***Center for Independent Experts (CIE) Independent Peer Review Report on the Gulf of Alaska Demersal Shelf Rockfish Assessment***

Kotaro Ono[[1]](#footnote-1)

****

Prepared for:

The Center for Independent Experts (CIE)

**Executive summary**

*The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.*

The Demersal Shelf Rockfish (DSR) review workshop (RW) took place on Sept 12-13, 2023 (8-12pm Norway time). The RW was conducted in an open, friendly, yet professional manner. Three CIE reviewers were present during the meeting, Kevin Stokes, Robin Cook, and Kotaro Ono, in addition to the lead stock assessor (Philip Joy), the NOAA project contact (Chris Lunsford), and several other observers from Alaska Department of Fish and Game (ADF&G).

Overall, I believe the development of the Bayesian state-space surplus production model (SS-SPM) is warranted. The proposed SS-SPM takes into account all sources of available information and the associated uncertainties. The SS-SPM goes one step beyond the currently-adopted standard i.e. the REMA model, as it directly models the population dynamics, while taking into account the catch history, and estimates reference point that can readily be used for management (e.g. FMSY, BMSY) without the need to use a proxy for FMSY (i.e. M) as it is currently the standard for Tier 5 stocks. The SS-SPM thus offers an approach that is in line with the National Standard 2: using the best scientific information available to device management measure as we bypass unnecessary subjectivity (i.e. assuming FMSY = M). Overall, I believe the SS-SPM offers a rational framework for devising quota advice for a stock where information on catch and abundance indices exist but with insufficient information to build an age-structure assessment model.

That being stated, before the present SS-SPM can be recommended to be used for management, it requires further development and changes which are summarized below:

* The International Pacific Halibut Commission (IPHC) survey CPUE is a key index used both in the REMA model and SS-SPM. However, the current approach for deriving the CPUE index is not adequate: raw CPUE data need to be properly standardized using an approach that can handle zero-inflation and heavy tail (e.g. model-based approach using Tweedie distribution or a delta model). Raw CPUE data should also be based on biomass (weight) instead of number of fish to be consistent with the unit in SS-SPM.
* Phases 1-3 of the SS-SPM need to be merged into a single joint framework to avoid “double-dipping” the data.
* A strategic construction of SS-SPM is recommended, starting with a simple model that is most similar to the currently adopted REMA model. Then one can add slowly some complexity to the model (e.g. adding more sources of uncertainty, going further back in time). In that way, one can test the impact of each step on model estimates and performance. In a similar vein, I suggest to re-parametrize the SS-SPM to reduce parameter correlation and increase computation speed and/or re-write the model in Stan.
* Last but not the least, a simulation framework (or MSE) needs to be developed to test the performance of the proposed SS-SPM for management (i.e. answering questions such as: Are SS-SPM-derived FMSY, OFL estimates unbiased? What are the risks to the stock (to fall below MSST) if using the estimated FMSY to set the TAC?) and compare its performance to the current status-quo approach (the REMA model).

To summarize, I encourage the authors to pursue with the initiated endeavor and further develop the SS-SPM modelling (and simulation testing) framework to potentially use it for management of “data-limited” stocks that have (at least) a catch time-series and an abundance index but not enough information to build an age-structure assessment model.

**Background**

*The main body of the reviewer report shall consist of a* ***Background****, Description of the Individual Reviewer’s Role in the Review Activities, Summary of Findings for each TOR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the TORs*

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 the USA’s marine living resources based upon the best scientific information available (BSIA). NMFS science products require scientific peer reviews that are strictly independent of all outside influences. Thus, these external reviews are essential to strengthening scientific quality assurance for fishery conservation and management actions.

The GOA DSR assessment has evolved much over time: it started with simple ROV survey-based biomass estimates, then transitioned to a full age-structured stock assessment model, before coming back again to a simpler state-space random effect model (REMA). Concomitantly to the 2022 assessment using REMA, the assessment authors developed a new Bayesian state-space surplus production model to be potentially used for yelloweye rockfish assessment. The work was presented to the North Pacific Fishery Management Council (NPFM) plan team in 2022 which, together with the Scientific and Statistical Committee (SSC), requested an official Center for Independent Experts (CIE) review of the model. The present RW hence focuses on evaluating the appropriateness of the developed approach for assessing and managing the DSR (namely the yelloweye rockfish) stock in the Southeast GOA following the guidelines and Terms of Reference (ToR) set in the statement of work (see Appendix 2 below).

**Description of the individual reviewer’s role**

*The main body of the reviewer report shall consist of a* ***Background****, Description of the Individual Reviewer’s Role in the Review Activities, Summary of Findings for each TOR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the TORs*

The role of the reviewer is set out in the CIE Statement of Work, Attachment A, attached here in Appendix 2 and reads as follows:

1. Review the following background materials and reports prior to the review meeting.
2. Attend virtually and participate in the panel review meeting.
3. After the review meeting, conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and TOR, in adherence with the required formatting and content guidelines.
4. Deliver their reports to the Government by the specified deadline.

**Summary of findings for each TOR**

*The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer’s Role in the Review Activities, Summary of Findings for each TOR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the TORs*

# TOR1. Evaluate the use of the proposed Bayesian state-space surplus production models, and model methods for use in management of the yelloweye rockfish portion of the GOA Demersal shelf rockfish complex. Specifically, determine if results suggest that the NPFMC Tier 5 approach is appropriate. Should biomass estimates from such a model be used in place of the survey average method currently used for most Tier 5 stocks managed by the NPFMC?

