Center for Independent Experts (CIE) Independent Peer Review Report

Gulf of Alaska Pacific Ocean Perch Assessment Review

Prepared for the Center for Independent Experts

By

Dr. Geoff Tingley

Gingerfish Ltd, Wellington, New Zealand
email: fishinnz@hotmail.co.uk

July 2021

Table of Contents

Executive Summary

- This document is the individual CIE Reviewer Report of the review of the stock assessment of the Gulf of Alaska stock of Pacific ocean perch (*Sebastes alutus*). The review was conducted during late March and into April 2021, with the review meeting held from 10.00 (PDT) daily from 30th March to 1st April using a virtual platform (Webex) due to the SARS-CoV-2 pandemic. The meeting timing was a day later for this reviewer due to the difference in time zone and started at 06.00 NZDT on the 31st March. This report solely represents the views of the independent CIE reviewer Geoff Tingley.
- The current and recent historical assessment documents for the stock, voiced presentations, a video as well as other relevant background documents, were provided in advance of the meeting. This material was posted on a dedicated Webex webpage, from where it could be downloaded. Some updates to presentations were posted quite close to the meeting start date. Additional supporting documents and analyses were made available during the meeting. All documents are listed in Appendix 1.
- The assessment for the stock was clearly presented and supported by clear documentation, including detailed descriptions of the input data and especially that from the research survey series, and included appropriate coverage of the main uncertainties. The assessment analyst and other participants fully engaged with the review in a highly professional and constructive manner. The assessment presentations were supported by clear and informative presentations on the Gulf of Alaska Bottom Trawl Survey, MACE Acoustic Survey, the North Pacific Observer Program, and the Gulf of Alaska Age and Growth Program, all of which were key to the understanding of the assessment being reviewed.
- The assessment model was age-structured and the design and implementation appropriate for the biology of Pacific ocean perch, the type, scale and extent of the fishery, and the available catch, abundance, and composition data. A particular strength of the assessment is the availability of a consistent timeseries of biomass estimates from the Gulf of Alaska Bottom Trawl Survey timeseries (in particular since 1996), including high quality age data from both the survey and fishery.
- The assessment appropriately considered and addressed the main uncertainties in the data and the assumptions necessary to develop and implement the model.
- During the review process, the CIE reviewers identified some minor shortcomings in the assessment. None of the identified issues were considered of major importance by this reviewer. There was open discussion about identified issues and, where possible, additional model outputs were produced and reviewed during the meeting.
- This stock assessment for Pacific ocean perch represents the best available science and exceeds the acceptability threshold for scientific and technical quality to be used for informing management. Given the assessment, the current management approach and the scale and intensity of the fishery, there are no current sustainability concerns for the stock of Pacific ocean perch. The stock is not experiencing overfishing, is not overfished, and projections indicate that, at assumed catch levels, this will likely remain the case for the short- to medium-term.

- A part of the review meeting time was spent on understanding the temporal-spatial nature of the fishery and survey data and the application of novel approaches to analyze these data and incorporate them into the assessment. Despite this, the review of this component could only be relatively superficial given the scale and complexity of the task and the on-going development of the approach.
- The applicability of the acoustic data generated by the MACE survey to provide an additional abundance index for Pacific ocean perch for use in future assessments was a specified item to be reviewed during the meeting. There was a lack of convincing evidence that the acoustic backscatter from the MACE survey could reliably and robustly be decomposed into its component parts to provide an estimate of the abundance of Pacific ocean perch given the mix of species found. This outcome strongly argues against accepting an acoustic timeseries from the MACE survey as an index of abundance for Pacific ocean perch until these issues have been fully explored and a robust and reliable index has been convincingly demonstrated.
- Specific recommendations aimed at improving both the input data and the stock assessment for Pacific ocean perch, as well as some general processes, are made by the reviewer as required by the reviewer Terms of Reference provided by the CIE.

Background

This review of the 2020 Stock Assessment Report for the Gulf of Alaska (GOA) Pacific ocean perch stock was conducted as part of an independent review for the Center for Independent Experts (CIE).

All views expressed in this report are solely those of the independent CIE reviewer.

The key assessment reports, for assessments in 2017, 2019 and 2020, together with supporting background documents and reports, were comprehensive, well written and clearly presented. The support provided by the local NMFS staff, particularly that provided by Paul Spencer (the meeting Chair), Pete Hulson (stock assessment analyst) and Jim Ianelli, was of a high standard and much appreciated by the reviewer. The reviewer appreciated the high quality and informative presentations by the various NMFS staff, particularly including the pre-recording of a video about the MACE Acoustic Survey and voice descriptions attached to some of the PowerPoint presentations.

The fishery for Pacific ocean perch has been experiencing a managed, long-term recovery since the mid-1990s as the stock has recovered and catch limits have been adjusted. The recent four-year catch average has been about 24,000 t (2016-2019).

The meeting Chair, Paul Spencer, ran the meeting, facilitated discussions within the group, and ensured appropriate support was provided, as required. The various presenters provided clear and informative background on their individual areas of expertise and responsibility for the review team, including fish biology, sampling, spatio-temporal data analyzes, ageing, and assessment. All presenters fully responded to questions raised by the review panel. This CIE reviewer considered all of the documents and presentations provided. All CIE reviewers asked questions of clarification and engaged to offer alternative approaches where they considered such approaches would lead to insights on or improvements in the assessment.

Description of Review Activities

This review was undertaken by Geoff Tingley between the 14th March and 15th April 2021. The timing of the virtual review meeting proceeded as scheduled from 30th March to 1st April 2021 PST. The review meeting was conducted using the Webex software platform, with CIE reviewers joining remotely from Canada, South Korea and New Zealand. The agreed agenda, was broadly followed, with the meeting finishing a little early most days. The additional time set for Friday morning was a sensible contingency arrangement but, in the event, was not needed for this review.

The supporting documentation for the review of the assessment were provided to the reviewers in electronic format adequately in advance of the review meeting. These documents included the current and two historic stock assessments for Pacific ocean perch, previous CIE review reports and a recorded video presentation on the MACE acoustic survey. Additional relevant documents detailing aspects of the stocks, sampling and other related science necessary for a full understanding of these fisheries and their assessments were also made

available electronically before and during the meeting as and when it became clear that these may be of use. Electronic copies of the various presentations and additional work conducted during the review meeting were also provided. The reviewer also accessed additional, publicly available reports relevant for understanding the assessment and supporting the review. All documents provided and used are listed in the Bibliography (Appendix 1).

All documents provided in advance of the meeting were reviewed prior to the start of the meeting and the assessment was reviewed against the specific, Terms of Reference (ToR) provided by the CIE in the Performance Work Statement (Appendix 2).

Information relevant to this review is presented in three appendices to this review report, as required by the ToR. These are, Appendix 1: Bibliography of documents; Appendix 2: CIE Performance Work Statement (which includes its own annexes describing the (1) Peer review report requirements, (2) the ToR for the peer review, and (3) the draft agenda for the review meeting; and Appendix 3: Panel membership and other relevant information and decisions, including the agenda (as agreed at the start of the meeting).

