Center for Independent Experts (CIE) Independent Peer Review Report for BSAI Rockfish Pacific Ocean Perch (POP)

Matthew D. Cieri

August 2022

Executive Summary

The peer review of Pacific Ocean Perch in the Bearing Sea Aleutian Islands was conducted remotely using WebEx from May 11th through May 13th, 2022. All information was made available online, including presentations and recorded presentations, before the meeting started. This stock is managed using quotas that are calculated using a well-formed harvest control rule that accounts for both scientific and management uncertainty.

Data used in this assessment included two fishery-independent surveys, age and length information from those surveys as well as the fishery via onboard observers. Additional information was provided by maturity studies as well as other information on catches.

The model used in this assessment was a statical catch at age model with fixed growth parameters as well as a fixed aging error matrix. Estimated parameters included fishing mortality, abundance/biomass, maturity, natural mortality, and fishery selectivity as well as others. Of importance is the fact that fishery selectivity is estimated yearly with a bi-cubic spline approach. Recruitment was basically examined as mean recruitment over the time period with deviations, an important management consideration.

The model seems to perform well. Estimates of natural mortality and selectivity are reasonable, even if odd in the case of selectivity. Model results indicated that the stock was above its management target for biomass, and below its target for fishing mortality. Sensitivities indicated that most of the results are robust, though the CV derived from the MCMC analysis appears fairly large.

That said, the model has a major issue. There is a lack of fit to the surveys by the model at the end of the time series. This results in a retrospective pattern that tends to underestimate biomass and overestimate fishing mortality relative to the terminal year. The causes for this lack of fit are not well known. Downweighting the length and age compositions resulted in better fits to the surveys, but it required their age and length information to be completely removed before this was resolved. As this is not a good result given it is the same information that informs year-class strength, the issue remains unresolved. One possible fix is to input the surveys as abundance rather than the current method which uses biomass. But more work on this issue should be done in future assessments to explore this issue.

Despite this issue, the assessment as presented appears feasible for management use. The estimates of F and biomass are likely biased and are in a more precautionary direction. This suggests that the estimates derived from this model could be seen as generally a lower bound for biomass.

A number of recommendations were made to explore and perhaps improve the model. Of these better sampling by the onboard observers for non-dominate hauls, either conducting or replacing the AI survey index, is likely the most important.

Background

Introduction

The peer review of Pacific Ocean Perch (POP) in the Bearing Sea Aleutian Islands (BSAI) was conducted remotely using WebEx from May 11th through May 13th. Presentations, documentation, data files, modeling files, and even recorded presentations were available through either a web page or a good drive folder after a few minor technical issues with access.

Pacific Ocean Perch (*Sebastes alutus*) in this region are relatively slow growing when compared to other groundfish, spawning roughly at age 8-10 and continuing to grow until about age 40. POP live to approximately 100+ years and have low natural mortality.

Spawning occurs during the fall months, with the larvae born alive in the Spring and are planktonic. Fecundity is quite high from 10,000 to 300,000 larvae per spawning female. Settled juveniles then start migrations to deeper waters at about age 3. As such adults and juveniles tend to be spatially segregated. The diet of adults tends to be similar to pollock in the region, typically planktivorous consuming mysids, copepods, and other zooplankton. Though they rarely occur in gut-contents collections POP are preyed upon by seabirds, rockfish, salmon, lingcod, and large demersal fish as juveniles. Adult predators include sablefish, Pacific halibut, and sperm whales.

POP is generally harvested in deep water using a bottom trawl at about 200-300m depth. Distribution is patchy with schooling behavior dominating. Like many others of the genus *Sebastes*, POP tends to be semi pelagic and is often encountered in the midwater trawl pollock fishery.

Management is chiefly done by setting appropriate quotas which limit the direct harvest. As a Tier 3 stock, reference points include B40% as a proxy for B_{target}, $F_{35\%}$ and $F_{40\%}$ as the F at OFL and F at ABC respectively. Quotas at $F_{35\%}$ and $F_{40\%}$ are calculated and applied to the biomass as shown in Figure 1. This results in both ABC and OFL quotas. Retention is prohibited when the catch exceeds the OFL, and fishery closure occurs when the catches exceed ABC. The NPFMC (North Pacific Fishery Management Council) often sets a TAC lower than the OFL due to management uncertainty and ecosystem concerns. This TAC has not been exceeded since 2003, but in most years this quota is fully or near fully utilized.

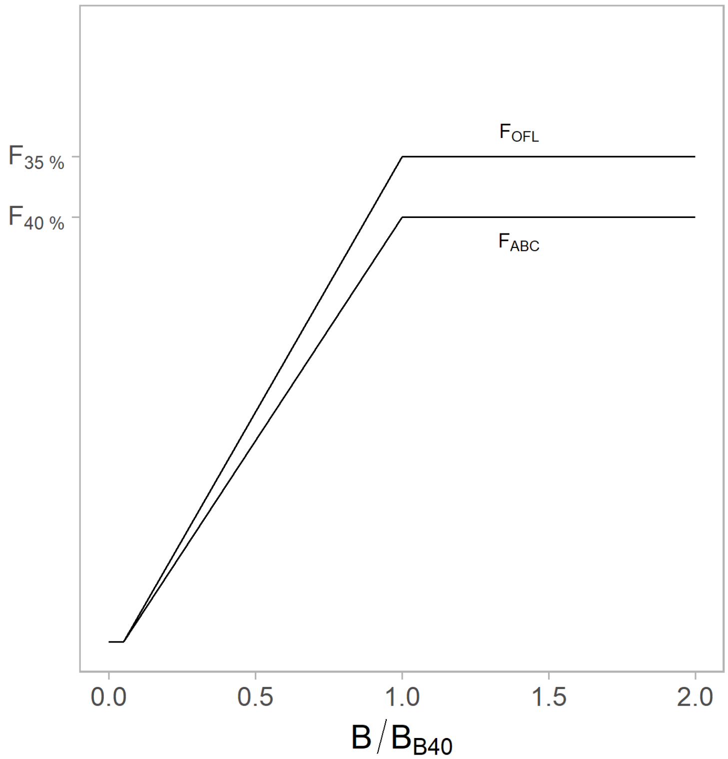


Figure 1. Harvest control rule for Tier 3 stocks in the North Pacific

Fishery dependent data

Fishery dependent data included overall catch by three areas; Western, Central, and Eastern Aleutian Islands (AI), which were summed to provide the overall catch of the stock. The total catch and catch by region can be seen in Figures 2 and 3. The large catches in the earlier period are dominated by the foreign fleet, while recently catches have been steadily increasing since about 2005. Catches are currently reported as a result of the mandatory logbook program, which is cross-checked with at-sea observers.

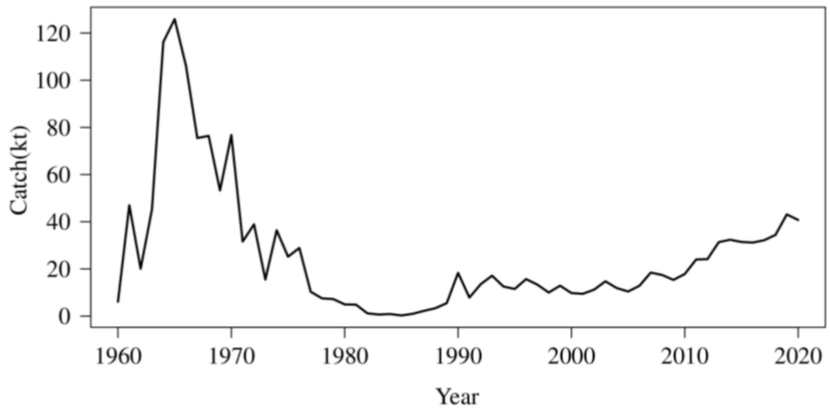


Figure 2. Catches of POP by year.

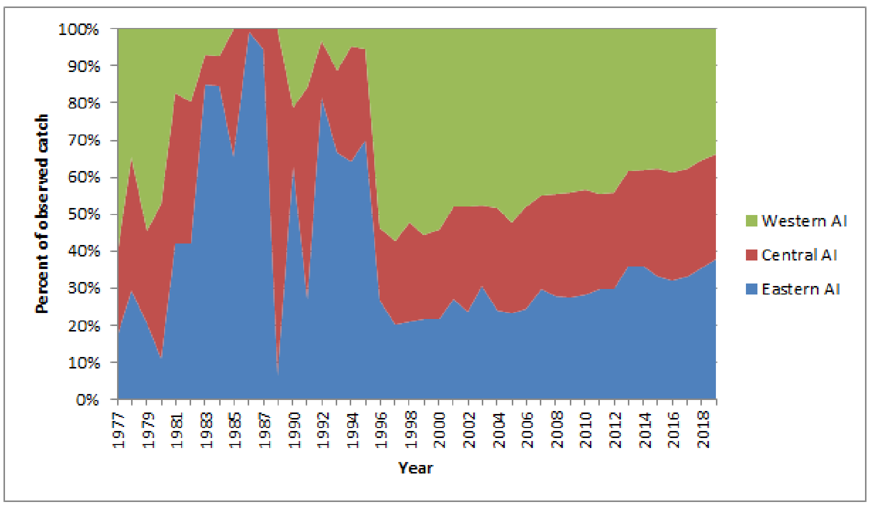


Figure 3. Catches as a percentage by area.

