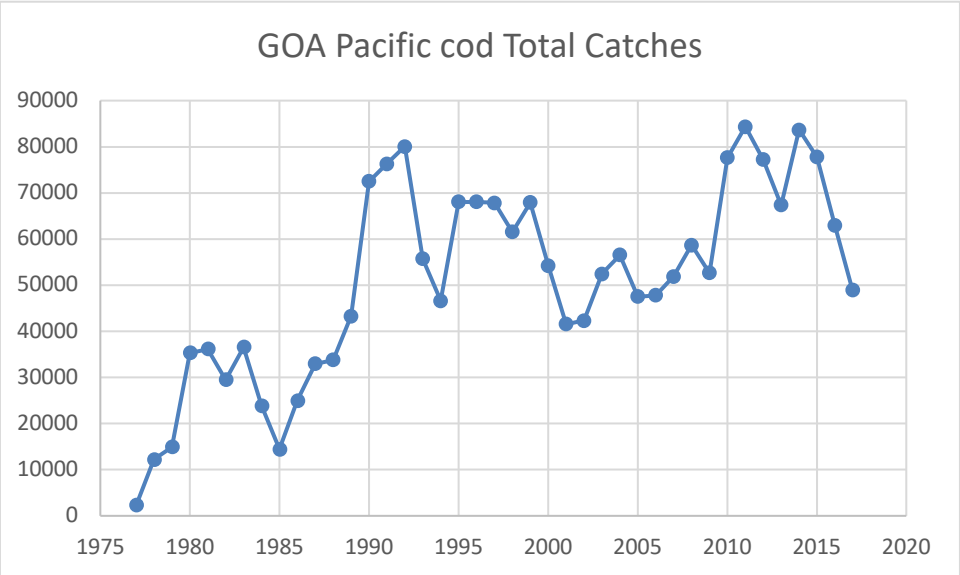
This paragraph covers the request by CIE reviewers. Using maturity at length rather than maturity at age (model 18.09.38LM) simply decreases the SSB over the entire time series. Not using survey age composition prior to 2007 (model 18.09.39No_Agepre2007) and changing standard deviation of M prior results in higher SSB for most of the time series, but with SSB in 2017 very close to the previous model. The fit to the LL index seems to be better, but there were several warnings that parameters were out of bound. Removing all survey age compositions (model 18.09.40No_Age) reduces the SSB for the whole time series, with the 2017 SSB very close to the previous model. Adding back the survey age compositions with ageing error and bias (model 18.09.41_bias) results in slightly lower SSB for the first half of the series, but very close values during 1998-2017. Adding the ADFG large mesh Westward trawl survey (model 18.09.42_BiasState) results in very similar SSBs during 2002-2017 and slightly higher SSB for earlier years, but adding the ADFG trawl survey when removing the AFSC trawl survey age composition (model 18.09.43NoAgeState) results in considerably higher SSB early in the time series and somewhat higher in recent years. This model did not converge, however, and it could be informative to identify why.



Note that several of the models described above had warnings, sometimes several of them. Those models with warnings would need to be improved if they were to be used to formulate management advice, but for the purpose of this review, the results were taken as given and included in the descriptions above.

It is comforting that the large majority of model configurations tell a similar story, particularly for the recent past, the most important for management purposes: current stock biomass is the, or close to the, lowest observed.

Conventional wisdom for Gulf of Alaska Pacific cod is that exploitation started at the end of the 1970s, and that catches prior to 1977 were minimal (catches used in the assessment are shown in the Figure below).



If it is really the case that Pacific cod was not significantly by-caught in foreign fisheries prior to extension of jurisdiction, the stock would be expected to be near virgin conditions at the beginning of the fishery. This implies that scenarios with steeply increasing biomass at the beginning of the time series (most of the 2016 models without tuning) would be unlikely, unless the resource experiences large fluctuations even in the absence of fishing as its nickname in Aleut implies: "*Atxidaq, the fish that stops*". If this were the case, the period of the fluctuations would be relatively long because stock trends over the 40 some years of the assessments are low frequency. The possibility that fishing mortality is too high cannot be eliminated however. I discuss this further under ToR 3 on the effect of the environment.

I conclude from the above that the simple model should be run during each assessment cycle as a baseline, but recommend that moderate complexity should be included in the base case assessment to account for specific features such as the lack of fit to size composition in 2005-2006, because there was no fishing in the B season, expected increases in M due to anomalous environmental conditions or changes in catchability in the longline surveys related to changes in distribution due to environmental condition.

b. What factors should be considered in data weighting and how should we assess the appropriateness of current methods applied for this stock?