An on-line (Webex) draft agenda was provided in advance of the meeting. At the start of the meeting the agenda was discussed, and minor, appropriate, and agreed changes were made to the agenda to accommodate the participation of some additional scientific staff during the review. The agenda was agreed. The meeting was conducted in an open, friendly and constructive manner throughout. Presentations were made with questions of clarification asked by members of the panel. All discussions were professional and good natured, being focused on clarification and clarity around the assessment under review. Other than the reviewers, presenters, and other supporting scientific staff, there were no other meeting attendees.

Additional output from model runs requested during the meeting, as well as responses to panel questions, were made available to the reviewers as soon as completed, and all reviewer requests were responded to before the end of the meeting or by email shortly afterwards.

None of the three CIE reviewers reported any major issues or concerns about the assessment during the three days of the meeting.

Summary of Findings

The assessment report was well written and together with the supporting documentation included virtually all the information necessary to support the review.

Focusing on a single assessment, as was done for this review, enabled a more comprehensive review than is possible when multiple assessments on different stocks or species are conducted.

Areas of the assessment where one or more reviewer considered there was opportunity to improve on the assessment approach were explored during the meeting. These are discussed in some detail below, and where appropriate, recommendations have been made.

Additional model runs for the assessment were developed during the review meeting. This testing of the assessment enabled reviewer concerns to be explored and clarification of whether those concerns were justified or not. Some areas where improvements could be made in future assessments were identified and are reported below. This approach also clearly demonstrated that the assessment was robust and of a high quality, representing the best available science and fully appropriate to use as a basis for providing management advice.

The types, amount and quality of data available to assess Pacific ocean perch are more than sufficient to enable an assessment of very high quality to be developed. There remain a small number of areas where some improvements can be made. Issues for future assessments for this stock to address include some early survey data suitability and an issue associated with incomplete spatial coverage of the normal survey area in one year.

The Gulf of Alaska Bottom Trawl Survey was found to be a high-quality data source, providing fishery-independent abundance (biomass) information. This has been developed with considerable care, with tow distribution planned to account for key fish species density distribution, depth, main habitat types, trawlable/untrawlable ground, and differences in gear and survey vessels within the timeseries. The earliest two survey points (1984 and 1987) have already been dropped from the bottom trawl survey abundance index based on concerns that the survey methodology for these two data points was sufficiently different from that used in later years (different vessels, different months, and different tow durations) and that these two points would not represent part of the same timeseries as later points. The reviewer concurs with this decision.

Two other areas of potential inconsistency in the survey timeseries were discussed, with some exploration though additional model runs, and recommendations made. These were reviewing the data points from 1990 and 1993 for time-series consistency, and reviewing how the abundance estimate was developed for 2001 when some of the normal survey area was not surveyed. The reviewer concluded that there was a basis for excluding these datapoints from the timeseries for future assessments.

Recommendations for research and development work for future assessments for this stock were considered and discussed. Some of the recommendations made with regard to the assessment of Pacific ocean perch are also relevant for assessments for other Gulf of Alaska stocks, and where this occurs these issues are noted under general recommendations.

By the end of the review meeting, the approach to modeling the Gulf of Alaska stock of Pacific ocean perch had been thoroughly explored, was considered thorough and sound, and appropriately addressed uncertainty to the principal assumptions through the range of models and sensitivities explored. The ranges of input data available and used were clearly described.

The overall outcome of this assessment, as reviewed, is that it meets the description of best available science and exceeds the acceptability quality threshold to be used to inform management.

Addressing the Terms of Reference for the Peer Review

Detailed findings and recommendations are presented below, as required by the ToR for the review.

Gulf of Alaska Pacific ocean perch

1. *Evaluate the data used in the assessments, specifically trawl survey estimates of biomass, and recommend how data should be treated within the assessment model*

The data used in the Pacific ocean perch assessment were generally all of a high standard.

Trawl Survey

The quality statement above is particularly true of the data collected by Gulf of Alaska Bottom Trawl Survey. The development approach of the abundance index for Pacific ocean perch from the survey data is also generally sound, with two minor exceptions. These exceptions pertain to the years 1990 and 1993, and to 2001, and are specifically addressed under Section 4.

The Pacific ocean perch fishery is reported to have started in the early 1960s when it was dominated by foreign flagged vessels, with very large catches made in the mid-1960s. This assessment used catch data from 1961 to 2020. There is, clearly, some uncertainty about the reliability of the early years catch data and any concerns about the influence that these data may have on assessment outcomes should be evaluated by running sensitivities to plausible alternative catch histories.

Catch History

Sensitivities to plausible alternative catch histories, particularly for the early years of the fishery, should be run when there are substantive changes to the assessment model structure or assumptions.

Composition Data

There are substantive amounts of composition data available, both from the survey timeseries and from the commercial fishery. There has been full observer coverage of this fishery since about 2008, and thus the composition data after this date are expected to be broadly representative of the fishery. The current assessment uses age-frequencies from 1990 to 2018 (though not for every year) derived using an age-length key. Length composition data are sparingly used in this assessment, with data from the fishery only incorporated in the assessment from the periods 1963–1977 and 1991–1997, with no length composition data from the survey used.

Considerable efforts have been applied in this assessment to ensure that the composition data have not been over-weighted, a surprisingly common issue in assessments. While the fits to these data, as evidenced by the various plots of age-and length-frequency fits and also residual plots, are acceptable, there does still appear to be some element of over-weight of these data in this assessment.

It would be appropriate to continue to explore different approaches to the appropriate weighting of the composition data, by using different statistical approaches but possibly also by careful quality control of these data, excluding data of known poorer quality.

At a future assessment it is recommended to try and incorporate all of the high quality length composition data from both the survey and the commercial fishery.

Plus Group

The 2020 assessment used a plus group at 25 years old. While this may have been appropriate for earlier assessments, with the managed reduction in fishing mortality and subsequent rebuilding of the stock, the numbers of older fish in both the survey and the fishery have been increasing. This can be seen in the steady progression in the proportion of the sampled fish in the plus group from 2004 to 2018, which is now approaching that seen in the samples from the fishery in 1990 (Figure 1). At some point an older plus group will become appropriate and this point may have already been reached. It is, therefore, appropriate to explore whether the plus group should remain at 25 or be increased. This exploration should be done in advance of, or as part of, the next assessment for Pacific ocean perch.