Fishery length and age are collected in the process of the 100% at-sea observer sampling program. It's notable that not all of the catch is sampled for lengths and ages per haul as shown in Table 1, only the top three species have samples collected on any given haul.

Table 1. Sampling priority of the at-sea observer monitoring in the BSAI region.

Rockfish * Species Ranking List 1. Pacific Ocean Perch 2. Northern Rockfish 3. Shortraker/ Rougheye 4. Thornyheads 5. Dusky Rockfish	**Every Sampled Haul** ~ 16 of the most predominant species in the list, chosen by rank in cases of equal predominance **and** ~ 8 from the second most predominant species on the rockfish Species Ranking List **and** ~ 4 from the third most predominant species on the rockfish Species Ranking List	**Every Sampled Haul** 3 otolith pairs from the ~ 16 s/l rockfish **and** 2 otolith pairs from the ~ 8 s/l rockfish **and** 1 otolith pair from the ~ 4 s/l rockfish	**Every sampled haul** ~ 10 Viability or Injury Assessments **for pelagic trawl hauls, record viability as unknown**

This can lead to cases where POP may not be well sampled, particularly in cases where it is not one of the dominant species encountered. This in turn could potentially bias the sampling, particularly if individual fish school according to size/age. Where older/larger individuals are in smaller schools, encountered as one of the less dominant species in a haul.

Further examination revealed that catch by the target of the haul, as measured by the percentage of the species in that haul, indicates a change in which fleets operating in the area catch POP (Figure 4).

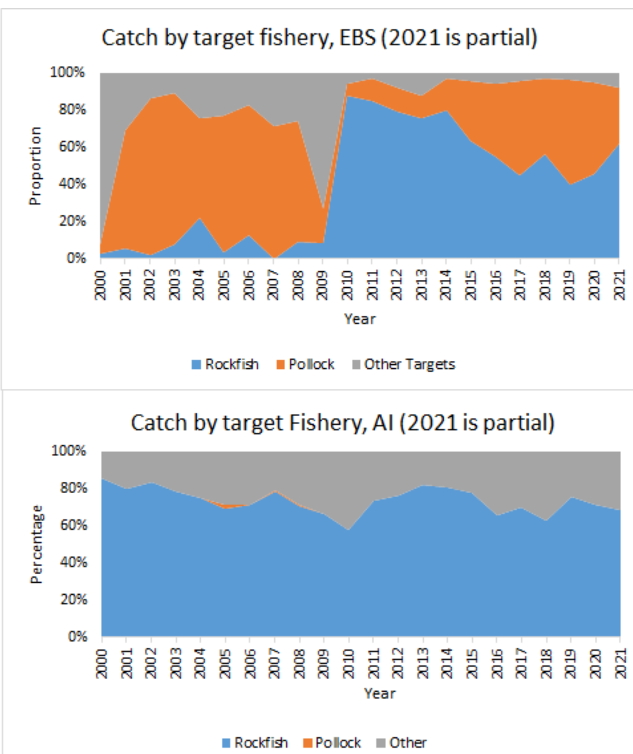


Figure 4. Catch by target fishery as a percentage in the Eastern Bearing Sea (EBS) and Aleutian Islands (AI).

Given that the pollock fishery tends to use more midwater trawls than bottom trawls, this is an issue that should likely be addressed. A more thorough recommendation will be made later in this document.

Ages and lengths are then analyzed back onshore. Two readers conduct aging to allow for an aging error matrix to be used as input into the model. As can be expected there is some drift in readers as the fish age (Figure 5). There has been an increase in both the number of age/length samples taken and the number of otoliths read over time as the catch has grown (Table 2).

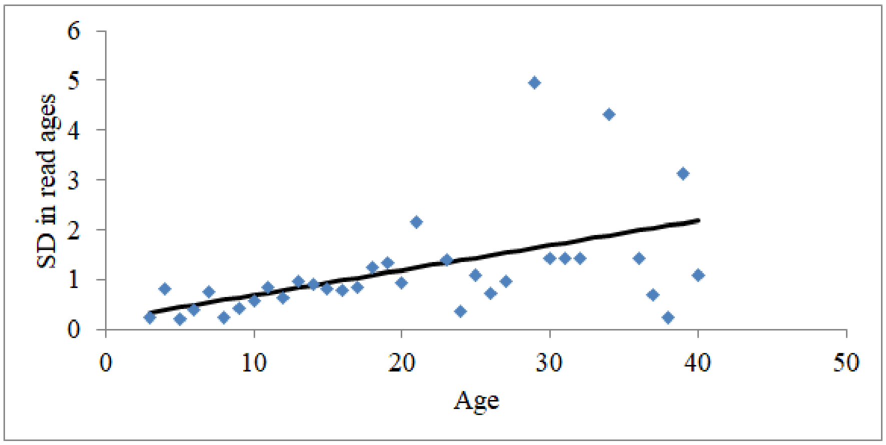


Figure 5. SD of aging by the two readers by age

Table 2. Otoliths collected and read by year 1981-2019.

Year	Collected otoliths	Read otoliths	Hauls (read otoliths)
1981	682	510	33
1982	226	222	16
1990	346	328	26
1998	848	823	67
2000	539	487	110
2001	576	524	113
2002	704	466	165
2003	1,171	397	210
2004	787	784	173
2005	591	581	126
2006	929	449	194
2007	1,362	675	275
2008	1,737	572	305
2009	1,438	719	294
2011	2,810	932	513
2013	4,213	1,043	661
2015	4,072	1,016	661
2017	3,292	1,080	771
2019	4,399	1,093	853

Prior to input into the model, a global Age-Length Key (ALK) is developed across the whole fishery. Examination during the review meeting revealed that decomposing the ALK into areas produced similar results.

Maturation schedules were derived from two main studies conducted in 2004 (Shaw unpublished data) and 2010 (TenBrick et al., 2010). But showed similar results as can be seen in Figure 6, with maturation as a percentage of individuals progressing rapidly from age 6 to age 13.

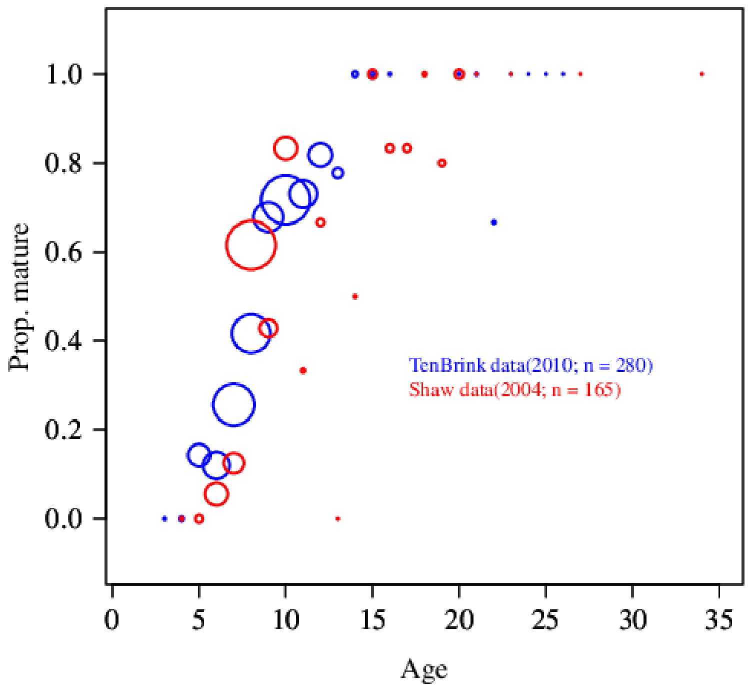


Figure 6. Maturation by age as determined by two studies conducted in 2004 and 2010.

Fishery independent data

Fishery independent data for the POP assessment hinged on the two trawl surveys covering the area, the Eastern Bering Sea (EBS) and the Aleutian Islands (AI) trawl time series. Both use the same gear type of trawl but have different configurations of doors and chains as they tend to fish in areas with different bottom types. Both surveys are non-overlapping yet show similar trends in their relative biomass over time (Figures 7 and 8).

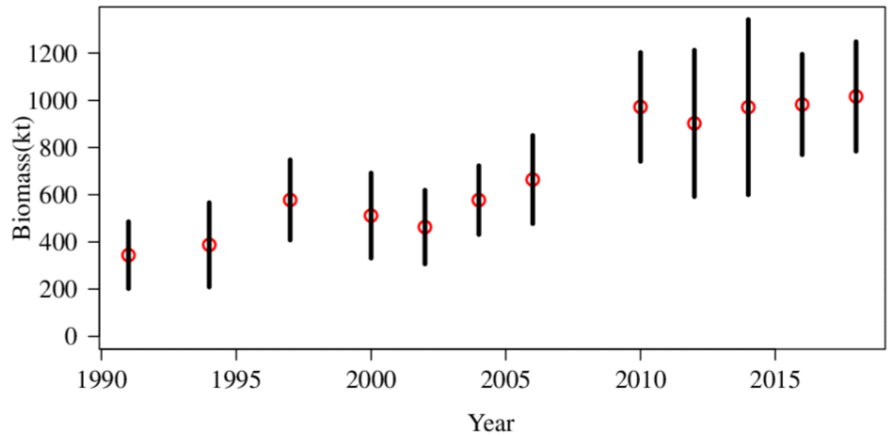


Figure 7. Eastern Bering Sea slope survey estimated biomass

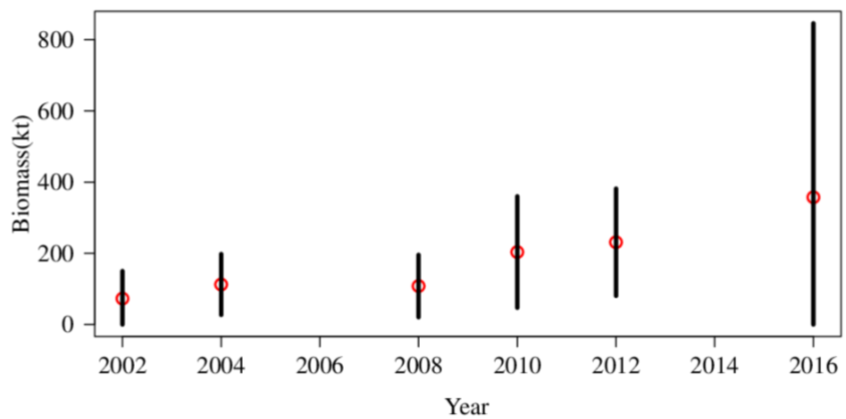


Figure 8. Aleutian Island Trawl survey biomass

Further examination during the review workshop indicated that these increases in biomass were consistent across depth and area. The number of positive tows also increased, suggesting that the population increase seen by both of these surveys was a result of population growth across all areas, rather than an increase in one area or at one depth.