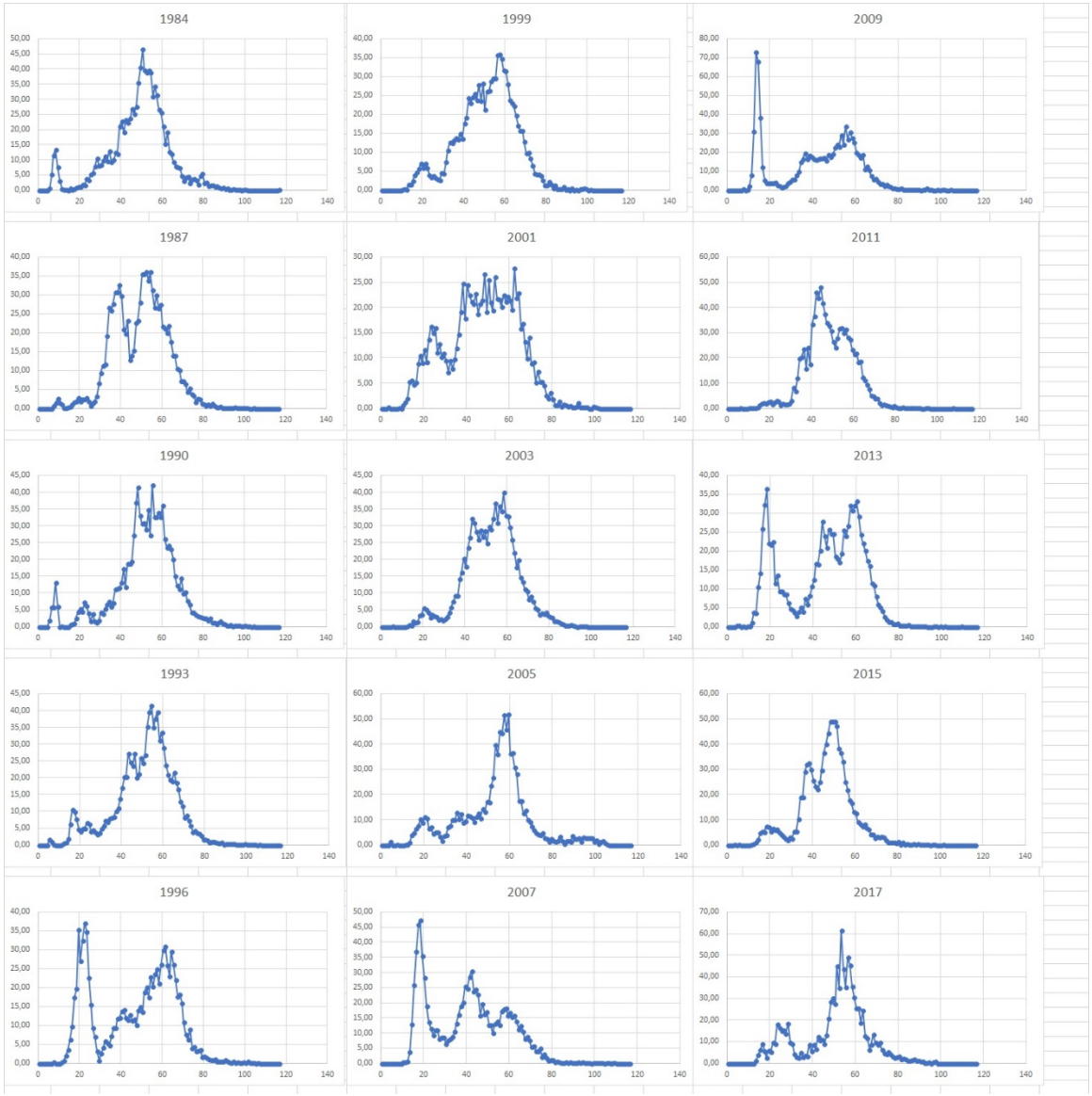
Francis TA1.8 method was used for data weighting in the 2017 GOA Pacific cod assessment. With the 2016 models, data weighting made changes early in the time series, decreasing biomass for the first 8-9 years, but the trends in subsequent years were similar to other runs without data weighting. With the 2017 models, data weighting did not provide results that stood out from models without data weighting.