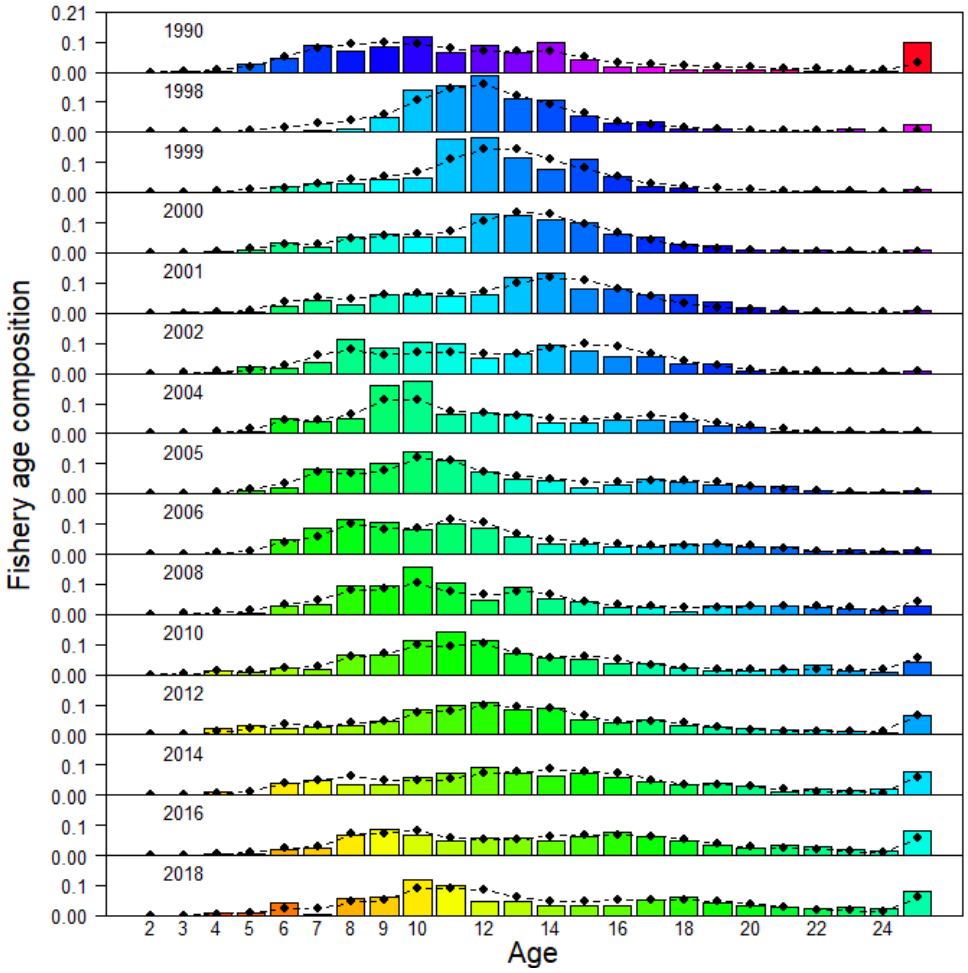

Figure 1: (Figure 9-2) Fishery age compositions for GOA POP. Observed = bars, actual age composition predicted from author recommended model = line with circles. Colors follow cohorts.

2. *Evaluate the stock assessment model for GOA Pacific ocean perch in general and comment on appropriateness of parameter estimates to assess stock status determinations*

The assumptions about stock structure are reasonable and appropriate.

The estimate of natural mortality, *M*, increased from the 2019 assessment (0.065) to this one (0.075). This was enabled by changes to the prior for *M*. Due to the influence of *M* on the productivity and the estimated biomass (the higher *M* is, the greater are the estimated productivity and estimated biomass). Therefore, understanding the uncertainty in the value of *M* for Pacific ocean perch and reducing this uncertainty is critical to ensuring that the estimated biomass is not unknowingly mis-estimated. Defining a realistic range of natural mortality for a stock is typically challenging, especially as the available data often contain little information on *M*. This is likely to be the case for Pacific ocean perch. The concern about possible mis-estimation of the population size of Pacific ocean perch is informed by the slightly degraded fit of the model to the most recent four bottom trawl survey biomass estimates, where the model consistently underestimated the biomass when compared to the survey estimates (Figure 2).

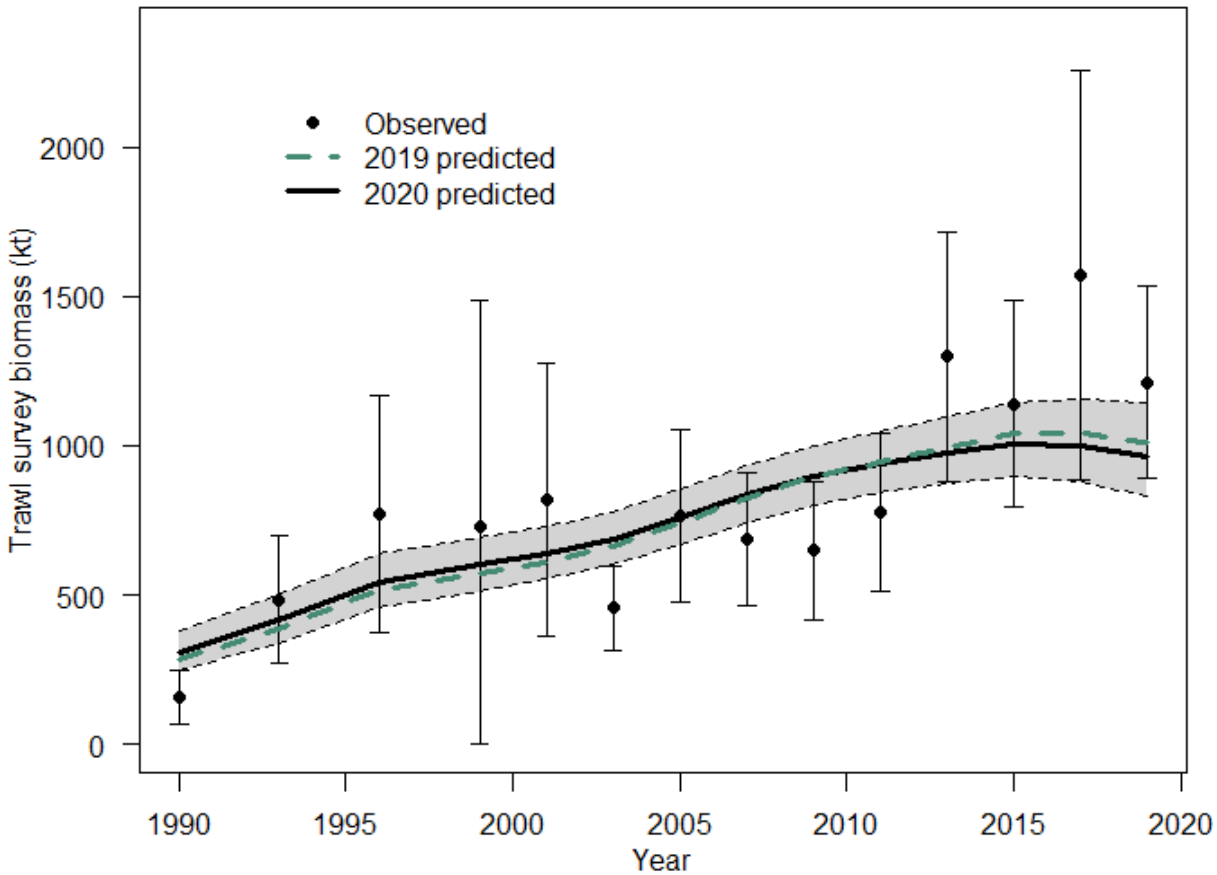

Figure 2: (Figure 9-4) NMFS Groundfish Survey observed biomass estimates (open circles) with 95% sampling error confidence intervals for GOA POP. Predicted estimates from the recommended model (black line, with 95% confidence intervals shown in grey shaded region) compared with last year's model fit (green dotted line).

While not conclusive, this is an indication that some caution may be required, and a need for further investigation of the most appropriate range of *M* is called for. Until this can be completed, managers should be provided with sensitivities using values of fixed *M* that bracket the estimated *M* to inform on the level of risk inherent in the current assumptions.

Given the lifespan of Pacific ocean perch and other rockfish, coupled with the historic overfishing and recent recovery in stocks, it is likely that some additional information is available to inform in the estimation of *M*. This information is likely to be contained in the right-hand limb of the age frequencies (Cordue, 2014).