## Is the model worth further development for use in managing yelloweye rockfish in the SEO?

The short answer is yes.

Generally speaking, both SS-SPM and REMA are state-space models which take the indices of abundance as inputs and estimate the underlying biomass of the stock. However, SS-SPM goes one step beyond the REMA model: in a way, REMA is a temporal smoother that links the observation and underlying biomass with the user-specified constraints (i.e. the error structure for both the process and observation errors). On the other hand, the SS-SPM tries to link the abundance indices to the underlying biomass by following a user-specified population dynamics model (here a Pella-Tomlinson model) and uses the catch data to adjust the scale of biomass, while including different sources of uncertainties as in REMA (and more). These additional steps in SS-SPM also allow estimating management-relevant reference points (e.g. FMSY, MSY, BMSY, which the REMA cannot) that can potentially be used directly to set the OFL and ABC for the stock.

Following the current status-quo (Tier 5 harvest control rule in the NPFMC), when using REMA to set the quota for yelloweye rockfish, one has to use the best “guestimate”of M to calculate the OFL and ABC. But M is notoriously difficult to estimate (and cannot be estimated within REMA), and this creates a disconnect between the biomass estimated from REMA and the TAC. Indeed, even if the estimate of current biomass from REMA is accurate, if one gets the value of M wrong, then the recommended TAC would be completely off and nothing can be done in REMA to prevent that.

On the other hand, SS-SPM estimates reference points (e.g. FMSY) whose accuracy and precisions are directly dependent on the information contained in the input data and parameters. As a result, improving the information content in the inputs (e.g. constructing a better indices of abundance, getting a more accurate and precise prior) can improve the estimates of these reference points and, subsequently, lead to a more rational quota allocation. In other words, the “correctness” of management recommendation using SS-SPM would be directly related to the performance of the assessment model (in estimating management relevant quantities) and not to a somewhat subjective choice of M value. In that sense, the SS-SPM offers an approach that is more in line with the National Standard 2 – using the best scientific information available to device management measure – as we would no longer need to use the subjective assumption that M = FMSY. Overall, I believe the SS-SPM offers a rational framework for devising quota advice for a stock where information on catch and abundance indices exist but with insufficient information to build an age-structure assessment model.

One thing needs to be avoided for sure: one should not use SS-SPM to estimate the stock biomass and use an estimate of M to set the TAC following the current Tier 5 Harvest Control Rule (HCR) system. This greatly reduces the benefit of SS-SPM and increases unnecessarily the subjectivity/complexity of the TAC allocation process which goes against the National Standard 2 guidelines. Past studies have used the value of M as an approximation to FMSY (Gulland 1970) but FMSY is readily estimated within SS-SPM. Moreover, recent studies suggested that M is likely different than FMSY for yelloweye rockfish. Zhou et al. 2012 has found that FMSY = 0. 694M was the best relationship for Scorpaenidae – which yelloweye rockfish is part of. Horbowy and Hommik (2022) conducted a simulation approach (using an age-structure model) and suggested that the actual relationship between FMSY and M is dependent on the steepness of the stock recruitment relationship.

## Suggestions for model development

While SS-SPM sets an attractive framework to perform a “data-poor” stock assessment and derive estimates of *FOFL* and *maxFABC* needed in setting the TAC, in practice, it is not that simple. The SS-SPM presented in the review was nice but the model was complex and included a lot of uncertainties. This makes it hard to identify sources of problems (if/when it happens). I therefore suggest the following:

1. The author should start with a much simpler version of SS-SPM. One that is “most similar” in terms of structure to the REMA model then build up on the complexity. That way, one can evaluate the effects of the different components of the SS-SPM and determine what creates differences compared to REMA model outputs (e.g. biomass trajectory). This also means focusing on the period after 1980 to begin with. Additionally, I suggest that the authors use informative priors (when they can) based on other studies (e.g. Thorson et al. 2012, Winker et al. 2020) to help model convergence and combine it with sensitivity tests to examine the impact of these informative priors. Furthermore, one can start with an SS-SPM without catch (just with process and observation error). Such model would enable examining the effect of including density dependence in the process dynamics instead of just using a random walk structure as in REMA. In this process, the authors need to make sure about the similarity to the REMA i.e. If the final REMA has a common process error variance among the 4 management units, then it makes sense to use the same “r” across the 4 areas in the SS-SPM as well (similar thinking needs to be done even if the model is re-parametrized). Then one can add the catch data, excluding any estimates of uncertainties around it to begin with. In the final step, one can add different uncertainty components one by one as included in the current SS-SPM. Such step-by-step analysis enables detecting which component(s) of the SS-SPM were responsible for creating differences (in biomass trajectory for example) compared to the currently adopted REMA. However, this will not tell anything about the reliability of the derived FMSY, BMSY, or final year biomass estimate. The later can only be tested using a simulation framework which leads to point b).
2. In order to test the performance of the SS-SPM, one should create a simulation framework. This means extending the simulation framework that the authors have already created to enable evaluating many more scenarios (i.e. testing the inclusion of different uncertainty component, at different amounts, and examine the impact to the management relevant parameters e.g. FMSY, OFL). Moreover, performance testing should not simply focus on parameter bias but should also evaluate risks e.g. what is the proportion of the simulations where the estimated FMSY, OFL were above 20% of the true value? Furthermore, the simulation framework can be extended to a management strategy evaluation (MSE) to enable answering questions such as: what are the risks to the stock falling below MSST if using the SS-SPM estimated FMSY to set the quota? The simulation framework could also be used to ask what-if questions e.g. what are the consequences on reference points estimates when using informative prior but the “wrong one”?
3. While the development of the SS-SPM would benefit from points a) and b), there is another issue that is relevant to both REMA and SS-SPM. And this relates to the IPHC survey CPUE standardization. The current approach is using a design-based estimate by management strata after filtering some data based on data quality criteria (from IPHC) or the amount of observed yelloweye rockfish CPUE being above 0. However, filtering out zeros is dangerous as it risks removing true zeros. Instead, the approach should focus on defining a habitable area for yelloweye based on some expert-judgement (e.g. depth limits) and use all the IPHC survey data within the defined domain for the CPUE analysis (while still excluding bad survey station data). Such data selection process should be reasonable as the IPHC survey stations selection process is random (not biased at least – needs verification though). The derived data will probably be zero-inflated but there are many approaches that can deal with such issues. Additionally, SS-SPM is based on weight, therefore, the CPUE index should also be based on weight and not on the number of individuals. In such a process, one should test different candidate models with different covariates combination and distributional assumptions. Below, I illustrate two CPUE standardization examples based on the data and code made available by the lead author (<https://github.com/commfish/seak_seo_dsr>) and using the CPUE in terms of weight. Also, it should be noted that geostatistical models might be hard to fit due to the limited data. A quick exploration using the R package *sdmTMB* indicated that the model had difficulty in estimating the spatial random field.

### Test 1: using a tweedie distribution (R code attached)

Survey$areayear = apply(cbind(Survey$Year, Survey$SEdist), 1, function(x) paste0(x, collapse ="\_"))

Survey$areayear = as.factor(Survey$areayear)

Survey$WCPUE = Survey$YE.obs \* Survey$mean.YE.kg / Survey$HooksObserved

Survey$WCPUE\_ratio = Survey$WCPUE / Survey$O32.Pacific.halibut.weight

gam1 <- gam(WCPUE ~ as.factor(SEdist) + as.factor(Year) + s(areayear, bs="re"), data = Survey, family = tw)

res1 <- simulateResiduals(gam1, n = 1000)

plot(res1, quantreg = TRUE)

A graph of a graph and a graph of a graph

Description automatically generated with medium confidence

predict\_data <- expand.grid(Year = 1998:2022, SEdist = c("EYKT", "CSEO", "SSEO", "NSEO"), AvgDepth.fm = 100)

predict\_data$areayear = apply(data.frame(predict\_data$Year, predict\_data$SEdist), 1, function(x) paste0(x, collapse ="\_"))

predict\_data$areayear = as.factor(predict\_data$areayear)

prediction = predict(gam1, predict\_data, type="response", se.fit =TRUE)

predict\_data$pred = prediction$fit

predict\_data$pred.se = prediction$se.fit

ggplot(predict\_data, aes(x=Year, y = pred)) + geom\_line() + facet\_wrap(.~ SEdist) + theme\_bw() +

geom\_errorbar(aes(ymin = pred - 1.96\*pred.se, ymax = pred + 1.96\*pred.se ))

### Test 2: Now using a delta model (binomial + gamma)

**# binomial**

gam1.pres <- gam(I(WCPUE>0) ~ as.factor(SEdist) + as.factor(Year) + s(areayear, bs="re"), data = Survey, family = binomial(link = "logit"))

bs="re"), data = Survey[which(Survey$WCPUE >0),], family = scat(link = "log"))

res1.pres <- simulateResiduals(gam1.pres, n = 1000)

plot(res1.pres, quantreg = TRUE)

A black line on a white background

Description automatically generated

**## Now the positive model using the gamma distribution with log link**

gam1.pos <- gam(WCPUE ~ as.factor(SEdist) + as.factor(Year) + s(areayear, bs="re"), data = Survey[which(Survey$WCPUE >0),], family = Gamma(link = "log"))

res1.pos <- simulateResiduals(gam1.pos, n = 1000)

plot(res1.pos, quantreg = TRUE)

A graph of a graph of a number of numbers

Description automatically generated with medium confidence

prediction.pres = predict(gam1.pres, predict\_data, type="response", se.fit =TRUE)

predict\_data$pred.pres = prediction.pres$fit

predict\_data$pred.pres.se = prediction.pres$se.fit

prediction.pos = predict(gam1.pos, predict\_data, type="response", se.fit =TRUE)

predict\_data$pred.pos = prediction.pos$fit

predict\_data$pred.pos.se = prediction.pos$se.fit

**# Now combining the two parts (for both the mean and variance)**

predict\_data$pred\_all = predict\_data$pred.pres\*predict\_data$pred.pos

predict\_data$pred\_all.se = sqrt((predict\_data$pred.pos)^2\*(predict\_data$pred.pres.se)^2 + (predict\_data$pred.pres)^2\*(predict\_data$pred.pos.se)^2 )