Care should be taken that data weighting gives too much weight to a single data source that has artificially low variance. This does not seem to be a problem here, and data weighting does not imply drastically different assessment results. Considering that data weighting gave different results for the early part of the data series for 2016 models, I recommend that the effects of data weighting should be monitored, and if data weighting changes the results considerably, data weighting should only be applied if there is strong justification for it.

c. How can we evaluate the appropriate level of time variability and appropriate pattern (i.e. blocking vs random walk) in fishery and survey selectivity patterns?

Time blocks should be based on known or expected changes in the fishery or on the biology of the species. Current blocks appear justified based on changes in management or in survey protocols.

Growth changes would be expected to result in corresponding changes in selectivity. Changes in the average size of GOA Pacific cod in the fishery were shown in the presentation, however the pattern suggest the passage of a dominant year class rather than actual changes in size at age. For the largest majority of gadoid stocks where there are sufficient age and size information, monotonous changes in size at age over time are observed. The reasons for these changes remain unclear, they may be due to density dependence or to environmental changes, but what is clear is that average size at age changes over time for most gadoid stocks in a non-random fashion, i.e. periods of decreasing size at age followed by periods of increasing size at age. To my knowledge, such changes have not been documented for GOA Pacific cod, but they can be expected to exist here as well. The size composition of the AFSC trawl survey (Figure below) shows that the size of the first mode varies over time. It can be expected that such differences would be maintained over time, and that slower growing year-classes would be maintained throughout the life of the cohort. It is not clear how SS3 handles such changes in size at age.



The analyst reported that if selectivity is allowed to vary in recent years this implies less weight to the survey data. We did not clarify if that happened only when the AFSC trawl survey was used by itself or when both the trawl and LL AFSC surveys are used. In either case, the survey should not be down weighted to allow a better fit to the size composition. This makes a further case to include more surveys (ADFG and IPHC) in the assessment. When there is complete catch at age information for a sufficient period of time, virtual population analysis used carefully can be helpful in identifying periods of changes in selectivity. In this context, I support completely the SSC recommendation "*that ageing additional fishery otoliths for this assessment be a priority, noting that the AFSC has an ongoing ageing-prioritization analysis which may guide their future efforts, and the author has recommended working with the age and growth lab on this project*".

I conclude that the variability in selectivity in the current assessment is acceptable and justified. Random walk could be tested to evaluate how it affects the results. Selectivity of the surveys would be expected to vary less than selectivity in the fishery, which may be affected by regulatory changes. Changes in catchability of the LL surveys due to changes in temperature would be difficult to disentangle from changes in selectivity. Given the results of the model with ageing bias, I concur with the SSC recommendation that "*aging bias should be explicitly included in the next assessment*".

Evaluate how ecosystem indicators are used in the assessment and provide recommendations how they can be better integrated into model development and stock management.

a. Should environmental indices be used to model natural mortality in the model? Is it appropriate to use a time block for the extremely warm period to adjust natural mortality?