There may be sufficient information available in more recent data, not only for Pacific ocean perch but also for a number of other long-lived rockfish, that can materially inform the estimation of the value of natural mortality.

The application of a *q*-prior in order to estimate catchability was a significant improvement over the approach used in earlier assessments.

There was good consideration given to the selectivity parameterization, including the need to be cautious when domed selectivities were indicated, so as to avoid generating any substantive cryptic biomass of older (larger) fish. This should, however, remain an area to monitor during future assessments.

There was some discussion about the interaction of *q* and selectivity, especially in relation to changes over time and whether time-varying *q* or time-varying selectivities would be more appropriate. Experience of this from New Zealand is that time-varying *q* for surveys simply undermines the value of the survey, in that the survey data are effectively down-weighted. Introducing time-varying selectivities does not have the same impact on the weight of the data, but there should be a requirement for some evidence of a shift in selectivity before really considering to apply this. The bigger picture here, however, is that even if some careful consideration of time-varying parameters may enable better model fits to be achieved, this still completely misses the important and otherwise ignored spatial component of these parameters. A possible solution to this challenge is to try and use the available spatio-temporal models, such as VAST, to interpret the complex of survey and fisheries data.

3. *Evaluate the strengths and weaknesses in the stock assessment model for GOA Pacific ocean perch, and recommend any improvements to the assessment model.*

The core strengths of this assessment are the sound basis of available data, especially that from the bottom trawl survey. In addition, there are substantive quantities of composition data available from both the survey timeseries and from the commercial fisheries. That the sampling from the commercial fisheries has been substantively comprehensive over recent years makes the majority of the commercial composition data likely to be representative of the fishery.

An important strength in the assessment is the scale and reach of the historic and on-going investment in scientific research. This provides background information that enables appropriate and informed consideration on data choices, data handing choices, and supports the rational development of necessary assumptions to underpin the assessment. For example, in the development of statistical approaches to the proper weighting of the composition data within the assessment, and current and on-going development of spatio-temporal approaches

to better understand and utilise the fisheries and survey data, where there are inherent spatial and temporal patterns that have previously been ignored in assessments.

The assessment also included the presentation of a range of informative diagnostics were provided to enable exploration of the nature and quality of the fit of the model to the various datasets. This strengthens confidence in the model.

The apparent weaknesses seen in this assessment are common to the vast majority of other high quality stock assessments worldwide. These are almost all determined by a lack of understanding of key parameters and interactions, often driven by limited data. For example, stock structure, appropriate values for *M*, and *h*, and estimation of catchability and selectivity. None of these are critical weaknesses or even weaknesses of particular importance in this assessment. Moreover, addressing these weaknesses has followed a rational and sensible approach within the development of the stock assessment. Principally this has included using what knowledge is available (e.g., in the development of priors), making only necessary and justified assumptions, and testing the sensitivity of the assessment to these different components.

Where some specific options exist for seeking to address identified weaknesses, these are detailed in the appropriate areas of the this report. However, overall, this is a very robust stock assessment.

4. *Evaluate and recommend how survey data are used for biomass indices within the assessment. Specifically, advise on trawl survey indices arising from design-based methods versus model-based approaches.*

As this reviewer has already noted the high quality of the Gulf of Alaska Bottom Trawl Survey, this section will focus on areas of possible, generally marginal, improvements.

This reviewer has already noted and agreed with the exclusion of the survey data points from 1984 and 1987.

It is further noted that the next two data points in the timeseries (1990 and 1993) also had issues of lack of consistency in methodology with the rest of the timeseries. These differences included the use of different vessels and used a longer tow duration of 30 minutes. The tow duration was reduced to 15 minutes in the 1996 and subsequent surveys but the implication of this change on catchability in general and for Pacific ocean perch in particular was not investigated. Previously, given a shorter timeseries of the Gulf of Alaska Bottom Trawl Survey, the inclusion of these early data points was both sensible and justifiable, but that with a longer and more consistent timeseries, this justification has become progressively weaker though time. Given this, this reviewer requested an additional model run dropping these two survey points (abundance and composition data) from the assessment. This additional run showed a good fit to the most important of the fitted data, the trawl survey abundance index (Figure 3), which supports dropping these two years of survey data, but a detailed investigation of the residuals and goodness of fit for other data was not possible in the time available. However, as there is good evidence of a lack of methodological consistency, and there are now many years of later, higher quality data, excluding these two points from future assessments would be a justifiable and rational decision.

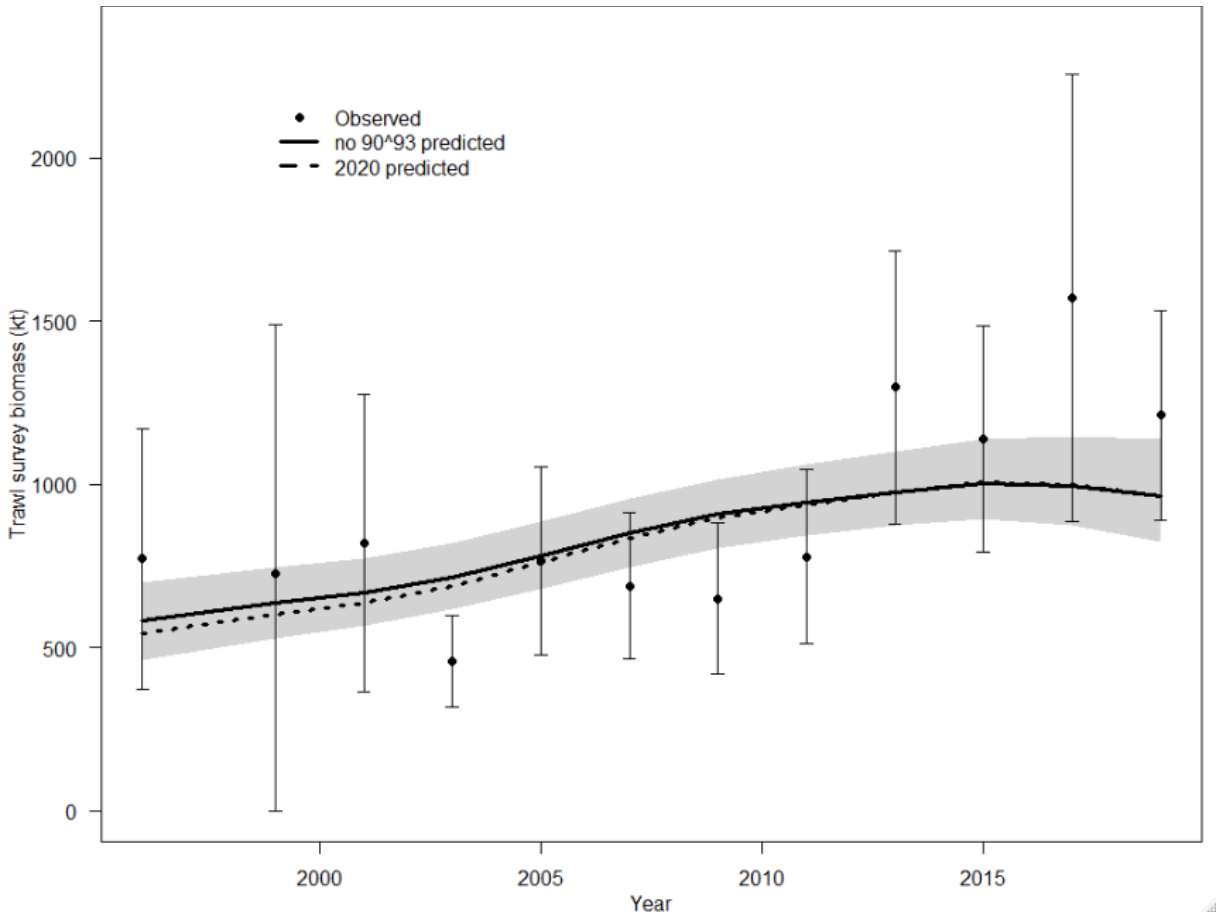

Figure 3: Model fit to the bottom trawl survey index, including (dashed line) and excluding (solid line) the 1990 and 1993 survey data points (a model run conducted during the review meeting at the request of the reviewer).