In addition, in the survey biomass, both surveys also provide age and length data that are useful to the modeling structure.

Data conclusions

The various data sources suggest an interesting and complex series of conclusions. Overall, this stock tends to grow and mature relatively rapidly for a *Sebastes sp.* With recruitment to the fishery and the surveys by age 3, a 50% spawning probability by age 10 and rapid growth from age 3 to 20. There appears to be little difference in growth among years and cohorts. Similarly, an examination of the Von Bertlanffy parameters by area (Table 3) reveals a slight difference in the parameters for the western AI, but not a dramatic difference. Estimates of growth rate parameters from the fishery and from the fishery independent surveys as well as mean size at age revealed little differences.

Table 3. Von Bertlanffy parameters by area for survey and fishery

Area	n	Linf	K	tzero
SBS	784	41.95	0.15	-1.19
EAI	793	42.27	0.14	-0.99
CAI	786	42.06	0.15	-1.06
WAI	784	41.77	0.12	-2.02
EBS slope	2424	43.21	0.13	-0.68

Maximum age was also fairly consistent, at 100 plus or minus a few years. Additionally, the population increases in the survey were echoed by the industry members in attendance. These individuals also indicated that the increases in abundance seem to be across the entire fishing area and at all depths, a finding in common with the results of both the AI and the EBS trawl survey.

An important piece of information given at the review workshop was the lack of sampling when POP was not a dominant species in the haul during the onboard observer sampling. This sort of sampling scheme can lead to sampling bias. Moreover, during the discussions, it would appear that a fairly large proportion of catches, at least 30% in some years are hauls targeting pollock. While there is no *a priori* reason to assume that there is a difference in the size or age of POP caught in targeted and non-targeted hauls is any different, this is an issue that needs more study.

A further concern is a lack of data from the AI survey post-2016. Discussions at the review workshop revealed that this survey is unlikely to be conducted at any time in the next few years. Given this is an important data stream for the overall assessment, a research recommendation will be made later in this report to continue or replace this survey with another data source.

Model structure

The model structure review was a forward projecting age-structured model with separable fishing mortalities, similar to a statical catch at age formulation. Such a model is regularly used in the region (i.e., POP in the Gulf of Alaska), and seems appropriate given the lack of growth changes across years. As such the model’s primary output currency was abundance at age.

Parameters estimated within the model as well as outside the model are given in Table 4 and Table 5.

Table 4. Parameters estimated within the model

Parameter type	Number
1) Fishing mortality mean	1
2) Fishing mortality deviations	61
3) Recruitment mean	1
4) Recruitment deviations	58
5) Unfished recruitment	1
6) Biomass survey catchabilities	2
7) Fishery selectivity parameters	25
8) Survey selectivity parameters	4
9) Natural mortality rate	1
10) Maturity parameters	2
Total parameters	156

Table 5. Parameters estimated or fixed outside the model

Parameter type	Number
Von Bertalanffy length at age parameters	3
Weight at length parameters	2
SD in length with age (polynomial)	5
Aging error (linear)	2
Proportion of stock in survey subareas	1

It is important to note that both fishing mortality deviations and recruitment had penalty terms associated with them to constrain the model. Likewise, a bi-cubic spline approach was used in yearly estimations of fishery selectivity. Priors for natural mortality and CV(natural mortality) were set at 0.05, as discussed below. Further, the McAllister-Ianelli weighing method was used with weighted the data sources in the model using effective sample size. A plus group of 40+ was used, and the recruitment to the fishery and surveys started at age 3.

Model results and diagnostics

Overall, the model showed that the stock appears to be in good health (Figures 9 and 10). F peaked in the late 1960s, declined rapidly through the mid-1980s, and has since slowly rebounded.

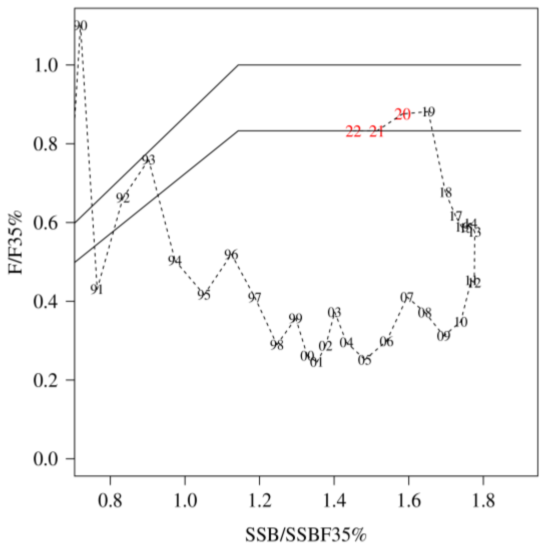


Figure 9. Phase-plane plot of the status of POP.

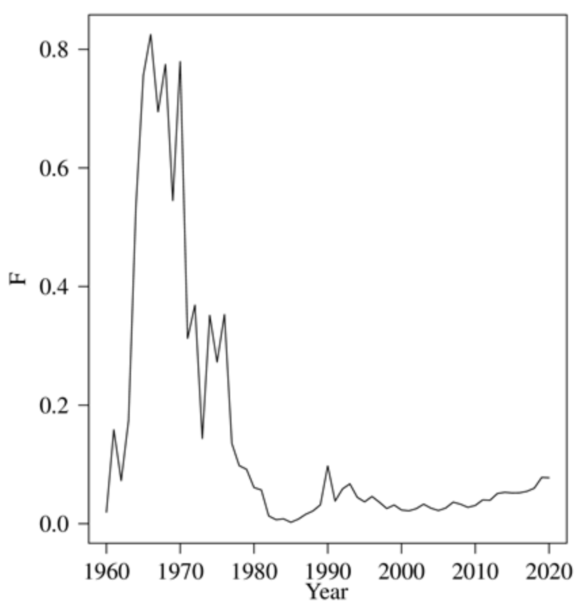


Figure 10. Fishing mortality 1960 to 2019.

However, the model does not seem to fit the surveys in the most recent years (Figure 11). Further examination of the likelihood profiles as well as dropping various data sources suggests some tension between the age and length composition data and the surveys (Figure 12). Because this lack of fit occurs only near the end of the survey time series, this results in a retrospective bias (Figure 13). In essence, each successive addition of survey data pulls the SSB further up while lowering the F. As such the bias could be thought of as conservative, with an underestimation of SSB and an overestimation of F relative to the terminal year.

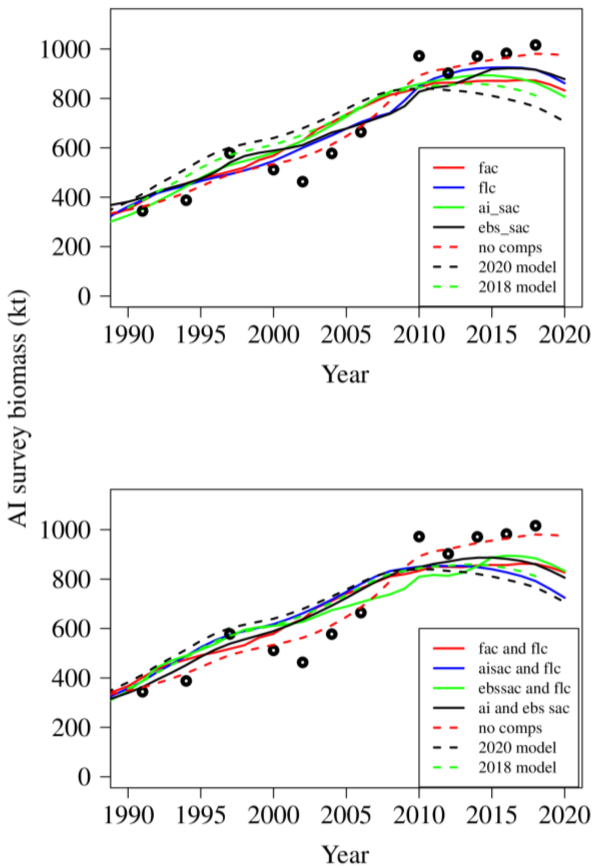


Figure 11. Model fits to the AI survey index with various data sources removed

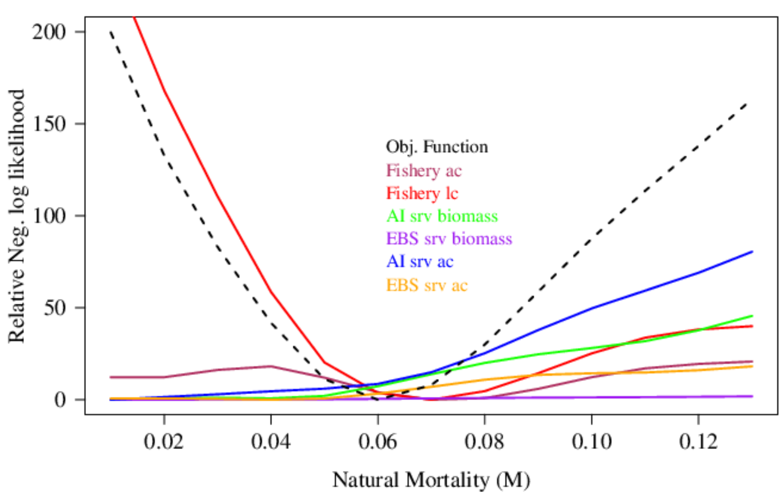


Figure 12. Likelihood profile for POP natural mortality by data source.