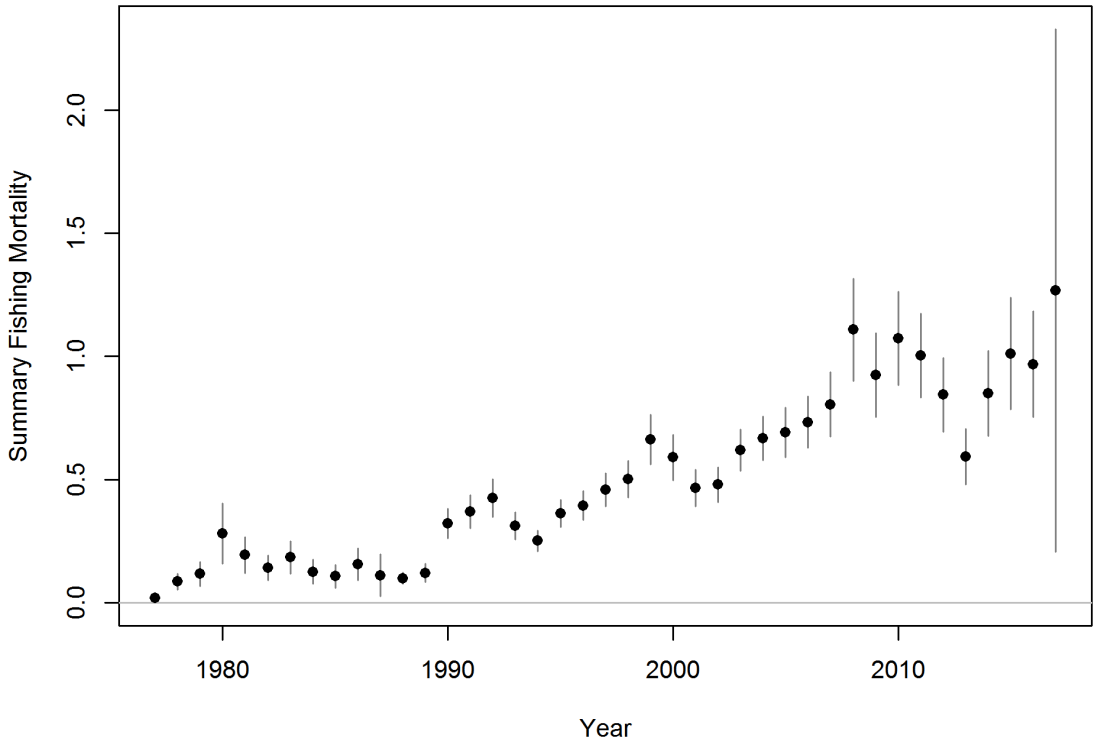
There is no doubt that environmental conditions in the Gulf of Alaska during 2014-2016 were considerably warmer than average earning the nickname of "endless summer". Big changes in the physical environment first observed in 2014 remained similar in 2015 and 2016. The biological components did not react immediately, but there were bad signs in both 2015 and 2016, albeit in different ways. Record low numbers of larval fish were collected in 2015 for several species, including Pacific cod while some species, e.g. Rockfish, had very high larval abundance. Capelin and sand lance declined during the heat wave while herring had declined earlier. There were reports of "sick" cod in the recreational fishery during 2015, more southern species (sunfish (Mola mola), pomfret, market squid) were seen in the GOA, seabirds experienced starvation, large whale mortality increased, few humpback whale calves were born, steller sea lions decreased, and tropical Pyrosomes were observed throughout the GOA in 2017.

Pacific cod experiences relatively rapid growth and short lifespans compared to other groundfish forage fish-eating predators. During 2015-2016 cod metabolic rate remained high year-round, requiring more food at the time when preys had declined. This could increase natural mortality. The analyst provided us with three indices of condition from the NMFS trawl survey by length category for Pacific cod (data were plotted for 2001-2017): i) % difference from mean Fulton's, ii) % difference from linear growth model, and iii) % difference from mean weight. For Fulton's condition, 2015 is generally lower than the average, but only by a few percentage points, and not lower than observed in previous years. For 2017, the values are average or above average. For the linear growth model, results by length category are more variable, but 2015 does not stand out as exceptional. The results for mean weights are similar, i.e. 2015 does not stand out as exceptional. These results weaken the argument for a possible increase in natural mortality during the warm period. It is, however, difficult to accept that the very exceptional physical and biological conditions in the Gulf of Alaska during 2014-2016 did

not have an effect on Pacific cod. I believe it likely that natural mortality of Pacific cod has increased in 2014-2016. However, the relatively small changes in the three indices the Pacific cod of condition examined suggests that further investigation is needed as to why such exceptional environmental conditions did not cause larger changes in the condition indices examined.