**# now let's compare the derived indices between the above two methods and the one used currently**

iphc\_index <- read.csv(paste0("Data\_processing/Data/IPHC.cpue.SEO\_non0\_2023.csv"))

ggplot(predict\_data, aes(x=Year, y = pred\_all)) + geom\_line(col="purple", cex=1.5) + facet\_wrap(.~ SEdist) + theme\_bw() +

geom\_errorbar(aes(ymin = pred\_all - 1.96\*pred\_all.se, ymax = pred\_all + 1.96\*pred\_all.se ), col="purple") +

geom\_line(aes(x = Year, y = pred\_tweedie), col="red", cex=1.5) +

geom\_line(data = iphc\_index %>% mutate(SEdist = mngmt.area), aes(x = Year, y = CPUE.bootmean\*2.5), col="black")

A screenshot of a graph

Description automatically generated

A comparison of the derived indices of abundance by SEO regions between the two CPUE standardization models (one using tweedie (purple) and the other one using a delta model approach (red)) indicated similar trend but were both smoother than the status-quo indices (black). Only the confidence interval (1.96\*SE) for the tweedie model was plotted in the above figure.

1. On a side note, the author faced a long model run-time. The author can try reparametrizing the model in terms of B/K as in Meyer and Millar (1999) or in terms of MSY and gamma as in Fletcher’s (1978). Additionally, the model can also be re-written in Stan (<https://github.com/stan-dev/rstan>) to possibly reduce run-time by taking advantage of the Hamiltonian Monte Carlo algorithm (Monnahan et al, 2017) after re-parametrization.
2. Finally, the author should avoid creating 3 phases of estimation (problem of double dipping the data) but should focus on merging into a single framework. This should be easier when building a SS-SPM focusing on period post 1980. For integrating data pre-1980, some tips might be required to link the regional model with the sub-regional model and the author can contact me if interested in exploring this topic more.

# TOR 2 - Evaluate the management of the GOA demersal shelf rockfish complex as a whole including examination of the use of available data for species other than yelloweye rockfish and determination of Tier level designations.

## Is this management approach (mixing tier levels into an overall complex) acceptable based on utilizing the available data by species?

I do not see any reason why the ABC from different tiers cannot be combined. That being said, if one is thinking about buffer for the complex-wise ABC, then a more thorough test needs to be developed to evaluate the best way to do so e.g. through the use of multispecies (maybe ecosystem) and data poor simulation model.

## Is SS-SPM an acceptable model for management purposes?

As stated in my answer for ToR 1, I believe that SS-SPM is an acceptable model for management purpose. Such model is already implemented elsewhere (e.g. in ICES for managing some data poor stocks (ICES 2018))*.* That being said, as stated in my answer in ToR 1, I also recommend some thorough simulation testing prior to its actual application.

## Appropriateness of the Tier level designation

# For the yelloweye rockfish

The currently adopted REMA model is used based on tier 5 HCR i.e. catch limits are set based on an estimate of M (that is not calculated internally within REMA) and the estimate of biomass from REMA. Thus, the performance of this HCR is not only affected by the performance of the REMA model. This consequently adds some subjectivity/complexity in determining the catch limits.

Based on the current Fishery Management Plan (FMP), tier 4 stocks require the development of an age-structure model at one point. The development of an age-structure assessment model could potentially be a useful contribution for yelloweye rockfish assessment. However, these types of models are more complex and have consequently many more parameters to estimate (than REMA or a surplus production model such as SS-SPM) which are seldom estimable with the available data (e.g. M, steepness of the stock-recruitment relationship, recruitment variability, fishery/survey selectivity, changes in growth, maturity). Moreover, prior development of an age-structure model for yelloweye rockfish has encountered many difficulties concerning model fit, stability, and uncertainty, which will likely persist and would require lots of investigation and man-hour.

The proposed SS-SPM does not fall in either of the two tier systems (see also my argumentation for ToR 1). However, it provides a good compromise between the REMA model currently in place and a potential age-structure assessment model based on data availability, time requirement (manhour required to develop a fully functional age-structured assessment model), and usability for directly setting catch limits. Therefore the SS-SPM goes directly in line with the National Standard 2: using the best available scientific information to set management measure, and with the National Standard 7: Costs and benefits.

1. For the other species in the DSR complex

All other species in the DSR complex are extremely data poor. I therefore think that the current tier system designation (Tier 6) is appropriate for these stocks.

# TOR 3 - Review the methods used for estimating yelloweye bycatch in the directed Pacific halibut fishery and for estimating total yelloweye catch removals.

## Review the method for estimating yelloweye rockfish bycatch in the directed Pacific halibut fishery

The yelloweye rockfish bycatch was estimated using a ratio estimator based on the International Pacific Halibut Commission (IPHC) survey data. The ratio was based on the relative biomass of yelloweye caught per unit of legal (or total) halibut biomass and the ratio was then raised based on the total landing per management area. The major assumption being that the catch pattern in the survey is similar/representative of the halibut fishery i.e. that the fishery does not happen in very different grounds/contexts than the survey. A posteriori, this assumption seems plausible considering that the estimated bycatch matches relatively well the landed bycatch data.