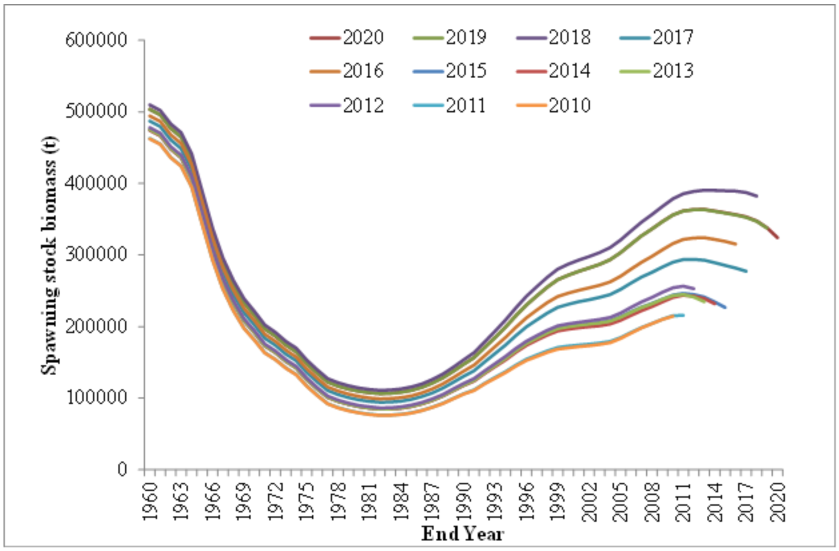


Figure 13. Retrospective pattern in SSB. Note surveys are only run every other year. Mohn's rho = -0.24

Error estimation and sensitivities

Estimates of uncertainty were produced via MCMC and showed fairly high CVs for SSB (Figure 14). However, even with the uncertainty the terminal year was well above $SSB_{35\%}$.

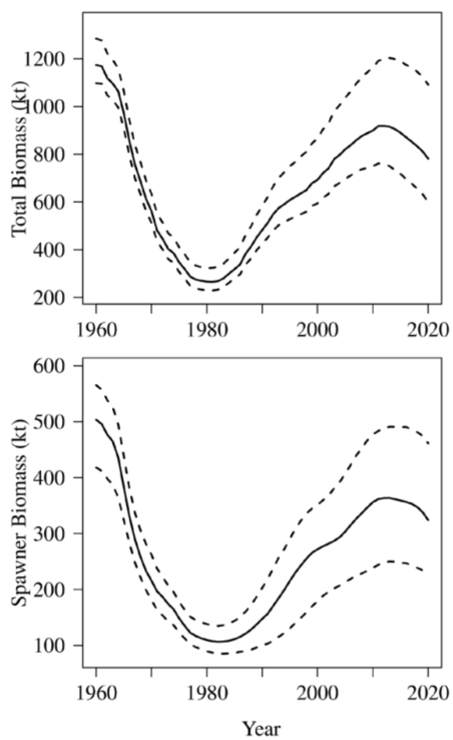


Figure 14. Total and spawner biomass for BSAI POP, with 95% confidence intervals from MCMC integration.

Sensitivities were conducted around Natural Mortality, Survey Catchability, and Data weightings.

For natural mortality, a total of five runs were completed including 1) 2020 base case, 2) Remove prior, 3) Blocks every 4 years, 4) 2 blocks, and 5) 4-year blocks, no prior. As stated earlier, M was estimated in the model with a 0.05 prior and a 0.05 CV on that prior. This was selected after examining the methods of Then et al (2015) across a range of maximum ages (Table 6).

Table X. Differing methods and maximum ages used to arrive at M.

:

Method	Model	Maximum Age 79	104	129
$Then_{1parm}$	$M = a/t_{max}$	0.065	0.049	0.040
$Then_{lm}$	$\log(M) = a + b \log(t_{max})$	0.067	0.051	0.041
$Then_{nls}$	$M = at_{max}^{b}$	0.090	0.070	0.057

Average across these methods and maximum ages equaled 0.059. With the prior, the M estimated for this stock was 0.056.

None of the aforementioned sensitivity runs indicated a better fit to the survey biomass (Figure 15)

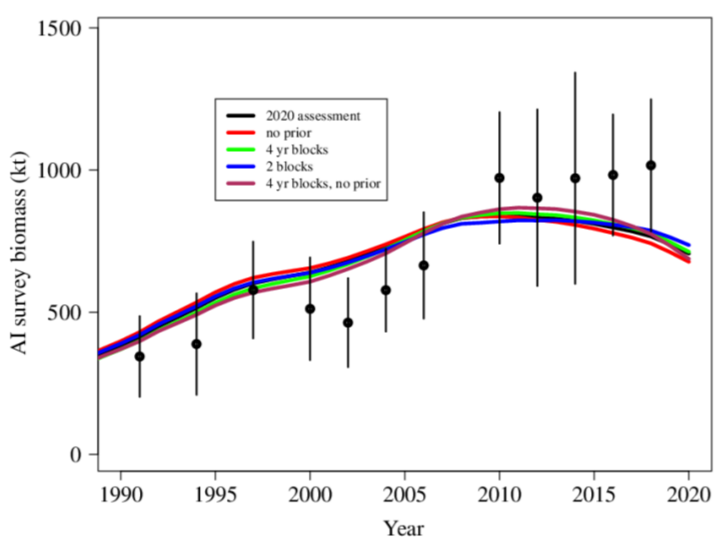


Figure 15. Fit to the AI survey across various sensitives of M.

As expected, these sensitivities to M did have a large impact on the estimate of SSB (Figure 16.)

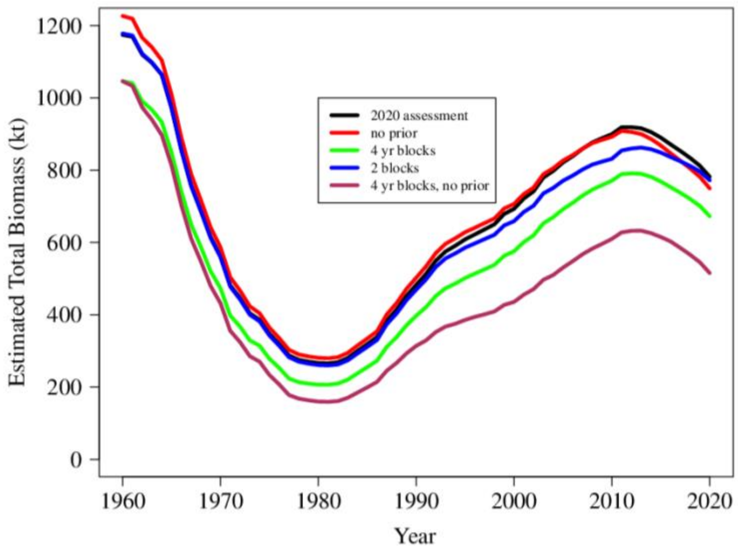


Figure 16. Estimated SSB across various sensitivities of M

However, it should be noted that the estimated M from each of these runs appears very different than the results of Then et al. (2015) (Figure 17).

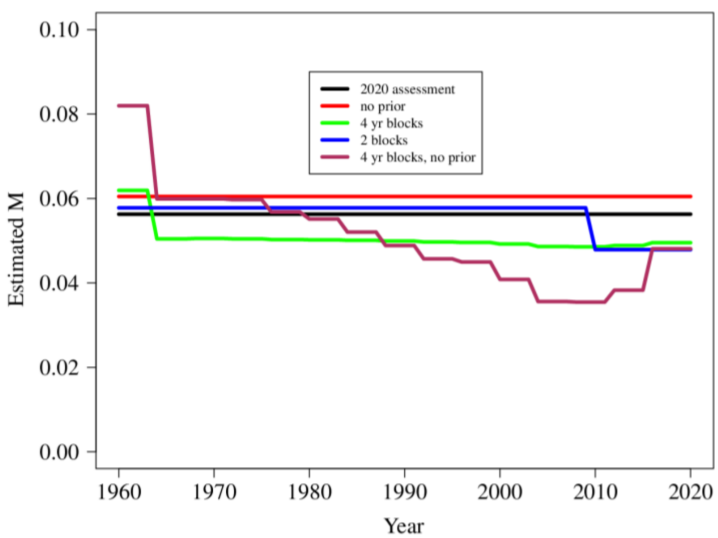


Figure 17. Estimated M across multiple sensitivities.