At this point in time, with increasing numbers of post-1993 surveys, it is probably appropriate to drop these two years (1990 and 1993) of survey data from the Pacific ocean perch assessment in order to reduce the uncertainty introduced by these data points. If considered desirable, these data points could be retained for sensitivity runs.

In 2001, the Gulf of Alaska Bottom Trawl Survey did not cover the extreme eastern part of the normal survey area. The approach used to address this lack of data was to estimate the 2001 missing data using data from the three prior surveys (Heiftz et al., 2001). While completely justifiable at the time, this approach does create additional and undefinable uncertainty in this component of the input data. This reviewer is strongly of the opinion that this is no longer the best, or even an acceptable approach, as it relies on an assumption that the source data to estimate the missing data are drawn from a broadly similar population (abundance, density distribution, age and size distributions) as the unsampled area in 2001. While this may be a tenable assumption for an annual survey, it is far less likely to be plausible for a biennial survey, as the Gulf of Alaska Bottom Trawl Survey was at that time. The time between surveys would permit considerable change in abundance, density distribution and in the age and size profiles of Pacific ocean perch available to the survey (e.g., Figure 4). Moreover, only data from earlier surveys were used, which would likely have introduced greater bias than changing to a mix of pre- and post-2001 data at some point after 2001, especially with an increasing

population, as was the case for Pacific ocean perch at that time. There are a number of alternative approaches that do not require estimating the missing data as per (Heiftz et al., 2001) and that would likely yield a better and more consistent outcome. Three of these more preferable, alternative approaches are described.

Proposed options for dealing with the missing eastern area data in the 2001 Gulf of Alaska Bottom Trawl Survey:

> Option 1: Completely drop the 2001 Gulf of Alaska Bottom Trawl Survey data from the assessment. This is an improvement on the current approach to this issue and would probably be the easiest approach. However, it does result in the loss of part of an important data series and may be more problematic should the 1990 and 1993 data points also be dropped, as discussed above. One way forward would be to run a sensitivity dropping the 2001 survey data from this year and assess whether this creates any difficult or unacceptable issues. This is the least preferred of these three proposed options.
>
> Option 2: The Gulf of Alaska Bottom Trawl Survey is divided into three spatial components, west, central, and east. It is then possible to develop the survey data as three separate, area-based indices of abundance and fit these separately within the model, dropping the data for 2001 from the eastern index only. This eliminates the need to use poorly defined data in the east and maximizes the retention of the high-quality data for west and central areas available from the survey timeseries in 2001. This would allow the missing data from the area not surveyed in 2001 to affect only the eastern area index for one year and represents a small part of the survey data series, with most of the survey data from 2001 retained within the west and central area indices, the two areas that also hold the bulk of the biomass of Pacific ocean perch. The implications of splitting the survey into three separate timeseries for the assessment would need to be explored in some detail. This is preferred to Option 1 and may be a practical option to be applied in the next assessment.
>
> Option 3: use one of the spatial-temporal data analytical approaches available (e.g., VAST) to develop a timeseries that lacks data for the unsurveyed area during the 2001 survey. Depending on how the development of the spatio-temporal analyses of the Pacific ocean perch data proceed, this may well be the best long-term option for developing area-based indices for Pacific ocean perch but is probably less practical than Option 2 in the short- to medium-term due to the substantive work required to developed and apply these spatio-temporal models.

It is recommended that the current approach of estimating the missing eastern data from the 2001 Gulf of Alaska Bottom Trawl Survey is discontinued for all future assessments of Pacific ocean perch. It is also recommended that one, or more, of the alternative approaches proposed are employed to reduce the uncertainty and improve confidence in the next assessment. Given the limited time between now and the next assessment, Option 2 is the recommended approach but if time is limited, Option 1 will still yield an improvement over the current approach. Option 3 is likely to take considerably more time to implement than is

currently available but should be explored in future as part of the general development and implementation of the spatio-temporal modeling approach.

The application of spatio-temporal models to the survey and fishery data shows considerable promise, even though this development is at a relatively early stage. This perspective is also supported by the early application of spatio-temporal models in other regions, including in New Zealand. The application of the VAST modelling framework to re-estimate the biomass index from the Gulf of Alaska Bottom Trawl Survey yielded a marked increase in abundance in recent years. However, without adequate diagnostics, it was not possible to evaluate whether this was an appropriate model to use or not. There remains an inherent risk of overestimating the stock biomass and trajectory by using an approach that is neither well understood nor testable in the usual sense by reviewing meaningful diagnostics. This is of specific concern as the fishery-independent Gulf of Alaska Bottom Trawl Survey abundance data are providing the principle driver of the stock assessment. The application of VAST, or other spatio-temporal models, to understand and interpret both survey and fishery data is completely appropriate and should absolutely continue. Substantive questions remain, however, about how to evaluate these approaches and whether they should be used in base case models at present or only in sensitivities to inform future development.

5. *Evaluate abundance estimates from summer acoustic-trawl data, and recommend how it may be used within the assessment.*

There was a detailed presentation on the MACE Acoustic Survey and the acoustic and other sampled data collected. The largest proportion of the backscatter was reported as coming from pollock, with Pacific ocean perch providing the next largest proportion. The analyses presented focused on obtaining an annual biomass estimate for Pacific ocean perch derived from the acoustic backscatter. Separating the acoustic backscatter attributable to Pacific ocean perch from that attributable to the other fish species was explored using the multiple frequency data available, unfortunately, as the relative frequency response for rockfish was very similar to that for pollock, making acoustic differentiation of the two commonest taxa impossible at present. Thus, separating the acoustic backscatter attributable to Pacific ocean perch from that attributable to the other fish species relied on the species composition from the fish identification tows taken during the survey. While this is a not an uncommon practice in ground-truthing acoustic surveys, this approach is mostly used for stocks that are both highly aggregated and virtually single species aggregations, neither of which really occur for Pacific ocean perch. The approach taken was to use the nearest tow to a specific fish aggregation, which may result in some individual tows being used more than once in defining aggregation species composition, an outcome that may introduce some bias into the estimates. The identification tows also occurred at a variety of different distances from the areas of backscatter that they are being used to decompose into Pacific ocean perch and other species biomass data.

If the spatial and temporal distributions of Pacific ocean perch were relatively constant, then some of the concerns expressed here about the relationship between the trawl and aggregation species compositions may be overstated. However, there is sufficient evidence to show that for Pacific ocean perch these distributions are highly variable, both within and

between years (Figure 4), sufficient to raise genuine doubts about the applicability of the methodology.