However, there was a lot of sparsity in the data to reconstruct bycatch pre-1982. What the authors did was to reconstruct the bycatch ratio back in time using the post-1982 data while allocating much uncertainty to the historical period. I believe the authors did well for this step, especially for accounting various levels of uncertainty. I do not have any major comments. One minor suggestion would be to fit a time series model to better track any potential trend in the ratio values but this would not have a large impact due to the large level of uncertainty around these values.

## Review the method for estimating total yelloweye removal:

Total yelloweye rockfish removal was calculated as the sum of directed fishery catch, bycatch by the halibut fishery (the main source of mortality in the non-directed fishery and described in section 3.1), potential foreign fleet bycatch in the historical period, recreational fishing, subsistence catch, and research catch.

# Examining the foreign fleet component

Foreign fleet fishing has been banned in the SEO since 1982. But before that period, the author needed to account for their contribution to the total yelloweye rockfish removal. The reconstruction work was done hierarchically. First, they estimated the amount of non-POP rockfish among all the rockfish caught by the foreign fleet in the Gulf of Alaska (GOA) using the more recent data to derive the proportional relationship. Then they determined the amount of yelloweye rockfish in the GOA among these non-POP rockfish, again basing it on more recent data to derive the proportion. Finally, they applied the same approach but this time to calculate the proportion of yelloweye rockfish in the SEO region compared to the yelloweye rockfish in the GOA. In all three steps, the authors used the mean value during the more recent time period. However, some of these proportions seem to have changed over time (e.g. the proportion of non-POP seems to have increased over time between 1973-1986), thus a time series approach would be preferable in that case. That being said, the authors carefully included the uncertainty around these estimates (quite large values) and while the average reconstructed foreign catch could differ between different approaches, the uncertainty will likely overwhelm the trend. So I suspect the end results would be almost the same but it would be worth checking.

# Removal from subsistence

# The approach seems reasonable. A minor comment (minor because of minor impact on total catch) is whether the creel sample is representative of the subsistence harvest?

1. The research catch

The approach seems reasonable.

1. Recreational fishery removals

The approach seems appropriate although I did not read in detail the method described in Howard et al. 2020 as it was not the main focus of this CIE review.

## Should any of this estimation process be adopted into the SPM?

The estimated total yelloweye catch removal included a lot of uncertainty in the historical period but this is how it should be to reflect the reality of the situation. As stated in the ToR 1 response, the inclusion of the different “components” of the SS-SPM should be simulation-tested (this includes the inclusion of the pre-1980 data with the corresponding uncertainty). If the SS-SPM performs well/better with the inclusion of the historical period then the catch reconstruction work is definitely worth doing. Moreover, some effort should also be put to test the sensitivity of the SS-SPM to changes in the reconstructed time series (see some of my suggestions above). That being said, the historical time period only contains information on catch (with large uncertainty and was not much larger than more recent periods) thus is unlikely to inform on current stock status nor on the absolute scale of the population. Additionally, one does not know if the yelloweye rockfish population characteristics have stayed the same since 1900 (i.e. that the population has not changed its carrying capacity, growth rate, process error rate).

On another note, a question of “double dipping” the IPHC survey data was mentioned during the RW i.e. the yelloweye rockfish survey catch data is used both in the CPUE standardization and for bycatch estimation. From my point of view, it is not a problem in this case as the IPHC survey data follows a proper sampling design (thus does not have a biased selection process). Moreover, the CPUE standardization uses the yelloweye rockfish CPUE data whereas the bycatch model uses the relative biomass of yelloweye caught per unit of legal (or total) halibut biomass. We are thus dealing with two different response variables (though they could be correlated). In species distribution modeling literature, people often model (jointly or independently) the distribution of multiple species using the same data and it has never been an issue. I believe that this issue could have been problematic if the data had some selection bias and that the analytical methods (i.e. CPUE standardization and the bycatch model) were not able to remove the selection bias. In which case, the bias could have been amplified and impacted the SS-SPM outcomes (but this is probably not the case).

# TOR 4 - Provide advice and recommendations on a framework for simulation testing the surplus production model to evaluate similarities between this application of surplus production models and age-structured methods typically used in the North Pacific Fishery Management Council Tier management system.

## Simulation recommendations to examine the use of yelloweye SS-SPM for management purposes by the NPFMC

Much of the simulation recommendation was already provided in my answer to ToR1. To summarize the authors should definitely perform some simulation-estimation testing to examine the contribution of different “components” of the SS-SPM to the different outcomes (e.g. biomass trajectory) but also conduct an MSE to determine the performance of SS-SPM in management context.

## Evaluate the similarity between surplus production model and an age structure model

The purpose of this question was hard to understand as I did not get the necessity to link the SS-SPM to an age-structure model. Especially knowing that the 2015 attempt to build an age-structure assessment model for yelloweye rockfish had failed with model fitting issue and lack of informative data. That being said, if such task of linking SS-SPM and an age-structured model is to be pursued, I recommend the authors to apply a method similar to that of Winker et al. 2020 i.e. simulate from a generic age-structure equilibrium model fitted to the yelloweye rockfish case study as much as possible based on existing information and expert-knowledge (while considering parameter uncertainty of course i.e. conducting many tests using ranges of possible values for the input parameters) and fitting the developed SS-SPM model (using priors for the Pella-Tomlison parameters generated as in Winker et al. 2020) to determine whether one can estimate reliably the management relevant parameters. The report from Cox et al. 2020 (as well as the results from the previous yelloweye assessment attempt in 2015) could become handy in narrowing down the range of life history parameters to be tested for the yelloweye rockfish age-structure model.