Given that POP is not a forage species, and doesn't appear to have dramatic changes in growth, coupled with the lack of predation information by predators on POP, arguments to include blocks or time-varying natural mortality are not parsimonious. This is particularly true in light of a lack of better fits to the surveys when using time-varying approaches.

As outlined further below, it seems that the prior on M, however, doesn't appear to be needed. The model apparently has enough information to make a fairly robust estimate of M that seems to be in keeping with the biology. A suggestion to run the final model with no prior on M is therefore encouraged given parsimony.

Additional sensitivities were run around the issue of survey catchability. These runs included: 1) 1) Base case (prior mean = 1.0), 2) prior mean for AI survey catchability = 1.15, 3) remove the prior, 4) remove prior and remove area proportion term, 5) blocks every 4 years, 6) blocks every 4 years, no prior, 7) 2 blocks. Effects on the biomass, survey fit, and the estimates can be seen in Figures 18 to 20 respectively.

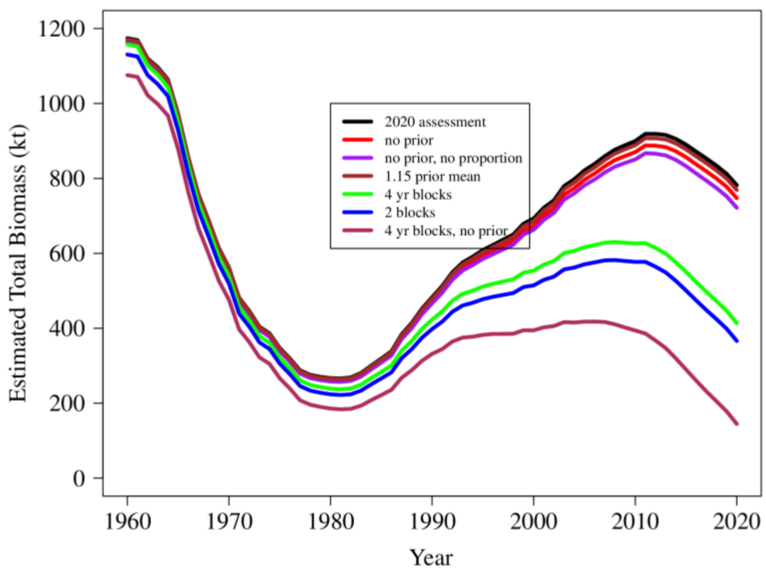


Figure 18. Estimated biomass by sensitivity run.

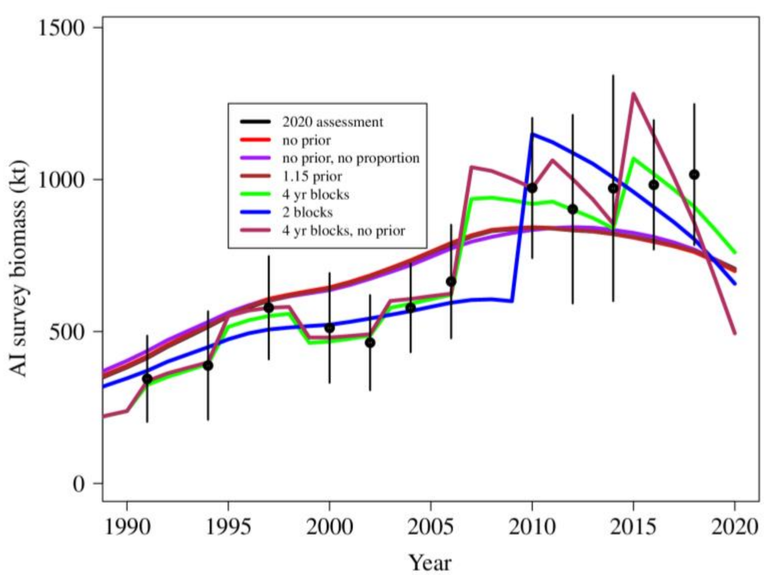


Figure 19. Fits to the AI survey by sensitivity run.

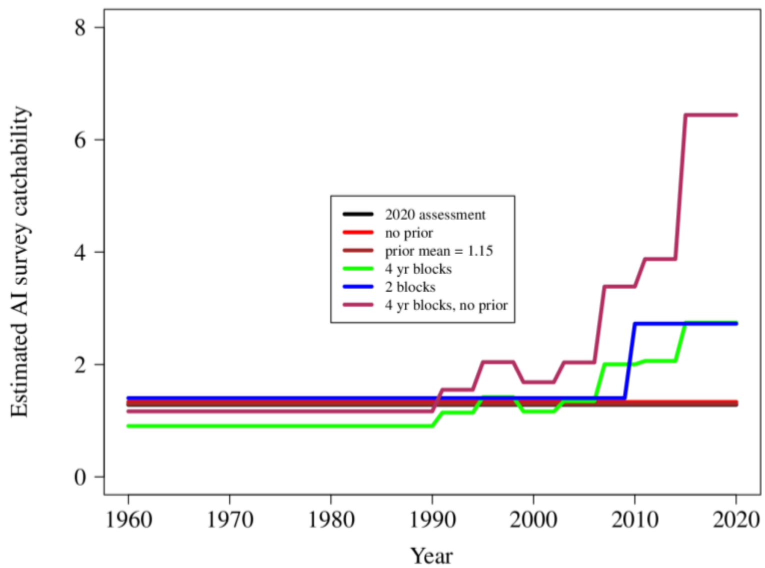


Figure 20. Estimates of catchability by sensitivity run.

The use of blocks in the catchability produced the greatest differences in the biomass estimated. While the use of either no prior, a prior of 1, and a prior of 1.5, produces similar biomass results. As can be seen, none of the sensitivities produced good fits for the AI survey with the sensitivities using blocks having the best fit at the expense of rather odd behavior. Likewise, the sensitivities with the blocks produced rather large estimates of q, well above what is18 reasonable.

While there could be some reason to indicate that q might be slightly higher than 1, there appears to be little evidence for it being much greater than 1. And there is further no evidence that q has changed much over time. However, like with natural mortality, the constraining prior doesn't seem to have a purpose.

Data weighting was the last issue in which sensitivities were carried out. Examined methods included the current McAllister-Ianelli method, the Francis weighting, and Dirichlet multinomial weighting. The McAllister-Ianelli method weights to the ratio of the effective sample size. The Francis weighting uses the inverse of the variance of the standardized deviations of the mean for both age and length. Meanwhile, the Dirichlet multinomial is similar to a regular multinomial but the probabilities for each age or length category are random variables that follow a Dirichlet distribution.

As can be seen in Figure 21, Francis and the current McAllister-Ianelli method produced similar results in more recent years, while the Dirichlet multinomial produced much greater biomass. None of the methods, including the current McAllister-Ianelli method, had better fits to the AI survey (Figure 22).

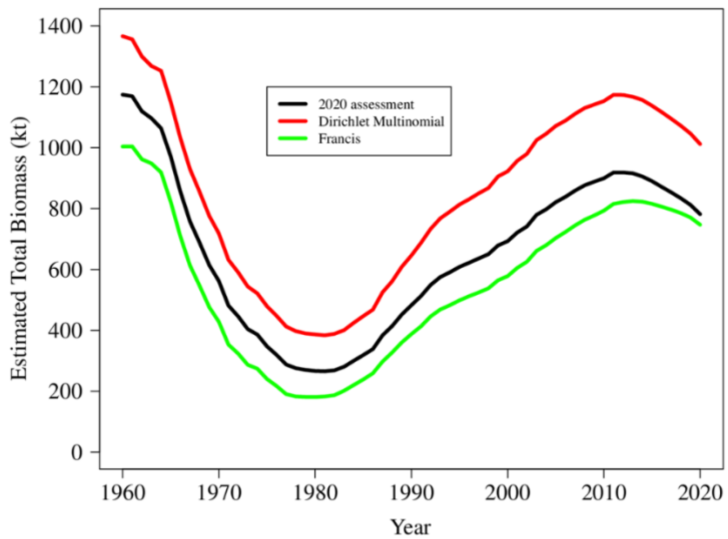


Figure 21. Biomass of POP by weighting method

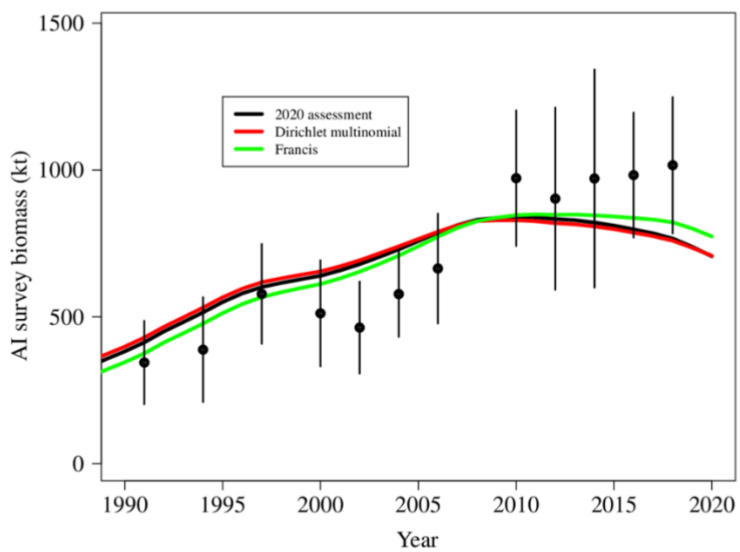


Figure 22. Fits to the AI survey using various data weightings.

