Center for Independent Experts (CIE) Independent Peer Review Report of the Big and Longnose Skate STAR Panel

By

Robin Cook

Prepared for
Center for Independent Experts
Independent System for Peer Review

Contents

Executive Summary

i. Assessments of big and longnose skate in Oregon and California, and Washington were reviewed at the Northwest Fisheries Science Center, Seattle, WA, during a formal, public meeting of fishery stock assessment experts from 3-7 June 2019. Two Center for Independent Experts (CIE) reviewers were included in the review panel.

ii. The assessments represent the best science available given the existing data, which is very limited, and the decision to use stock synthesis as the principal modelling tool. Both stocks are estimated to be not over-fished and are not subject to over-fishing.

iii. There are a number of fishery independent surveys available. The indices appear to contain some population signal, especially the west coast shelf survey (WCGBT) but these do not to contribute much to the estimated stock biomass trajectory from the assessment.

iv. The estimates of stock biomass for both species is largely determined by the WCGBT and this is mostly driven by the highly informative q prior. It means that biomass reference points are highly uncertain though ratio reference points such a percent depletion will be more robust. During the review the prior on q for big skate was revised and appears to be an improvement on the pre-star assessment.

v. The choice of weighting method in SS3 had an important effect on the estimates of growth parameters, especially for longnonse skate. The use of Dirichelet-multinomial weighting appeared to improve growth and natural mortality estimates for this species. The sensitivity of the assessments to weighting method is a concern given the poor fit to the survey data which in turn affects the growth parameter estimates.

vi. The WCGBT survey appears to offer more information on stock dynamics than is apparent from the SS3 assessments. It seems to be under-utilised and more focus should be placed on analysis of the survey to better understand the population dynamics of the stocks. Simple stage-based or projection matrix models may be options to explore.

vii. The catch data in the assessment are treated as exact and fixed in the model. While this is probably a necessary assumption it is clearly unrealistic. Discards dominate the early catch data and are subject to very high uncertainty. However, it appears that the fishery catch is a minor contributor to the stock dynamics and represents a small fraction of the total mortality.

viii. Natural mortality is estimable in the models and accounts for most of the mortality acting on the population. As it is assumed constant by size and over time, the assessments are not able to capture true population dynamics. Consideration should be given to modelling M by size using a relationship such as that estimated by Lorenzen and scaled to a mean value given by the Hamel or similar method.

ix. Priority should be given to the collection and processing of more age samples for both stocks. Data collection needs to be maintained to create a coherent time series of observations. This will facilitate the estimation of growth parameters and the estimation of any recruitment signal.

x. Thought needs to be given to the appropriate level of model complexity to ensure that the final base model fitted to the data also has the appropriate forecasting properties. A procedure needs to be developed to identify the most parsimonious model using an information statistic and the parameter correlation matrix.

xi. Stock Synthesis software (SS3) provides an impressive range of diagnostics to aid model development. In its present implementation it provides asymptotic variance estimates for the parameters and quantities of interest. This is something of a limitation as it hinders identifying problematic model fits and understanding the relative contribution of priors and data to the estimates. MCMC runs drawn from the reference models would produce more realistic estimates of posterior distributions and should be a routine output of the analyses.

xii. The review meeting was constructive and productive with effective excellent co-operation from the STAT teams. Meeting facilities were good and the local staff provided great support to the reviewers. There were no major disagreements between Panel members or the STATs.

Background

The National Marine Fisheries Service and the Pacific Fishery Management Council held a stock assessment review (STAR) panel meeting in June 2019 to evaluate and review benchmark assessments of Pacific coast big and longnose skate stocks.

Historically, big and longnose skates were mainly discarded but in more recent years a market has developed that has encouraged the landing of these fish. Both species were assessed as unit stocks in US waters. This is the second full assessment of longnose skate, which was last considered in 2007, and the first full assessment of big skate to determine the population status off the west coast of the United States.

The technical review of pre-STAR assessments took place during a formal, public, meeting of fishery stock assessment experts from 3-7th June in Seattle, WA. Two CIE reviewers were included in the review panel. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2.**

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

Materials for the review were made available on 20th May. These were studied prior to the meeting in preparation for the review. During the meeting the reviewer took an active role in discussions. Requests for additional analyses for the STAT were noted and responses collated into a summary for the STAR panel report. The STAR panel report was prepared and agreed by correspondence after the meeting.

Summary of Findings for each ToR

Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.

The draft stock assessment documents were reviewed. These covered assessments of big skate and longnose skate in Washington, California and Oregon. In addition, material relating to previous STAR panel reviews were studied.

Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.

Both assessments use data quantifying total catches by fleet, indices of abundance, length compositions and conditional length at age data.

Catch data

These data are perhaps the most important input to the assessment as they provide information on fishing mortality and normally help scale the assessment to the real fishable biomass. Each assessment attempts to characterize removals dating back to the beginning of the fishery which is judged to be 1916. It is generally considered that "good" catch data are available from the mid-1990s onwards and that the early data are subject to much uncertainty. This is in part the result of the way species were recorded historically and due to the high rates of discarding in earlier years. For big skate, historical discards have been estimated by dividing the landings by the discard rate. Not only are both these quantities subject to estimation error but the discard rate is very high (more than 90%), which means the scale of the derived discards is very sensitive to small errors in the assumed discard rate. It would be highly desirable to find a more robust method. The problem is less severe for the longnose skate, because discards are estimated from the landings of associated species such as Dover sole and the issue of the expansion factor is mitigated.