# APPENDIX 1 - BIBLIOGRAPHY

*The reviewer report shall include the following appendices:*

1. *Appendix 1: Bibliography of materials provided for review*
2. *Appendix 2: A copy of the CIE Performance Work Statement*

# Prior to the Workshop, reference materials were provided to the CIE reviewers. These materials were comprehensive and relevant for understanding the SS-SPM model. During the workshop one presentations were given, and additional materials were provided on request. The presentations were made available a few days prior to the meeting.

**Bibliography of materials provided for review**

Prior to the Workshop, the following materials were provided:

* Joy et al. 2022. 14: assessment of the demersal shelf rockfish stock complex in the southeast outside subdistrict of the Gulf of Alaska (<https://apps-afsc.fisheries.noaa.gov/Plan_Team/2022/GOAdsr.pdf>)
* <https://meetings.npfmc.org/CommentReview/DownloadFile?p=32eee72a-2fc4-46f6-bd2b-9011ea8e3577.pdf&fileName=C5%20GOA%20Groundfish%20Plan%20Team%20Minutes.pdf>
* <https://meetings.npfmc.org/CommentReview/DownloadFile?p=d95d28fe-3540-4e74-baa3-f029ce6a3a7d.pdf&fileName=SSC%20Report%20Oct%202022_Final.pdf>
* <https://meetings.npfmc.org/CommentReview/DownloadFile?p=39d6577b-136c-49e4-b17e-03dd78659c41.pdf&fileName=C5%20GOA%20Groundfish%20Plan%20Team%20Minutes%20November%202022.pdf>
* <https://meetings.npfmc.org/CommentReview/DownloadFile?p=d94f4b3e-7e21-4e4f-92fd-e39141acfc4a.pdf&fileName=SSC%20Report%20Dec%202022_DRAFT%20to%20COUNCIL.pdf>
* <https://www.npfmc.org/fisheries-issues/fisheries/goa-groundfish-fisheries/>

During/following the workshop, additional references were mentioned and/or examined:

* Cox, S.P., Doherty, B., Benson, A.J., Johnson, S.D.N., and Haggarty, D.R. 2020. Evaluation of potential rebuilding strategies for Outside Yelloweye Rockfish in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/069. viii + 135 p.
* Henrik Sparholt, Bjarte Bogstad, Villy Christensen, Jeremy Collie, Rob van Gemert, Ray Hilborn, Jan Horbowy, Daniel Howell, Michael C Melnychuk, Søren Anker Pedersen, Claus Reedtz Sparrevohn, Gunnar Stefansson, Petur Steingrund, Estimating Fmsy from an ensemble of data sources to account for density dependence in Northeast Atlantic fish stocks, ICES Journal of Marine Science, Volume 78, Issue 1, January-February 2021, Pages 55–69, <https://doi.org/10.1093/icesjms/fsaa175>
* Horbowy, J., & Hommik, K. (2022). Analysis of *F*msy in light of life-history traits—Effects on its proxies and length-based indicators. *Fish and Fisheries*, 23, 663–679. <https://doi.org/10.1111/faf.12640>
* Williams, R., Thomas. 2009. L. Cost-effective abundance estimation of rare animals: Testing performance of small-boat surveys for killer whales in British Columbia. Biol. Conserv. <https://doi.org/10.1016/j.biocon.2008.12.028>
* Thorson J.T., Cope J.M., Branch T.A., and Jensen O.P. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. *Canadian Journal of Fisheries and Aquatic Sciences*. **69**(9): 1556-1568. <https://doi.org/10.1139/f2012-077>
* ICES (2018). Technical Guidelines - ICES reference points for stocks in categories 3 and 4. ICES Technical Guidelines. Report. <https://doi.org/10.17895/ices.pub.4128>
* Winker, H., Carvalho, F., Sow, F.N., Ortiz, M., 2018. Unifying parameterizations between age-structured and surplus production models: An application to Atlantic blue marlin (Makaira nigricans ). ICCAT-SCRS/2018/092 1–16. <https://www.iccat.int/Documents/CVSP/CV076_2019/n_4/CV076040219.pdf>
* Fletcher, R.I. 1978. On the restructuring of the Pella–Tomlinson system. Fish Bull. 76: 515–521. <https://fisherybulletin.nmfs.noaa.gov/sites/default/files/pdf-content/1978/763/fletcher.pdf>
* Meyer, R., Millar, R.B., 1999. BUGS in Bayesian stock assessments. Can. J. Fish. Aquat.Sci. **56**(6): 1078-1087. <https://doi.org/10.1139/f99-043>
* Monnahan, C.C., Thorson, J.T. and Branch, T.A. (2017), Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. Methods Ecol Evol, 8: 339-348. <https://doi.org/10.1111/2041-210X.12681>