Because the Francis weightings appeared to be the best of the sensitivities for fitting and didn't produce large differences in biomass from the base case, a further sensitivity was requested during the review to half the Francis weights. Biomass results and fits to the surveys can be seen in Figure 23 and Figure 24 respectively.

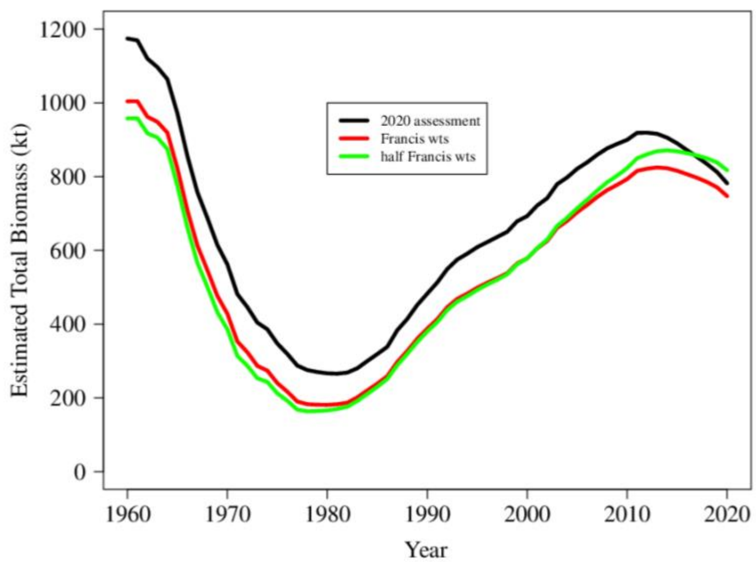


Figure 23. Biomass using different Francis weightings.

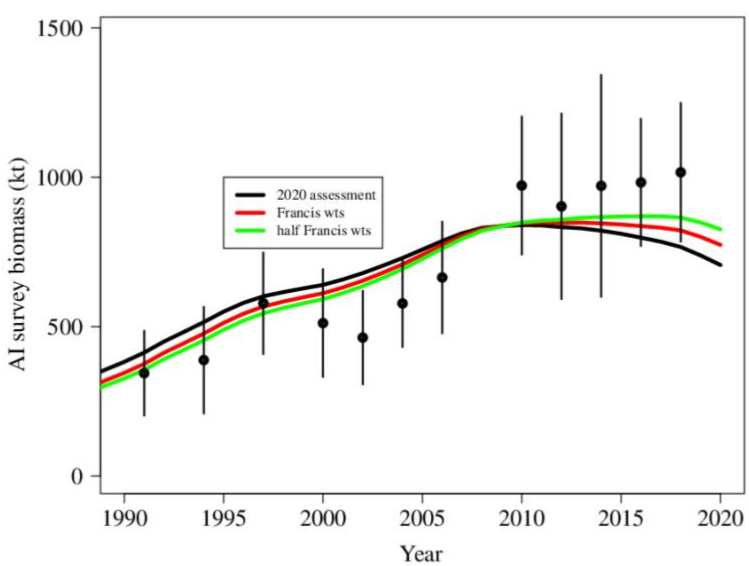


Figure 24. Fits to the AI survey using different Francis weightings.

This leads to the inevitable conclusion, that downweighting the age and length comps improve the fit to the surveys while increasing the estimated biomass. This was seen above where the complete removal of those data sources increased the fit to the surveys and reduced the retrospective, while also increasing the biomass.

There is no a priori reason to believe the surveys over the age and length comp data, at least not to the degree that is needed to resolve the fit to the survey in this model. In essence fitting, the survey requires the near-complete elimination of the age and length comp data to produce satisfactory fits. And it seems unlikely that those data sources don't contain important information on the population. Likewise, downweighting the surveys is also not an option, as those surveys not only show similar trends but show the population has increased simultaneously in multiple areas and depths over time; in keeping with industry observations.

Other sensitivities and requests during the review did not resolve this issue.

After extensive discussion, the model proposed for POP seems reasonable, though some further work (below) could help resolve this and other issues. Overall, not fitting the last few years of the survey is in some cases precautionary, with the results of the terminal year biomass likely above model estimates. However, this type of model behavior is concerning, particularly should the estimates of biomass/abundance from the survey data start to decline, or should there be more uncertainty as a result of the AI survey not being conducted in the future.

As such the current model represents the best available information on stock status. Overall, the stock appears to be in good condition, with F below the reference points while biomass is above its reference points. If anything, the terminal estimates of F and biomass are conservative, being mostly biased in a conservative direction.

Description of the individual reviewers' roles

The role of this reviewer for the POP assessment was to read all pertinent information provided, attend the review meeting, and provide constructive input into deliberations. After the meeting, the role is to provide an objective, independent, and impartial review of the methods used as well as provide justifications for statements made as well and a critique of the assessment and the process. Further, the role was also to provide suggestions on how to improve the assessment, including, but not limited, to research recommendations as well as model structural changes, if need be.

Terms of Reference for the Peer Review

Evaluation of the data used in the assessments, specifically trawl survey estimates of abundance, and recommendations for processing data before use as assessment inputs

This assessment has multiple sources of data informing the model. Age and length data seem appropriately processed as well as having a large number of samples in the recent time frame to inform model parameters.

The difficulty in age and length sampling is that there appears to be some bias in sample collection. Because the bulk of the samples from the fishery are taken when POP are the dominant catch in a haul, the incidental catch is not well represented. This can be fairly high in some years (30-40%). Moreover, many of these hauls are targeting pollock, which uses a different gear type (midwater trawl) rather than the standard bottom trawl. Additionally, because POP tends to school by size and age, you would expect that smaller schools might contain some of the larger/older individuals in the population. However, without good sampling on hauls where POP is dominant, this will remain an uncertainty. A recommendation is made further in this document to help address some of this uncertainty.

The surveys appear to be very informative on the stock's status. Both surveys show a very similar pattern in population growth over time, showing increases not only in biomass per tow but also the number of positive tows and an increase generally across all depths and strata. While catchability appears to be high (1.3) it is noted that POP tends to school and there is likely some vertical herding going on as with other *Sebastes sp.* Around the world. All this information tends to suggest a growing population of POP; a fact reiterated by the industry members that attended the review.

One of the uncertainties here is the AI survey. This survey has not been conducted since 2016 and will not be conducted in at least the near future. This is a difficult uncertainty for this approach as the AI survey is an important part of the data stream for this assessment.

Another uncertainty, and the subject of numerous questions during the review meeting, was the survey stratification of the surveys. The optimization appears to be based on 14 different species using an economic measure, price per lbs. This is an odd optimization given it's more of an economic analysis rather than biological. It seems unlikely that much work will be conducted on this, however, as it will affect not only POP but other stock assessments in the region.

A final issue relates to how the surveys were incorporated into the current modeling framework. While putting them in as separate surveys are ideal, the currency used seems inconsistent as the model uses the number of individuals while the survey is incorporated as biomass. As stated earlier in this report, the currency used in this type of modeling approach is numbers, rather than biomass. While this was examined during the review workshop to a limited extent (Figure 25) with little difference, a further examination is likely warranted given the hint of a difference in the most recent years, and a recommendation to that effect is made below.

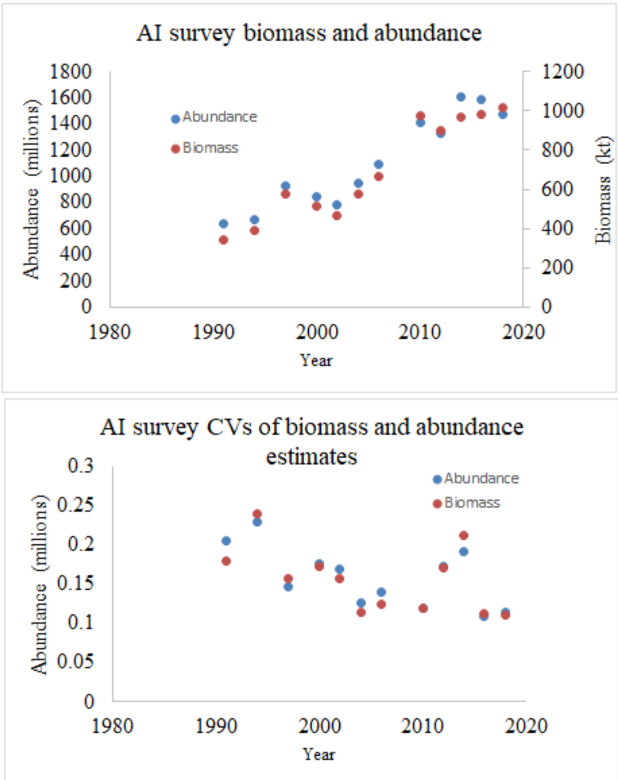
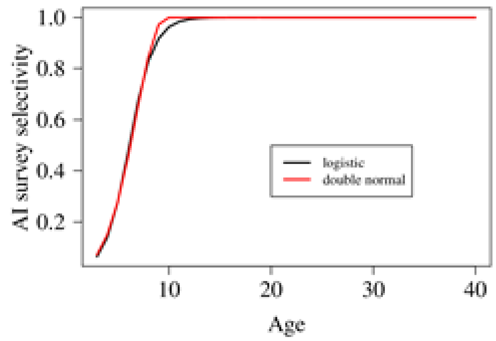


Figure 25. Abundance, biomass, and CV estimates from the AI trawl survey.