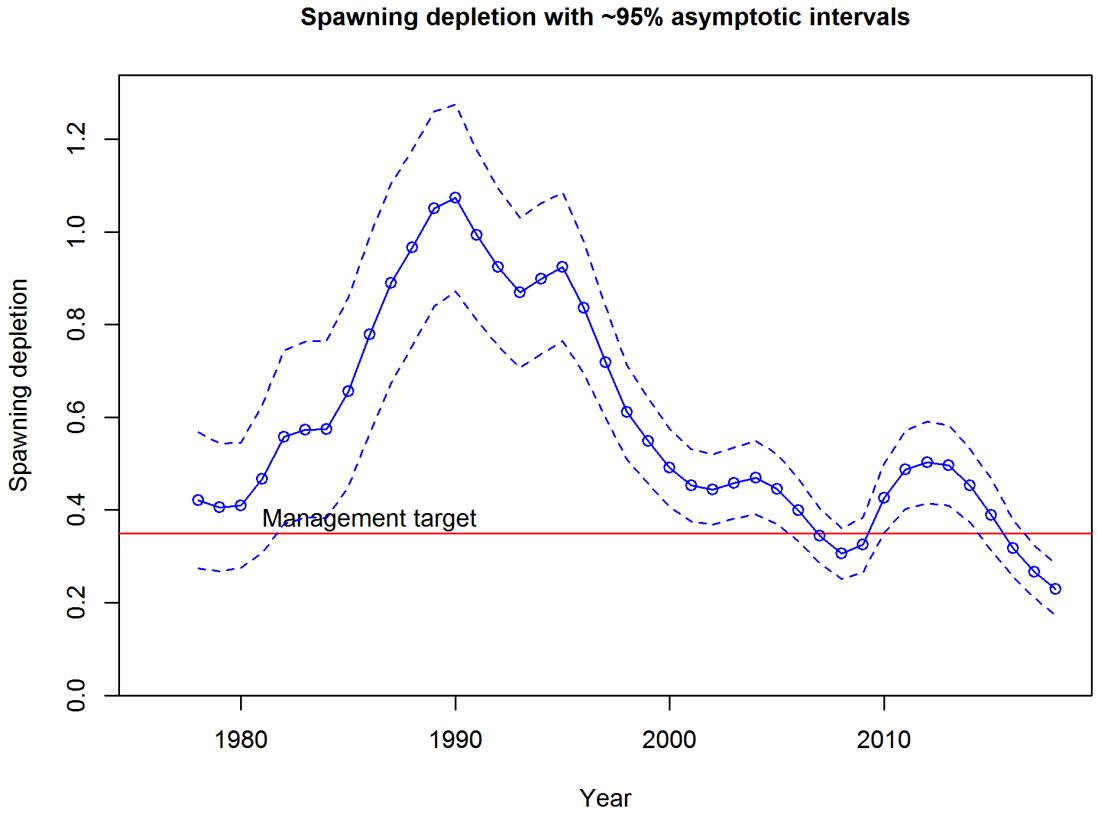
Environmental conditions appeared to have reverted to more or less normal in 2017. The information presented in the ecosystem assessment shows that conditions were anomalous and that several species were affected, including Pacific cod.

While it seems plausible that natural mortality may have increased during 2014-2016, the assessment shows that fishing mortality is estimated to have increased steadily from near zero in the late 1970s to values close to and higher than 1.0 in recent years in model 17.09.35, the author's preferred model (Figure below).



On page 35 of the assessment document in section Amendment 56 reference points, F_{OFL} is defined as $F_{35\%}$ and calculated to be F=1.045 for the current conditions in the fishery. Fishing mortality appears to have been near or exceeded F_{OFL} in recent years. This is also shown in the estimated spawning depletion for the same model 17.09.35 (Figure below) where the management target has been breached. While it is acknowledged that fisheries management in this area has generally been very careful, fishing mortality on this stock could be dangerously high and it would be prudent to

identify, based on the stock and recruitment scatterplot, a spawning stock biomass at which drastic management measures would be taken to prevent the stock from declining further.



An increase in natural mortality is one of the possible causes of an unexpected large decrease in abundance, but migration out of the sampled area or increases in unreported catches could also explain a decrease in survey abundance. While fishermen in this area are recognized as respectful of the regulations and conservation oriented, it is not impossible that misreporting did occur. From the June 2017 presentation on the observer program to the North Pacific Fisheries Management Council, Slide 12 states: "*There is strong evidence of bias in unobserved trips relative to observed trips, and some vessels conducting an entire fishing season without carrying an observer*" while Slide 17 states: "*Compliance and enforcement issues remain a problem within the observer program that are contributing to bias and in fact seem to be getting worse*". From the discussion during the review, it is not clear that if misreporting were to occur, current mechanisms would be able to detect it in the short term. I doubt that underreporting of catches or increases in discards are the reasons for the observed decline in abundance, but I recommend that these issues be investigated to eliminate them with certainty.

According to the presentations, Gulf of Alaska Pacific cod is assumed to be a single discrete stock, distinct from Hecate Straight cod and those further south, there is

evidence for separation of GOA cod from Aleutian Islands cod, but GOA cod and Unimak pass cod appear to be more closely related. These conclusions are also supported by tagging data. Stock boundaries, however, are not fixed in time: the Gulf of St. Lawrence stock of cod was caught outside of the management unit in the late 1980s and the management unit was changed to take that into account. The point here being that GOA Pacific cod could have moved outside of the management unit. Empirical evidence does not support the hypothesis that GOA Pacific cod migrated out of the area: there was an increase in the northern Bering sea that could match the decrease observed in the Eastern Bering sea, but there was no report of increases in the GOA neighboring stocks matching the decrease in the GOA.