# APPENDIX 2 - Performance Work Statement for Kotaro Ono

***Gulf of Alaska Demersal Shelf Rockfish Assessment (CLIN 0003) CIE review***

## 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[[2]](#footnote-2).

## Scope

The stock assessment for Demersal Shelf Rockfish Complex in the Southeast Outside Subdistrict of the Gulf of Alaska (GOA) provides the scientific basis for the management advice considered and implemented by the North Pacific Fisheries Management Council (NPFMC). This stock assessment is conducted by the State of Alaska in partnership with the Alaska Fisheries Science Center (AFSC) and provides management advice in the federally regulated Gulf of Alaska Fishery Management Plan. An independent review of this stock assessment is requested by the Alaska Fisheries Science Center’s (AFSC) Auke Bay Laboratories Division (ABL) and the Alaska Department of Fish and Game (ADF&G). 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. In 2022, the stock assessment authors introduced new methodologies for assessing a major component species of the DSR Complex, Yelloweye Rockfish. Review of these new methods and their application to the NPFMC harvest control rules are requested. The specified format and contents of the individual peer review reports are found in Annex 1. The Terms of Reference (TOR) 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 TOR below. The reviewers shall have a working knowledge and recent experience in the application of complex stock assessment methods and in particular application of Bayesian state-space surplus production models (SSSPM).

Additionally, the CIE reviewers shall have:

* Expertise with measures of model fit and evaluation, uncertainty, forecasting, and biological reference points;
* Expertise in the application of surplus production models in providing management advice;
* Familiarity with Alaska groundfish fisheries and management;
* Familiarity of the assessment and management of stock complexes;
* Working knowledge of the use of fishery and survey data in stock assessment;
* Familiarity with North Pacific Fishery Management Council harvest control rules and determination of reference points used in Alaska fisheries management;
* Excellent oral and written communication skills to facilitate the discussion and communication of results.

This panel review will also have a Chair, who is in addition to the CIE reviewers. However, the Chair’s participation (e.g., labor and travel) is not covered in any way by this task order.

**Tasks for reviewers**

1. Review the following background materials and reports prior to the review meeting. Two weeks before the peer review, the NMFS and ADF&G Project Contacts 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. Attend virtually and participate in the panel review meeting. The meeting will consist of presentations and discussions with the ADF&G stock assessment authors, NMFS staff, and supporting experts of fishery and survey information to facilitate the review.
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 TOR, 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 TOR.
5. Deliver their reports to the Government by the specified deadline.

Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

**Pre-review Background Documents:** Review the following background materials and reports prior to the review:

The following document is the most recent Demersal Shelf Rockfish Complex stock assessment (November, 2022).

Joy et al. 2022. 14: ASSESSMENT OF THE DEMERSAL SHELF ROCKFISH STOCK COMPLEX IN THE SOUTHEAST OUTSIDE SUBDISTRICT OF THE GULF OF ALASKA .

<https://apps-afsc.fisheries.noaa.gov/Plan_Team/2022/GOAdsr.pdf>

NPFMC peer-review bodies provided comments and recommendations to the authors regarding the implementation of the Bayesian state-space surplus production models. The links below provide comments from the NPFMC GOA Groundfish Plan Team and the NPFMC Scientific and Statistical Committee (SSC).

September, 2022 GOA Groundfish Plan Team Minutes (page 3)

<https://meetings.npfmc.org/CommentReview/DownloadFile?p=32eee72a-2fc4-46f6-bd2b-9011ea8e3577.pdf&fileName=C5%20GOA%20Groundfish%20Plan%20Team%20Minutes.pdf>

October, 2022 NPFMC SSC Minutes (page 23)

<https://meetings.npfmc.org/CommentReview/DownloadFile?p=d95d28fe-3540-4e74-baa3-f029ce6a3a7d.pdf&fileName=SSC%20Report%20Oct%202022_Final.pdf>

November, 2022 NPFMC GOA Groundfish Plan Team Minutes (page 6)

<https://meetings.npfmc.org/CommentReview/DownloadFile?p=39d6577b-136c-49e4-b17e-03dd78659c41.pdf&fileName=C5%20GOA%20Groundfish%20Plan%20Team%20Minutes%20November%202022.pdf>

December, 2022 NPFMC SSC Minutes (page 46)

<https://meetings.npfmc.org/CommentReview/DownloadFile?p=d94f4b3e-7e21-4e4f-92fd-e39141acfc4a.pdf&fileName=SSC%20Report%20Dec%202022_DRAFT%20to%20COUNCIL.pdf>

Management advice provided by NPFMC stock assessments conforms to specified harvest control rules. The NPFMC uses an Overfishing Control Rule (OFL) referred to as the “Tier System”. Any model output or management advice must provide reference points applicable to these harvest control rules.

<https://www.npfmc.org/fisheries-issues/fisheries/goa-groundfish-fisheries/>

Two weeks before the peer review, the NMFS Project Contact will send by electronic mail or Google drive to the CIE reviewer all necessary background information and reports for the peer review. In addition to the documents cited above, the Project Contact will provide pertinent code and model development materials. 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, for example:

**Test:** Prior to this peer virtual review, the CIE reviewers will participate in a test to confirm that each reviewer has the necessary technical specifications provided in advance of the panel review meeting to participate virtually.

**Virtual Peer Review:** The CIE reviewers will attend and participate in a virtual peer review. The meeting will consist of presentations and discussions with the NMFS Project Contact, assessment authors, and other staff to facilitate the review.

**Independent Review Report:** Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and TOR, and shall not serve in any other role unless specified herein. Modifications to the PWS and TOR cannot be made during the peer review, and any PWS or TOR modifications prior to the peer review shall be approved by the Contracting Officer’s Representative (COR) and the CIE contractor.