Evaluation of analytical methods used in assessments, particularly in regard to selectivity, modeling of natural mortality, and data weighting assumptions

As outlined in the report, there seems to be no difference between using a logistic survey selectivity and allowing for doming. Both produce a similar selectivity pattern (Figure 26). And as such the assessment use of a logistical selectivity seem appropriate.



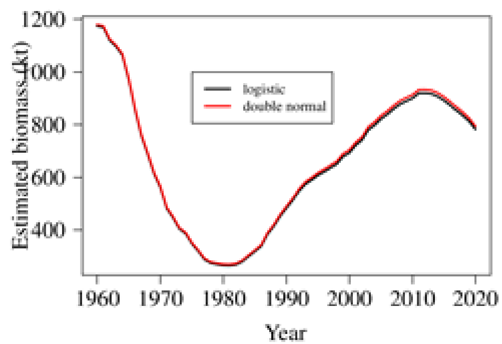


Figure 26. Survey selectivity with and without allowing doming.

Selectivity for the fishery was accomplished using the bicubic spline approach with some priors to mitigate doming on the downward slope at older ages. This appears appropriate as there doesn't seem to be any evidence that older fish are not exposed to the fishery, though some light doming is seen.

Interestingly there is an odd pattern to the selectivity (Figure 27). Overall, the most recent years show a saddle-like pattern, with peaks in selectivity at approximately age 15 and 35. This pattern appears to be increasing over time. While the exact cause of this is unknown, one possibility relates back to the sampling issues by the at-sea sampling mentioned above. If there are in effect two "fleets", one directed and one non-directed and using different gears, this might be the cause of this peculiar shape. A recommendation to examine fleet structure and the sampling of the non-targeted hauls is made below.

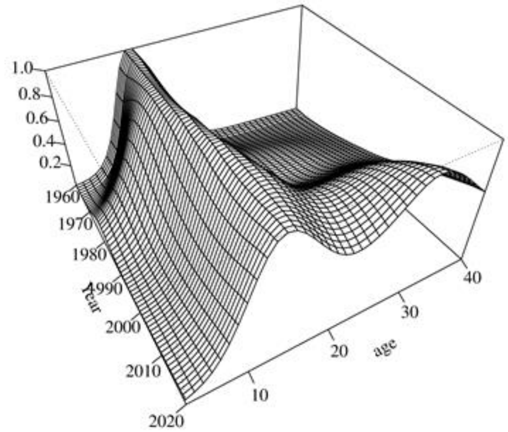


Figure 27. Selectivity by age and year for POP

Natural mortality seems well informed. The use of Then et al. (2015) in providing accessory evidence to back up the model estimates of M was well done. Overall, the model estimates of natural mortality seem in line with both the model's diagnostics, as well as published literature from other stocks. As mentioned earlier there appears to be no evidence that there should be an expected change in natural mortality across years, and as such a constant M is likely a more robust assumption. That said, the prior on the natural mortality estimate appears to not be of much use given where the model estimates the final M.

The review workshop spent a good amount of time discussing data weightings for this assessment. Given the choice of the base McAllister-Ianelli method, the Francis weighting, and the Dirichlet multinomial, the current McAllister-Ianelli method is acceptable. While the Francis weighting appears to give slightly better fits to the surveys, it didn't produce very different results in the time series of biomass, especially at the end of the time series. Additionally, even halving the effective sample size for the Francis weighting did not produce acceptable results. Given this, the problem of fits to the survey, and resulting retrospective patterns are unlikely to be resolved by the weighting method. The only resolution at this time is to remove the age and length comps from the model entirely, an unacceptable result given the lack of information left to resolve year-class strength.

Evaluation of the ability of the stock assessment model for BSAI Pacific Ocean perch to provide parameter estimates to assess the current status of the stock

The current model uses a statistical catch at age approach for modeling the population. Modeled parameters include fishing mortality, Recruitment deviations, R_0, survey catchability natural mortality, maturity, and both fishery and survey selectivities. All of these seemed well resolved in the model even without their priors, in some cases.

With one exception, the model appears to be performing well. That exception, the lack of fit to the surveys in the recent part of the time series, is the cause for some consternation as it leads to a retrospective bias that tends to underestimate the SSB and overestimate fishing mortality relative to the terminal year.

This is a vexing problem, and not one easily solved. Changes in weights and removing certain data sources indicate it is a result of tension between age and length information and the surveys themselves. While not ideal, the issue is in the precautionary direction, suggesting that model estimates of F and SSB are likely to be revised downwards and upwards respectively in subsequent years. This coupled with the survey information itself indicates that the stock is in good shape, with increases in biomass across spatial strata, surveys, and depths. The increase in positive hauls on the survey is also telling, as it indicates not only greater abundance but an increase in the areas where POP are found.

That said the current model formulation should represent the best available science. POP does not appear overfished, and overfishing is not occurring. For reasons mentioned above, the assessment is likely underestimating biomass, suggesting that the stock is likely higher than reported. However, there is a lot of important issues the model is not addressing well, such as the lack of fit to the latter part of the survey time period, as well as others. The current placement of POP in Tier 3 for formulating ABC and OY is therefore appropriate.

Evaluation of the strengths and weaknesses in the stock assessment model for BSAI Pacific Ocean perch.

The modeling approach and implementation used for POP has a number of strengths. There appears to be good accounting of aging, including the use of an aging error matrix. Further, the model seems to be performing rather well, with the likelihood profiles showing good contrast and a lack of a flat surface across many parameters. This indicates that these parameters are rather informative to the model.

Other strengths include the estimation of maturation, which is a nice addition. The use of the bicubic spline approach to modeling selectivity is also welcome, as it allows for some flexibility in formulating fishery selectivity. In particular, the use of selectivity by year is an additional strength, even if the resulting selectivity curves are odd.

There are a number of weaknesses in the model as currently implemented, however. Chiefly the lack of fit to the survey data in the latter portion of the timeframe suggests that there is a data source that is in tension with the other sources of data. Further examination reveals that the age and length comps are conflicting with the survey estimations of biomass. This results in a retrospective pattern that overestimates F and underestimates SSB relative to the terminal year. This happens as each successive survey draws the model's estimate of biomass up in previous years.

Multiple sensitivities were attempted to help resolve the issue. While changing the data weights improved the fits marginally, the complete resolution required the total removal of the age and length comps. Given that this would lead to difficulties in resolving year-class strength, this is not a good solution. After examination, the issue remains unresolved, though some suggestions on future avenues to explore are made below.

Another weakness is the bias in sampling associated with the at-sea observer coverage. As mentioned earlier, most of the samples come from hauls where POP was the dominant species. However, a growing portion of the catch is coming from hauls in which POP was not the dominant catch. Further, some of the hauls in which POP was not the dominant catch are with another gear type/fleet, the midwater trawl fishery for pollock. Some exploration of different fleets as well as an increase in sampling might be useful in future assessments, as outlined below.

That said the model as formulated is appropriate for the determination of stock status and in the formulation of management advice. As indicated above the model's bias tends to be in the more precautionary direction and is reinforced by the surveys which indicated an even higher population size than the model estimates.

Recommendations for improvements to the assessment models

While the current model is acceptable for use as a management tool, there are a number of recommendations that can be made to help improve the assessment and the model for the next review.

One of the most important issues that come up in the review process was the lack of sampling on hauls that did not have POP as one of the dominant species. This indicates that there may be a bias in the current sampling regime for POP, resulting in non-targeted hauls being undersampled. Given that POP likely school by size/age, this might be an important issue to resolve. While there is currently no reason to assume POP caught on targeted hauls have different sizes/ages than those on non-targeted hauls, this potential difference should be studied to at least rule it out, if nothing less. **As such it is recommended that a study be conducted with samples of POP for age and length by onboard observers when POP is not one of the dominant species.**

A related recommendation revolves around the fleet structure used in the model. Should the study mentioned above find some differences in size/age between hauls where POP was dominant and those where it was not, breaking out a "directed" vs a "non-directed" fleet may have some advantages in model diagnostics, allowing for perhaps a better fit to the survey biomass. Further, many of the non-dominate hauls were the result of pollock fishing activity. As the pollock fishery generally uses midwater, rather than bottom trawl gear, this re-examination of fleet structure might be important. A related, but interesting, side note is that the fishery selectivity curve for POP is indicating a saddle-shaped curve. Such a curve could be the result of two separate fleet selectivities, each with its own selectivity peak, being combined. This, however, is speculation and there could be other, more biological/ecological reasons for this saddle-shaped curve in selectivity. **That said, it is recommended that fleet structure be reexamined prior to the next peer review for POP.**

An additional recommendation revolves around the use of the plus group (40+) in the model. While it is agreed that 40 is useful as it is near L_{inf}, and has achieved maximum maturity, this results in nearly 60-years classes in the plus group. While the assessment team indicated that such sensitivities had been performed in the past with both this stock of POP and the Gulf of Alaska, it may be worth redoing those sensitivities in the near future. While this will likely greatly increase the uncertainty in parameter estimates, it is important as it might help to rectify some of the lack of fit to the surveys. **As such it is recommended that a full sensitivity analysis be performed on the choice of a 40+ group.** Such sensitivities should examine the later (50 -60+) grouping.