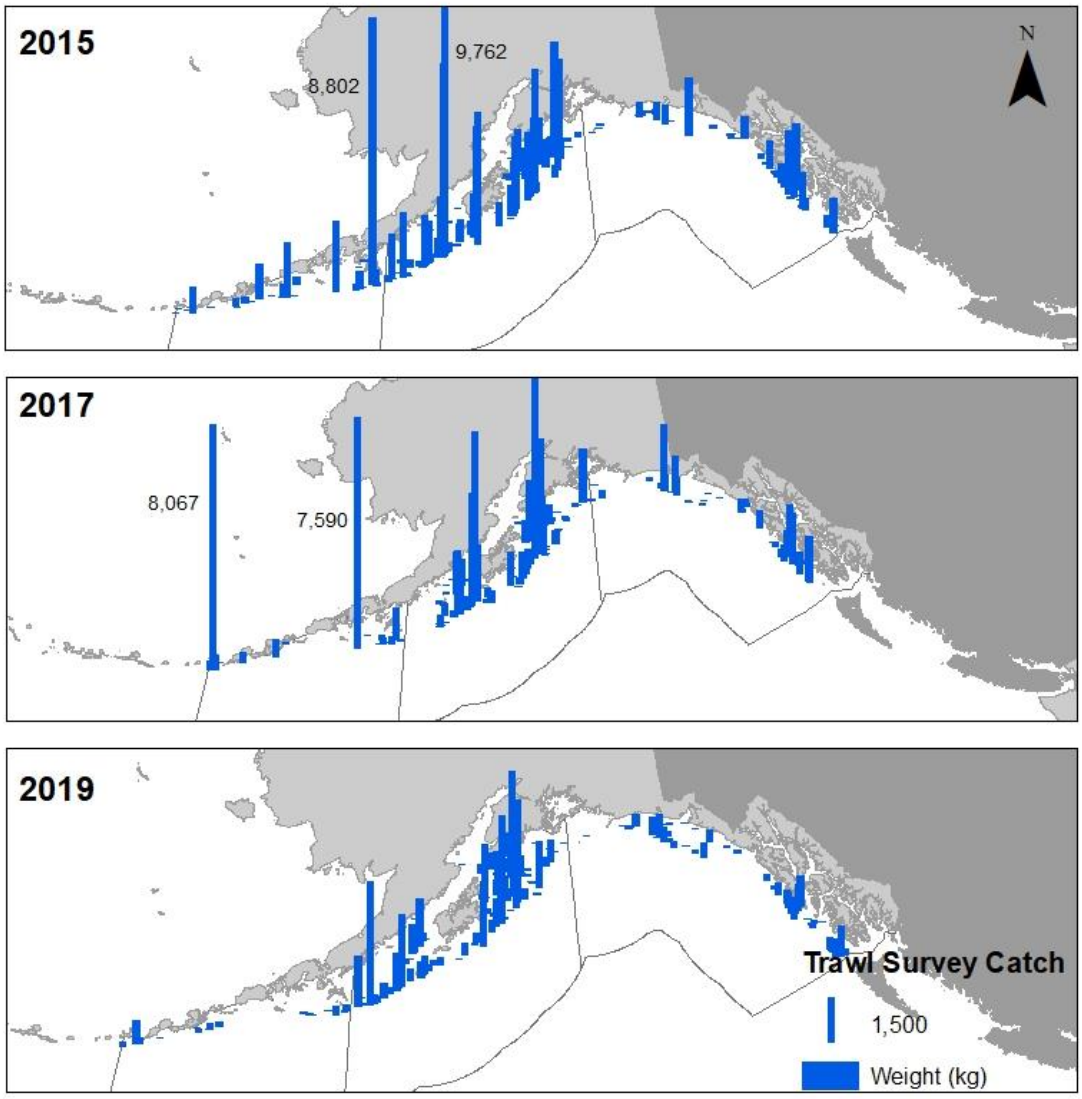

Figure 4: (Figure 9-5) Distribution of GOA POP catches in the 2015-2019 GOA groundfish surveys.

No evidence was presented that the identification tows used were adequate or appropriate to make the inference on species composition in a fish aggregation at variable distances acceptable. No evidence was presented to demonstrate that the different distances between the tow and the backscatter being decomposed made no difference to the species composition and biomass outcomes. No evidence was presented to show how consistent any of these differences were between years. These issues almost certainly contain a substantial proportion of the total uncertainty associated with developing an acoustic abundance index and this needs to be adequately quantified.

Also implicit in the simple assumption that the catch composition in the trawls represents the species composition in the acoustic backscatter is a series of assumptions about catchability. Essentially there are a series of assumptions about catchability (sampling efficiency) of the two sampling methods, trawl and acoustic sampling. These assumptions include, for example,

that the relative catchabilities for the main sources of backscatter in the acoustic signal are similar to those of the fish in the trawl gear (at whatever distance the tow is made from the fish aggregation being decomposed), and that the relative catchabilities are either spatially and temporally invariant or they vary in proportion. It is highly unlikely that these assumptions about catchability are valid.

To be convincing as a high-quality dataset to be used as part of a stock assessment, all of these identified, potential shortcomings would need to be explored and quantified. At present there is too much uncertainty inherent in the methodology to consider that a timeseries developed in this way would be a meaningful abundance index for Pacific ocean perch.

There are a number of possible ways to explore the relationship between the acoustic and trawl data in a simple way using the currently available data. For example, for each area of backscatter analysed, a simple substitution of the current trawl station used to provide the species composition (the closest tow location) with another one from further away, either the next nearest or a randomly selected one could be made. This substitution may provide information about how repeatable the comparison is and also the impact distance has on the predicted aggregation species composition and on the biomass estimates.

An alternative approach to evaluating the reliability of using the trawl species composition to decompose the acoustic backscatter data would be to define a minimum acceptable distance between the location of paired datasets for acoustic backscatter and trawl data. Various minimum distances could be defined and the impact of increasing this distance on the predicted aggregation species composition and the on biomass estimates explored.

Alternative methods using only acoustic data to decompose the backscatter data may also become available in the future, although the similarity of the relative frequency responses for pollock and rockfish is not encouraging. For example, Ryan and Kloser (2016) developed an approach using multiple lines of evidence to quantitatively estimate biomass of orange roughy in spawning aggregations. This was principally based on separating the backscatter from fish with swim bladders from those without swim bladders using duel-frequency acoustics, supported by a number of other analyses.

As the intention here is to seek to provide a fishery-independent abundance index that will be a principle driver of future stock assessments, the quality threshold for accepting an acoustic index should be set very high, otherwise the overall quality of any assessment that uses the data will likely be impaired.

It is recommended that attempts to develop an acoustic abundance index for Pacific ocean perch for use in stock assessments should be discontinued until the evidence-base supports a substantially increased likelihood that the processed acoustic backscatter represents a reliable abundance index for Pacific ocean perch. Currently, there is insufficient evidence that the acoustic data can be reliably and consistently decomposed into Pacific ocean perch and other species solely using the species composition data from the trawl tows.

It is recommended that the existing MACE acoustic and trawl data are further explored in detail to ascertain whether the backscatter data can be reliably and robustly be decomposed into Pacific ocean perch and other species or not.