**Contract Deliverables - Independent CIE Peer Review Reports:** Each CIE reviewer shall complete an independent peer review report in accordance with the PWS. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each TOR as described in Annex 2. A draft meeting agenda is provided in Annex 3.

## Place of Performance

Each CIE reviewer shall conduct an independent peer review following a virtual review meeting, therefore no travel is required.

## Period of Performance

The period of performance shall be from the time of award through **November 2023**. Each reviewer’s 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 |
| Two weeks prior to the review | Contractor provides the pre-review documents to the reviewers. |
| **September 12-13, 2023** | Panel review meeting |
| Within two weeks after review | Reviewers submit draft peer-review reports to the contractor for quality assurance and review |
| Within two 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 (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

**Travel**

Since this is a virtual review, travel is neither required nor authorized for this contract.

# Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

# NMFS Project Contacts:

Chris Lunsford

Supervisory Fish Biologist

NOAA/NMFS/AFSC

[Chris.Lunsford@noaa.gov](mailto:Chris.Lunsford@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 or not the science reviewed is the best scientific information available.
2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer’s Role in the Review Activities, Summary of Findings for each TOR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the TOR.
3. The reviewer report shall include the following appendices:
   1. Appendix 1: Bibliography of materials provided for review
   2. Appendix 2: A copy of the CIE Performance Work Statement

**Annex 2: Terms of Reference for the Peer Review**

CIE reviewers are contracted to complete their independent peer review based on the TOR. Therefore, the CIE-NMFS review and approval process is based on whether the CIE independent reports addressed each TOR. The AFSC requests a virtual review in September 2023 to review the applicability and use of Bayesian state-space surplus production models in the Demersal Shelf Rockfish Complex stock assessment in the North Pacific, specifically related to survey averaging methods. CIE reviewers shall address the following TOR during the peer review and in the CIE reports.

1. Evaluate the use of the proposed Bayesian state-space surplus production models, and model methods for use in management of the yelloweye rockfish portion of the GOA Demersal shelf rockfish complex. Specifically, determine if results suggest that the NPFMC Tier 5 approach is appropriate. Should biomass estimates from such a model be used in place of the survey average method currently used for most Tier 5 stocks managed by the NPFMC?
2. Evaluate the management of the GOA demersal shelf rockfish complex as a whole including examination of the use of available data for species other than yelloweye rockfish and determination of Tier level designations.
3. Review the methods used for estimating yelloweye bycatch in the directed Pacific halibut fishery and for estimating total yelloweye catch removals.
4. Provide advice and recommendations on a framework for simulation testing the surplus production model to evaluate similarities between this application of surplus production models and age-structured methods typically used in the North Pacific Fishery Management Council Tier management system.

**Annex 3: Alaska Demersal Shelf Rockfish CIE Agenda**

Virtual Panel Review

Git link to DSR assessment files: <https://github.com/commfish/seak_seo_dsr>

Agenda Start/End Times:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | AK Time | NZ Time | Scotland Time | Norway Time |
| Start | 10:00am | 6:00am | 7:00pm | 8:00pm |
| End | 2:00pm | 10:00am | 11:00pm | 12:00am |
| Dates | Sept 12-13 | Sept 13-14 | Sept 12-13 | Sept 12-13 |

September 12, Alaska Standard Time

|  |  |
| --- | --- |
| 10:00am | Introductions/Questions (Lunsford) |
| 10:15am | DSR stock assessment overview (Joy) |
| 10:45am | Survey inputs to DSR assessment (Joy) |
| 11:15am | Yelloweye fishery catch estimates (Joy) |
| 12:00pm | Break |
| 12:15pm | Yelloweye surplus production model (Joy) |
| 2:00pm | End |

September 13, Alaska Standard Time

|  |  |
| --- | --- |
| 10:00am | Summary of Sept 12 (Lunsford) |
| 10:15am | Followup questions from Day 1 |
| 10:45am | Comparing surplus production results to NPFMC Tier system (Lunsford) |
| 11:15am | Review of TOR 1 (yelloweye model evaluation) |
| 12:00pm | Break |
| 12:15pm | Review of TOR 2 (DSR complex management) |
| 12:45pm | Review of TOR 3 (estimating yelloweye removals) |
| 1:15pm | Review of TOR 4 (recommendations for comparing to accepted NPFMC methods) |
| 1:45pm | Wrapup |
| 2:00pm | End |

Agenda specifics to stock assessment - Phil Joy

1. SEO DSR description
2. A brief history of the assessment
3. Rational for considering a surplus production model
4. Abundance indices
   1. ROV based biomass estimates
   2. IPHC FISS cpue estimates
5. Random Effects Model (current management)
   1. Basics and application to SEO yelloweye
6. Catch
   1. Known catches
   2. Bycatch and discard estimates from the halibut fishery
7. Tier 6 stocks (non-yelloweye DSR)
8. Surplus Production Model
   1. Model review
   2. Prior development
   3. Model results
   4. Risk analysis
   5. Preliminary simulation results
9. Age and length data
   1. What’s available?

1. Institute of Marine Research, Bergen. Norway

   Email: kotaro.ono@hi.no [↑](#footnote-ref-1)
2. <https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf> [↑](#footnote-ref-2)