The allowance of maturity to be a parameter in the modeling was a novel approach, one that is not often seen in this type of model. While well done, it is only informed by two studies that actually measure maturity. **Given that they are not very far apart and the fact that the studies have not been recently redone, it is recommended that a maturity study be conducted in the near future to help anchor the data in the model.**

Another recommendation is again aimed at the surveys and their incorporation into the model. During the review, it was highlighted that the surveys were incorporated into the model using biomass. One request was to examine the surveys in terms of abundance, rather than weight or biomass. **While the results were not very different in the overall trend it is recommended that incorporating the surveys be done in abundance rather than as biomass, at least as a sensitivity and that this analysis is fully explored prior to the next review.** While this is unlikely to completely resolve the issue of contrasting data sources, it may reduce the overall lack of fit, and thereby the retrospective pattern as well.

Similar to the above recommendation, this final recommendation deals with the surveys. It was clear during the review that the AI survey has not been conducted since 2016 and there are no plans to resume the survey in the near future. While this survey shows a similar trend to the EBS, the loss of this survey's signal, as well as its length and age data, will likely make the results of this assessment even less certain. **Therefore, a recommendation to continue the AI survey is suggested.** As this recommendation was suggested during the review workshop to be unlikely given budget and logistics, there is another recommendation, therefore, to make. As suggested in the review meeting by another reviewer, a fishery-dependent CPUE could be useful in the assessment of this stock. While such a CPUE index might suffer from hyper stability given the schooling nature of the species, it is worth exploring in the near term, particularly as data from the AI ages. **As such an additional recommendation is made to explore a standardized fishery-independent CPUE index prior to further review.**

Overall conclusions

The assessment team did a wonderful job with this assessment. The model chosen seemed to be a good fit for the data in hand, the sensitivities were rigorous and well explained, and the document presented had few issues. Unlike other assessments, this assessment had a completed document for the reviewers to work from.

One issue brought up in the review was the age of the assessment. This review was conducted on the assessment as completed in 2020, with most data through 2019. While this is understandable given the process, it should be noted a bit better in the initial presentation to basically give the process a bit more emphasis.

Additionally, one over-arching theme during this assessment was the use of priors to constrain important parameters from deviating too far. While this is acceptable, philosophically it is best if priors can be removed where possible to allow parameters to be freely estimated. Though this might be more of a personal preference, freely estimated and unconstrained parameters that arrive at reasonable estimations bolster the argument about model robustness. In short, it's usually wise to reinvestigate the need for priors and penalty terms for each assessment, rather than constraining estimation because it was done in previous assessments.

Based on the materials provided during this review, the assessment seemed particularly well done. The assessment information including presentations as well as recorded presentations were made available beforehand, though there were a few technical glitches with sharing information. The set-up for POP was one of the best examples of providing just the right amount of information far enough in advance, to really help the process along.

This was an enjoyable review, despite there being less time to meet (three days) when compared to other recent assessment assessments. This was in part due to the organization of the materials as well as the availability of the documentation. However, this review as well as others is best done in person rather than remotely. While the other reviewers on this assessment may disagree, in-person meetings for reviews are vital. Much is lost when there are no side-bar or dinner conversations about the assessment. While it is unlikely that there would have been a different outcome for this assessment had it been in person, in-person meetings allow for a more collaborative experience and a better review product.

It was good to work with the reviewers as well as the assessment team for POP again. All involved did a great job in stimulating discussion and in formulating thought-provoking questions. Overall, a very fun review.

References

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68:1124-1138.

TenBrink, T,T, and P.D. Spencer. 2013. Reproductive biology of Pacific Ocean perch and northern rockfish in the Aleutian Islands. N. Am. J. Fish. Man. 33:373-383.

Then, A.Y., J.M. Hoenig, N.G. Hall, and D.A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science, 72(1): 82–92. Doi:10.1093/icesjms/fsu136.

Appendix 1: Bibliography of materials provided for review

A full list of the reports and files can be found at https://apps-afsc.fisheries.noaa.gov/Plan_Team/2022_pop_cie/

Additional literature included:

Jones, D., Rooper, C., Wilson, C., Spencer, P., Hanselman, D., Wilburn, R. 2021. Estimates of availability and catchability for select rockfish species based on acoustic-optic surveys in the Gulf of Alaska. Fisheries Research Volume 236, https://doi.org/10.1016/j.fishres.2020.105848

Legault, C, and Palmer, M. 2015. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences. 73(3): 349-357. https://doi.org/10.1139/cjfas-2015-0232

Then, A.Y., J.M. Hoenig, N.G. Hall, and D.A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science, 72(1): 82–92. Doi:10.1093/icesjms/fsu136.

Thorson, J. T., K. F. Johnson, R. D. Methot, and I. G. Taylor. 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research 192: 84-93.

Appendix 2: A copy of this Performance Work Statement

Performance Work Statement (PWS)

National Oceanic and Atmospheric

Administration (NOAA) National Marine Fisheries

Service (NMFS)

Center for Independent Experts (CIE)

Program External Independent Peer

Review

May 9-13, 2022

BSAI 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 Bering Sea/Aleutian Islands Pacific ocean perch provide the scientific basis for the management advice considered and implemented by the North Pacific Fisheries

Management Council. An independent review of these integrated stock assessments is requested by the Alaska Fisheries Science Center's (AFSC) Resource Ecology and Fisheries Management (REFM) Division. The goal of this review will be to ensure that the stock assessments represent the best available science to date and that any deficiencies are identified and addressed.

The goal of this review will be to ensure that the stock assessments represent 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 age-structured stock assessment methods in general and, in particular, application of ADMB in stock assessment.

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 trawl survey design and estimation of stock biomass
- 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 send by electronic mail or make available at an FTP site to the CIE reviewer all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review.

2) Attend and participate in the panel review meeting. The meeting will consist of presentations by NOAA scientists, including the stock assessment authors and survey team members to facilitate the review, provide any additional information and answer questions from the reviewers.

3) 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 according to the specified milestones dates.

Foreign National Security Clearance

When 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 reviewers who are non-US citizens. For this reason, the 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/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national- registration- system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be at the contractor's facilities, and in Seattle, WA.

Period of Performance

The period of performance shall be from the time of award through July 2022. The CIE reviewers' duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables

The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	**Milestones and Deliverables**
Within two weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
May 9-13, 2022	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
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 $8,000.

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

Project Contact(s):
Paul Spencer
NMFS, Alaska Fisheries Science
Center 7600 Sand Point Way NE,
Seattle WA 98115
Phone: 206-526-4248
paul.spencer@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. Evaluation of the data used in the assessments, specifically trawl survey estimates of abundance, and recommendations for processing data before use as assessment inputs
2. Evaluation of analytical methods used in assessments, particularly in regard to selectivity, modeling of natural mortality, and data weighting assumptions
3. Evaluation of the ability of the stock assessment model for BSAI Pacific ocean perch to provide parameter estimates to assess the current status of the stock
4. Evaluation of the strengths and weaknesses in the stock assessment model for BSAI Pacific ocean perch
5. Recommendations for improvements to the assessment models

Annex 3: Tentative Agenda

Virtual Panel Review

BSAI rockfish – Pacific Ocean perch

May 11-13, 2022

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

As provided in advance of the meeting via the Weblink site.

All times below are Pacific Daylight Time. Daily breaks at 11:30am and 3:45pm, Lunch 1pm-2pm

Wednesday, May 11		
10:00 am – 11:30 am: Introduction/Background		
	Introductions and agenda	Pete Hulson
	Overview of rockfish biology, fishery, and history of assessment	Pete Hulson
	Current management of Alaska rockfish	Paul Spencer
11:45 am - 1:00 pm: Discussions		
2:00 pm – 3:45 pm: Input data		
	Survey data	
	Abundance, distribution, and age composition	Ned Laman, Paul Spencer
	Update on model-based abundance	James Thorson
	Fishery data – Catch, observer program, ages, lengths	Raul Rameriz, Paul Spencer
	Age determination, lengths, maturity, and growth	Delsa Anderl, Paul Spencer
4:00 pm - 5:00 pm: Discussions		
Thursday, May 12		
	Pre-recorded presentations to review: Field-based catchability	
10:00 am – 11:30 am: Assessment model		Paul Spencer
	Model structure, likelihoods, data weighting, parameter estimates, data fit, diagnostics	
11:45 am - 1:00 pm: Discussions		
2:00 pm – 3:45 pm: Assessment model		Paul Spencer
	Catchabilities, selectivities, model fits, diagnostics	
4:00 pm - 5:00: Discussion		
Friday, May 13		
10:00 am – 11:30 pm: Model Developments		Paul Spencer

	P	Alternative data weighting
		Alternative specification for natural mortality
		Other miscellaneous model developments
11:45 am - 1:00 pm: Discussion		
2:00 pm - 3:45 pm: Requested topics/model runs		
4:00 pm - 5:00 pm: Summarize, revisit Terms of Reference		

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

Chair	Pete Hulson	AFSC, Seattle
Members	Noel Cadigan	Memorial University, St. Johns, Newfoundland, Canada
	Geoff Tingley	Gingerfish Ltd, Wellington, New Zealand
	Matthew Cieri	Maine Department of Marine Resources, US