The Icelandic cod assessment makes allowance for the return, as adults, of larvae that drifted to Greenland. This is achieved by allowing for immigration/emigration event to happen at a specific age and a specific year. This acts like a recruitment event but on spawning age fish. The last time such an event was allowed to happen was age 6 in 2009 (the 2003 year class). This is an ad-hoc approach, based on information in the catch at age matrix when unusually large amounts of 6 years old appeared in 2009 catches compared the catch history of the same year-class at age 3, 4 and 5. It may be difficult to apply this approach to GOA Pacific cod because it implies relatively high sampling intensity for age-composition which may not be available.

The decrease in Pacific cod survey estimates could also be due to changes in the depth distribution with Pacific cod being more widespread in the water column and caught in a smaller proportion than usual particularly in the trawl survey. This could be investigated from acoustic records from the survey. In addition, there is expertise in joint trawl/acoustic surveys in this laboratory.

I consider that the environmental information used in the assessment is handled correctly.

b. Is the temperature-catchability relationship modeled for AFSC surveys being modeled appropriately?

In the 2017 assessment, the temperature adjustment to catchability was applied only to the AFSC longline survey. The addition of the 2017 trawl survey suggested that the adjustment was not necessary for that survey. The thinking behind the adjustment is that the longline survey covering deeper waters than the trawl survey may be affected more by changes in depth distribution of Pacific cod as a result of changes in temperature.

The SS User Manual (February 11, 2015, page 66) Section 9.3.16 Catchability, describes how to establish a parameter to create environmental effect on catchability. The catchability section on page 22 of the 2017 assessment report seems consistent with what the SS user manual says. I have no reason to question this approach.

Conclusions and Recommendations

I conclude that the four surveys (NMFS trawl, NMFS LL, ADFG trawl, IPHC LL) should be used in the assessment (index and compositions where available), and I recommend that their influence on results be evaluated by running an assessment with each index separately and by removing each one at a time.

I do not recommend that the ADFG trawl survey in the Central Region be used in the assessment because of the limited geographical and depth coverage, and because its possibly spurious good correlation with the AFSC trawl survey could mislead the assessment.

I conclude that the simple model should be run during each assessment cycle as a baseline, but recommend that moderate complexity should be included in the base case assessment to account for specific features such as the lack of fit to size composition in 2005-2006, because there was no fishing in the B season, expected increases in M due to anomalous environmental conditions or changes in catchability in the longline surveys related to changes in distribution due to environmental condition.

I conclude that the variability in selectivity in the current assessment is acceptable and justified. Random walk could be tested to evaluate how it affects the results. Selectivity of the surveys would be expected to vary less than selectivity in the fishery which may be affected by regulatory changes. Changes in catchability of the LL surveys due to changes in temperature would be difficult to disentangle from changes in selectivity.

Considering that data weighting gave different results for the early part of the data series for 2016 models, I recommend that the effects of data weighting should be monitored, and if data weighting changes the results considerably, data weighting should only be applied if there is strong justification for it.

I concur with the SSC recommendation that "*aging bias should be explicitly included in the next assessment*".

I support completely the SSC recommendation "*that ageing additional fishery otoliths for this assessment be a priority, noting that the AFSC has an ongoing ageing-prioritization analysis which may guide their future efforts, and the author has recommended working with the age and growth lab on this project*".

The relatively small changes in the three indices the Pacific cod of condition examined suggests that further investigation is needed as to why such exceptional environmental conditions did not cause larger changes in the condition indices examined.

While it is acknowledged that fisheries management in this area has generally been very careful, fishing mortality on this stock could be dangerously high and it would be prudent to identify, based on the stock and recruitment scatterplot, a spawning stock biomass at which drastic management measures would be taken to prevent the stock from declining further.

I doubt that underreporting of catches or increases in discards are the reasons for the observed decline in abundance, but I recommend that these issues be investigated to eliminate them with certainty.

With regards to process, as indicated above, the analyst and local participants were very helpful, efficient and forthcoming, and the meeting took place in a collegial and pleasant atmosphere. Documents and analyses were made available on a shared drive, which was helpful, but this could be improved. The North East Fisheries Science Center has developed a really helpful data portal for its demersal stock assessments (https://www.nefsc.noaa.gov/saw/sasi/sasi_report_options.php) where ALL information is readily available. This greatly facilitates the review process. I recommend that such a system be implemented for the AFSC peer reviews.