Peer Reviewer Recommendations for Gulf of Alaska Pacific Ocean Perch

General Recommendations

- Filling data gaps by creating, extrapolating or interpolating input data should, as a general rule, be avoided as this is likely to introduce undefined uncertainty or bias into the results, and the age-structured modeling framework is designed to handle such gaps in datasets. This recommendation specifically applies to how a lack of some spatial coverage in the 2001 datapoint from the Gulf of Alaska Bottom Trawl Survey was dealt with in estimating the biomass for this survey point.
- Consider using detailed, fine-scale commercial spatial location data to assist in defining the relationship between trawlable and untrawlable ground in the Gulf of Alaska Bottom Trawl Survey and to provide detailed spatial information for future spatio-temporal models.

Recommendations for Pacific ocean perch

1. *Evaluate the data used in the assessments, specifically trawl survey estimates of biomass, and recommend how data should be treated within the assessment model*
    - Sensitivities to plausible alternative catch histories, particularly for the early years of the fishery, should be run, but only when there are substantive changes to the assessment model structure or major assumptions.
    - Continue to explore different approaches to the appropriate weighting of the composition data, by using different statistical approaches but possibly also by careful quality control of these data, excluding data of known poorer quality.
    - At a future assessment, it is recommended to try and incorporate all of the high-quality length composition data from both the survey and the commercial fishery, at least in a sensitivity.
    - Prior to or as part of the next assessment, explore whether the plus group should continue to start at age 25 or whether an older plus group starting age is more appropriate.

2. *Evaluate the stock assessment model for GOA Pacific ocean perch in general and comment on appropriateness of parameter estimates to assess stock status determinations*
    - Exploration of additional information to better define the realistic range of *M* for Pacific ocean perch is recommended. This should consider data available for Pacific ocean perch and for other long-lived rockfish species.

3. *Evaluate the strengths and weaknesses in the stock assessment model for GOA Pacific ocean perch, and recommend any improvements to the assessment model*
    - In the absence of better information about the likely magnitude of *M*, sensitivities using values of fixed *M* that bracket the estimated value *M* should be run in future stock assessments to inform on the level of risk inherent in the current assumptions about *M*.

4. *Evaluate and recommend how survey data are used for biomass indices within the assessment. Specifically, advise on trawl survey indices arising from design-based methods versus model-based approaches.*

   - Continue to exclude the 1984 and 1987 survey biomass estimates and survey composition data from all future assessments as these are clearly not part of the longer survey timeseries due to the use of differences in vessels, trawl gear, tow duration and survey timing.
   - Exclude the 1990 and 1993 Gulf of Alaska Bottom Trawl Survey biomass estimates and the survey composition data from all future Pacific ocean perch (and other species) assessments (or include them only in sensitivities, possibly including them as a separate timeseries). These two years do not appear to be part of the longer survey timeseries due to different timing, tow duration and survey structure.
   - It is recommended that the current approach of estimating the missing eastern data from the 2001 Gulf of Alaska Bottom Trawl Survey is discontinued for all future assessments of Pacific ocean perch and that one of the proposed approaches, or an alternative approach, is used so as to reduce uncertainty in the next assessment.
   - Continue to support the development and application of spatio-temporal models (such as VAST) for use in stock assessments. In order to make this effective, there need to be a rapid development of a suite of informative diagnostics for spatio-temporal models in a fisheries stock assessment context. Until such time as suitable diagnostics are available, it is recommended that these spatio-temporal models are only used in sensitivity model runs and not in the base case from which management advice is developed.

5. *Evaluate abundance estimates from summer acoustic-trawl data, and recommend how it may be used within the assessment.*

   - It is recommended that attempts to develop an acoustic abundance index for Pacific ocean perch from the MACE Acoustic Survey data for use in assessments should be discontinued until the evidence base supports a substantially increased likelihood that the processed acoustic backscatter represents a reliable abundance index for Pacific ocean perch.
   - It is, however, also recommended that the existing MACE acoustic and trawl data are further explored in detail to ascertain whether the backscatter data can be reliably and robustly be decomposed into Pacific ocean perch and other species or not.

Appendix 1: Bibliography

Pacific Ocean Perch Assessment Documents

Hulson, P-J.F., Hanselman, D.H.,Lunsforn, C.R. and Fissel, B. (2017). Assessment of the Pacific ocean perch stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE (2017) 913-922.

Hulson, P-J.F., Hanselman, D.H.,Lunsforn, C.R., Fissel, B. and Jones, D. (2019). Assessment of the Pacific ocean perch stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE (2019) 78p.

Hulson, P-J.F., Lunsforn, C.R., Fissel, B. and Jones, D. (2020). Assessment of the Pacific ocean perch stock in the Gulf of Alaska. NPFMC Gulf of Alaska SAFE (2020) 79p.

Previous CIE Reviews

Dichmont, C. (2013). AFSC Rockfish Assessment Review. Juneau, Alaska 9-11 April 2013 56p.

Klaer, N. (2013). CIE Reviewer's Independent Report on Alaska Rockfish Stock Assessment. Alaska Fisheries Science Center, Ted Stevens Marine Research Institute, 17109 Pt., Lena Loop Rd, Juneau, AK 99801 April 9-11, 2013 32p.

Kupschus, S (2013). CIE Review of Alaska Rockfish Assessments, 9-11 April 2013, Alaskan Fisheries Science Center, Juneau, Alaska 36p.

Additional Papers Provided

Jones, D.T., Rooper, C.N., Wilson, C.D., Spencer, P.D., Hanselman, D.H., and Wilborn, R.E. (2021). Estimates of availability and catchability for select rockfish species based on acoustic-optic surveys in the Gulf of Alaska. Fisheries Research 236.

https://doi.org/10.1016/j.fishres.2020.105848

Thorson, J.T (2019). Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research, 2010, Pages 143-161. https://doi.org/10.1016/j.fishres.2018.10.013

Thorson, J.T, Cunningham, C.J., Jorgensen E., Havron, A., Hulson, P-J.F., Monnahan, C.C. and von Szalay, P. The surprising sensitivity of index scale to delta-model assumptions: Recommendations for model-based index standardization. Fisheries Research 233. https://doi.org/10.1016/j.fishres.2020.105745

Thorson, J.T., Shelton, A. O., Ward, E.J., and Skaug, H.J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science, 72: 1297-1310.

Additional Papers Sourced

Conrath, C.L and Knoth, B. (2013). Reproductive Biology of Pacific Ocean Perch in the Gulf of Alaska, Marine and Coastal Fisheries, 5:1, 21-27, DOI: 10.1080/19425120.2012.751941. https://doi.org/10.1080/19425120.2012.751941

Cordue, P.L. (2014). The 2014 orange roughy stock assessments. New Zealand Fisheries Assessment Report 2014/50. 135 p. https://fs.fish.govt.nz/Page.aspx?pk=113&dk=23684

Heifetz, J., Courtney, D.L., Clausen, D.M., Fujioka, J.T. and Ianelli, J.N. (2001). Slope Rockfish. https://apps-afsc.fisheries.noaa.gov/refm/docs/2001/GOAsloperck.pdf

Ryan, T. E., and Kloser, R. J. (2016). Improved estimates of orange roughy biomass using an acoustic-optical system in commercial trawlnets. ICES Journal of Marine Science, 73: 2114–2126. https://academic.oup.com/icesjms/article/73/8/2112/2198257

Wetzel, C.R. and Cronin-Fine. L. (2017). Status of Pacific ocean perch (Sebastes alutus) along the US west coast in 2017. https://www.pcouncil.org/documents/2017/11/agenda-item-f-4-attachment-1-2.pdf/

Appendix 2: Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
Virtual External Independent Peer Review

Gulf of Alaska Rockfish – Pacific ocean perch

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 stock assessment for Gulf of Alaska Pacific ocean perch provides the scientific basis for the management advice considered and implemented by the North Pacific Fisheries Management Council. An independent review of this integrated stock assessment is requested by the Alaska Fisheries Science Center's (AFSC) Auke Bay Laboratories Division (ABL). 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 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 and recent experience in the application of statistical age-structured

stock assessment methods in general and, in particular, assessments developed with software such as ADMB.

Additionally, the CIE reviewers shall have:

- Expertise with measures of model fit, identification, uncertainty, forecasting, and biological reference points;
- Familiarity with federal fisheries science requirements under the Magnuson-Stevens Fishery Conservation and Management Act;
- Familiarity with groundfish fisheries and management;
- Working knowledge of the application of spatio-temporal models to population index estimation;
- Experience with application of acoustic data collection within stock assessment;
- Excellent oral and written communication skills to facilitate the discussion and communication of results.

Tasks for Reviewers

1. Review the following background materials and reports prior to the review meeting. Two weeks before the peer review, the NMFS Project Contact will make all necessary background information and reports available electronically 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. 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 panel review meeting. The meeting will consist of presentations and discussions with the stock assessment authors, NMFS observer program staff, and survey scientists to facilitate the review. After the 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.

4. Each reviewer should assist the Chair of the meeting with contributions to the summary report, if required in the terms of reference.

5. Deliver their reports to the Government by the specified deadline.

Place of Performance

This review will be conducted via virtual meeting software.

Period of Performance

The period of performance shall be from the time of award through April 14 2020. 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.

Within two weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
March 30-April 1, 2021	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.

Project Contact(s):
Pete Hulson
Auke Bay Laboratories
NMFS, Alaska Fisheries Science Center
17109 Point Lena Loop Rd., Juneau, AK, 99801
Phone: 907-789-6060
pete.hulson@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

1. Evaluate the data used in the assessments, specifically trawl survey estimates of biomass, and recommend how data should be treated within the assessment model
2. Evaluate the stock assessment model for GOA Pacific ocean perch in general and comment on appropriateness of parameter estimates to assess stock status determinations
3. Evaluate the strengths and weaknesses in the stock assessment model for GOA Pacific ocean perch, and recommend any improvements to the assessment model.
4. Evaluate and recommend how survey data are used for biomass indices within the assessment. Specifically, advise on trawl survey indices arising from design-based methods versus model-based approaches.
5. Evaluate abundance estimates from summer acoustic-trawl data, and recommend how it may be used within the assessment.

Annex 3: Tentative Agenda
Gulf of Alaska rockfish – Pacific ocean perch

TBD

March 29-31, 2021

NMFS Point of contact: Pete Hulson (pete.hulson@noaa.gov)

Appendix 3: Panel membership and other pertinent information from the panel review meeting

Panel Membership

Name	Role and Affiliation	
Paul Spencer	Chair	ASFC, NMFS
Noel Cadigan	CIE Reviewer (Canada)	
Saang-Yoon Hyun	CIE Reviewer (Republic of South Korea)	
Geoff Tingley	CIE Reviewer (New Zealand)	
Peter-John Hulson	Presenter – Stock Assessment Lead	ASFC, NMFS
Darin Jones	Presenter – MACE (Acoustic) Survey	ASFC, NMFS
Chris Lundsford	Presenter – Rockfish Management	ASFC, NMFS
Ben Williams	Presenter – Fishery Overview	ASFC, NMFS
Wayne Palsson	Presenter – GOA BT Survey	ASFC, NMFS
Jason Connor	Presenter – VAST	ASFC, NMFS
Tom Holland	Presenter – NP Observer Program	ASFC, NMFS
Chris Gburski	Presenter – Age & Growth	ASFC, NMFS

Other attendees

Jim Thorson	AFSC (Habitat & Ecological Processes Research), NMFS
Patrick Ressler	AFSC (Resource Assessment and Conservation Engineering), NMFS
Denise McKelvey	AFSC (RACE), NMFS
Madison Hall	AFSC postdoc (RACE), NMFS
Jim Ianelli	ASFC, NMFS
Kari Fenske	AFSC, NMFS
Dana Hanselman	AFSC, NMFS
Cindy Tribuzio	AFSC, NMFS
Dan Goethel	AFSC, NMFS
Kristin McQuaw	Alaska Groundfish Data Bank, Oregon
Julie Bonney	Alaska Groundfish Data Bank, Alaska

Other pertinent information from the panel review meeting

Following discussion with the reviewers about the need for a summary (panel) report, as referred to in the CIE reviewer TOR, the Chair stated that, following checking, no summary (panel) report was necessary and none would be prepared for this meeting. This is relevant to Peer Reviewer Report Requirement 2c; noting that, in the absence of a summary report all areas of interest have been fully covered in this individual report.

Agenda (as agreed at the start of day 1)

Gulf of Alaska Rockfish – Pacific Ocean Perch

March 30 - April 1, 2021
Time Zone PDT
Virtual (Webex) Meeting

Monday, March 29, 2021

10.00-10.30	Pre-meeting connectivity and software test (as required in the review ToR).

Daily breaks at 11:30 and 15:45, lunch 13:00-14:00

Tuesday, March 30, 2021

10:00-11.30	Introductions and background including roles and working arrangements
	Introductions and agenda (Paul Spencer)
	Overview of rockfish biology, fishery, and history of assessment (Ben Williams)
	Current management of Alaska rockfish (Chris Lunsford)
11:45-13.00	Discussions
14:00-15:45	Input data
	Survey data
	Abundance, distribution, and age composition (Pete Hulson, Wayne Palsson)
	Model-based abundance (Pete Hulson, Jason Conner)
	Fishery data - Catch, observer program, ages, lengths (Pete Hulson, Tom Holland)
	Age determination, lengths, maturity, and growth (Pete Hulson, Chris Gburski, Delsa Anderl)
16:00-17:00	Discussions

Wednesday, March 31, 2021

	Field-based catchability
10:00-11.30	Assessment model (Pete Hulson)
	Model structure, likelihoods, data weighting, parameter estimates, data fit, diagnostics
11:45-13.00	Discussions
14:00-15:45	Parameters, priors, diagnostics (Pete Hulson)
	Catchabilities, selectivities, model fits, diagnostics
16:00-17:00	Discussions

Thursday, April 1, 2021

	Pre-recorded presentations to review: Acoustic survey
10:00-11.30	Model developments

	Incorporation of acoustic information (Pete Hulson, Darin Jones)
	Incorporation of model-based index (Pete Hulson)
	Internal review model scenarios (Pete Hulson)
11:45-13.00	Discussions
14:00-15:45	Requested topics/model runs
16:00-17:00	Summarize, revisit Terms of Reference

Friday, April 2, 2021

10:00-14.00	Additional time, as needed