Appendix 1: Bibliography of materials provided for review

Several background papers, in addition to the material to be reviewed were provided.

Barbeaux, S.J. 2017. Gulf of Alaska Pacific cod model updates – September 2017. 31p.

Barbeaux, Steven, Aydin, Kerim, Fissel, Ben, Holsman, Kirstin, Palsson, Wayne, Shotwell, Kalei, Yang, Qiong, and Zador, Stephani. 2017. Chapter 2: Assessment of the Pacific cod stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE. 144p.

Barbeaux, Steven, A'mar, Teresa and Palsson, Wayne. 2016. Chapter 2: Assessment of the Pacific cod stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE. 150p.

Barbeaux, Steven, A'mar, Teresa and Palsson, Wayne. 2016b. GOA Pacific cod assessment 2016. PDF presentation to the NPFMC Plan Team, Nov. 16, 2016. 57 slides.

Barbeaux, Steve, Holsman, Kirstin, and Zador, Stephani. 2018. Gulf of Alaska Pacific cod. PDF presentation to the ComFish Alaska, Kodiak, Alaska. March 22, 2018. 29 slides.

Zador, Stephani and Yasumiishi, Ellen [ed]. 2017. Ecosystem Considerations 2017 Status of the Gulf of Alaska Marine Ecosystem. NPFMC Gulf of Alaska SAFE. 215p.

Appendix 2: A copy of the CIE Statement of Work

Statement of Work

External Independent Peer Review by the Center for Independent Experts

Assessment of the Pacific cod stocks in the Gulf of Alaska

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[1]. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

Scope

The Gulf of Alaska Pacific cod stock assessment has had a large number of alternative models over the years. In 2016, the model was rebuilt from scratch and greatly reduced in complexity from the previous model. Of particular concern is that this stock has experienced a precipitous decline since 2015 and there is concern that the simpler model may not adequately address the important biological complexities to appropriately manage this stock in the face of climate variability. However, review is requested of all aspects of the stock assessment models. The Pacific cod fisheries in the Gulf of Alaska is of great economic importance garnering $103 million ex-vessel value annually (29% of all Gulf of Alaska groundfish fisheries). The individual review reports

[1] http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf

are to be formatted with content requirements as specified in **Annex 1**. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements
Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment methods in general, and in Stock Synthesis in particular.

Tasks for Reviewers
Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: 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 reviewers the necessary background information and reports for the peer review. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for this peer review.

2016 Assessment of the Pacific cod stock in the Gulf of Alaska (150 p.)
2017 Assessment of the Pacific cod stock in the Gulf of Alaska (144 p.)
2017 Ecosystem Considerations Status of the Gulf of Alaska Marine Ecosystem (215 p.)
Comments on the final 2016 and 2017 Gulf of Alaska (GOA) Pacific cod assessments by the Plan Team and Scientific and Statistical Committee (*SSC*)

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with this SoW and ToRs, and shall not serve in any other role unless specified herein. Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein.

This review meeting will include three main parts: The first will consist of a series of presentations with follow-up questions and discussions by CIE reviewers, and will be chaired by an AFSC scientist or supervisor. The second will consist of real-time model runs and evaluations conducted in an informal workshop setting, and will be chaired jointly by the CIE reviewers. The third, time permitting, will consist of initial report writing by the CIE reviewers, with opportunity for additional questions of the assessment author.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website:
http://deemedexports.noaa.gov/
http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
Participate during the panel review meeting scheduled in Seattle, WA during May 1 - 4, 2018.
Approximately three weeks after the conclusion of the panel review meeting, each CIE reviewer shall submit an independent peer review report addressed to the CIE. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**.

Place of Performance
The place of performance shall be at the contractor's facilities, and Seattle, Washington.

Period of Performance

The period of performance shall be from the time of award through June 2018. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

March 26, 2018	CIE selects and confirms reviewers. Reviewer contact information is sent to the NMFS Project Contact
April 16, 2018	NMFS Project Contact sends the reviewers the pre-review documents
May 1 - 4, 2018	Each reviewer participates and conducts an independent peer review during the panel review meeting
Approximately three weeks later	CIE receives draft reports
Approximately two weeks later	CIE 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 (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $12,000.

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

NMFS Project Contact:
Steven J. Barbeaux, Alaska Fisheries Science Center
7600 Sand Point Way NE
Seattle, WA 98115
Phone: 206-526-4211
Steve.Barbeaux@noaa.gov

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall 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 main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role 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 should describe in their own words the review activities completed during the panel review meeting, including providing 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, and especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the Summary Report that they feel 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 CIE independent 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 CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

Assessment of the Pacific cod stocks in the Gulf of Alaska

Evaluate and provide recommendations on data used in the assessment models. In particular:
What are the benefits vs disadvantages of including data from the ADFG small-mesh trawl and the IPHC longline surveys in the assessment?
Evaluate and provide recommendations on model structure, assumptions, and estimation procedures. In particular:
How would you evaluate the appropriate level of complexity in the stock assessment model given that we have historically used simple and more complex models to manage this stock?
What factors should be considered in data weighting and how should we assess the appropriateness of current methods applied for this stock?
How can we evaluate the appropriate level of time variability and appropriate pattern (i.e. blocking vs random walk) in fishery and survey selectivity patterns?
Evaluate how ecosystem indicators are used in the assessment and provide recommendations how they can be better integrated into model development and stock management.
Should environmental indices be used to model natural mortality in the model? Is it appropriate to use a time block for the extremely warm period to adjust natural mortality?
Is the temperature-catchability relationship modeled for AFSC surveys being modeled appropriately?

Annex 3: Tentative Agenda
CIE Review of the GOA Pacific cod stock assessment models
Alaska Fisheries Science Center
7600 Sand Point Way NE, Seattle, WA 98115
May 1 - 4, 2018
Building 4; Room 2039
Review panel chair: Grant Thompson, Grant.Thompson@noaa.gov
Senior assessment author: Steven J Barbeaux, Steve.Barbeaux@noaa.gov
Security and check-in: Sandra Lowe, Sandra.Lowe@noaa.gov (206)526-4230
Sessions will run from 9 a.m. to 5 p.m. each day, with time for lunch and morning and afternoon breaks. Discussion will be open to everyone, with priority given to the panel and senior assessment author.
Tuesday, May 1
Preliminaries:
09:00 Introductions and adoption of agenda—Grant Thompson
Data sources (current and potential):
09:10 Overview of data types used in the assessments—Steve
09:20 Catch accounting system and in-season management—AKRO SF Division (via WebEx)
09:50 Observer program—AFSC FMA Division
10:20 Break
10:30 GOA trawl survey—AFSC RACE Division
11:00 AFSC longline survey—AFSC Auke Bay Laboratory (via WebEx)
11:30 IPHC longline survey—IPHC
12:00 Lunch
13:00 ADFG surveys— ADFG (via WebEx)
13:30 GOA Ecosystem assessment—AFSC REFM – Stephani Zador
Assessment models:
14:00 Assessment history—Steve
15:00 Break
15:10 Current assessments—Steve
16:10 Discussion—Everyone
16:40 Assignments for models to be presented on Wednesday—Panel
Wednesday, May 2 and Thursday, May 3
Review of models assigned the previous day—Steve
Discussion, real-time model runs—Everyone
Assignments for models to be presented the following day—Panel
Friday, May 4
Review of models assigned on Thursday—Steve
Discussion, real-time model runs—Everyone
Report writing (time permitting)—Panel

Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

GOA Pacific cod CIE review May 1-4, 2018

List of presenters:

1. Steve Barbeaux (AFSC)
2. Mary Furuness (NMFS Alaska Region)
3. Marlon Concepcion (AFSC)
4. Wayne Palsson (AFSC)
5. Dana Hanselman (AFSC)
6. Allan Hicks (IPHC)
7. Kally Spalinger (ADFG)
8. Mike Byerly (ADFG)
9. Stephani Zador (AFSC)

List of CIE reviewers:

1. Jean-Jacques Maguire
2. Henrik Sparholt
3. Kevin Stokes

List of other in-person participants

1. Delsa Anderl (AFSC)
2. Jim Armstrong (North Pacific Fishery Management Council)
3. Craig Castelle (AFSC)
4. Anne Hollowed (AFSC)
5. Jim Ianelli (AFSC)
6. Sandi Neidetcher (AFSC)
7. Chad See (Freezer Longline Coalition)
8. Grant Thompson (AFSC)
9. Tom Wilderbuer (AFSC)