A considerable amount of effort has gone into the reconstruction of the catch time series but ultimately it is reliant on pragmatic assumptions about the development of the fishery, discard rates, discard survival rates and splitting aggregate skate landings into each species. These uncertainties mean that the catch data are subject to error and possibly bias. The assessments assume the catch is fixed (and by implication error free) which means errors and bias in the data are forced into the estimates of stock biomass and exploitation rate. While it is not possible to recover accurate data from historical records, a study which attempted to quantify likely uncertainty and bias would make a valuable contribution to understanding the veracity of the estimated stock trends from the assessments and provides a basis for well-designed sensitivity tests. In the longnose skate, for example, a sensitivity test that doubled or halved the historical catches does not really provide much insight into uncertainty in annual variability and simply rescales the stock biomass with little impact on trend. It would be of value to consider how inter-annual variation in the catch resulting from uncertainty in the catches alters the perception of stock changes.

Indices of abundance

A number of fishery independent surveys were available for the assessments. These include the NWFSC shelf (WCGBT) and slope surveys, the AFSC slope survey, the triennial survey and the IPHC survey.

For big skate, only the triennial survey and the WCGBT surveys were included. The IPHC was excluded because of low encounter rates with the species. The two-survey series used barely overlap temporally, which means that in effect there is only one series available for any time period to inform stock trends, and given the very different survey design, this will add to uncertainty. Indices were standardized using the VAST package assuming either lognormal or gamma errors and compared with a design based swept area index. All three derived indices show similar trends, but with some change of scale and differences in the estimated CVs. The lognormal tended to give the largest estimated CVs, however the gamma model gave better diagnostics and was chosen for use in the assessment. The similarity of the indices between models provides reassurance of their robustness to the choice of model, but there is clearly an issue with the adequacy of the estimates of precision.

For the longnose assessment, all five indices were used. With the exception of the IPHC, the indices were derived in the same way as the big skate assessment using VAST with gamma errors. Comparison of these with indices derived from design-based estimates showed only minor differences in scale, trend and precision. For the IPHC a GLM approach was used assuming binomial errors to reflect catch per hook. The indices tend not to show strong temporal trends, although the 2004 triennial index is noticeably higher than earlier years, a feature not seen in the other indices.

Length compositions

Available length composition data for surveys and the fishery began in the mid-1990s for both species. Annual sample sizes at fleet level are generally low while the WCGBT survey provides the greatest samples. The latter is likely to be the most important fishery independent data source for both species. Visual inspection of the length frequency data does not show much evidence of year class strength signal. Age data, therefore, are likely to be critical for the estimation of growth parameters and mortality rates.

Given that most data are for the post-1995 period, it means that there is almost no information on age structure for the early period of the assessment. With the uncertainty in the early catch data the interpretation of the estimated stock trends pre-1995 requires considerable caution.

Age compositions

A limited amount of age data is available from about 2004 onwards. For big skate, most of the fishery age compositions are from Oregon and may not be representative of the whole stock. More data are available from the WCGBT survey from 2009 and are likely to represent better areal coverage. For longnose skate some age data are available for four years from 2003 but sampling levels are low. Age reading studies suggest that age determination is reasonably accurate. However, estimates of growth suggest skate cease growing at older ages which calls into question whether age reading beyond the age of growth cessation is really feasible.

Age structured data is most effective when a year class is sampled regularly throughout its lifetime, so that an accurate picture of its survival rate can be estimated. This tends to be lacking in the data for these assessments, especially for longnose skate and means that estimation of recruitment deviations is not possible. For these stocks, the age data are most likely to enable the estimation of growth parameters in the model and contribute to the estimation of total mortality (natural mortality and fishing mortality combined), but are unlikely to facilitate the estimation temporal trends in mortality due to samples being confined to a few recent years.

Evaluate model assumptions, estimates, and major sources of uncertainty.

Model framework

The assessments make use of the latest version of Stock Synthesis (SS3). This is a flexible modelling framework that can make use of a variety of disparate data and is particularly useful when time series data are discontinuous or where there are intermittent observations on length or age. It is therefore an appropriate choice for the assessments considered at the meeting.

Maximum likelihood forms the basis for parameter estimation but can be modified through the use of penalty functions referred to as priors. The model is therefore founded in maximum likelihood but leans toward a Bayesian approach by incorporating prior information. However, as currently implemented, parameter estimates are characterized by point estimates with approximate asymptotic variances rather than their full posterior distributions. MCMC sampling of posterior distributions in other SS3 assessments suggests that the posteriors are typically asymmetric, and the use of asymptotic variances may not therefore adequately characterize parameter uncertainty.

Size composition model

The underlying population model is fully age structured, but it also models the size composition of the population. This is done by assuming growth follows a specified model with dispersion around the mean. The size composition of the population is then reconstructed from the age composition using the length at age distribution. In the assessments considered here, observed length distributions were assumed to follow invariant growth rates. This inevitably raises the question as to whether this somewhat rough growth assumption is sufficiently robust in the light of possible changes in growth by cohort, month and year. This issue may be of most importance for hindcasting the population back to

1916 since there is no data pre-1995 to estimate growth and the projected trajectory is predicated on constancy of growth over a period of nearly 80 years.

The big skate assessment uses a growth cessation model in which fish grow more or less linearly to a maximum length. This maximum length is assumed to be sex specific and is supported by the conditional age-at-length data.

For longnose skate, a conventional von Bertalanffy curve is applied with both sexes following the same growth curve and is supported by sensitivity analysis.

Model parsimony

Each stock assessed reported the parameters that were estimated. These were generally in the region of 40-50. The number of parameters is large when considering the available data. Some thought should be given as to whether such highly parameterized models are really justified by the information in the data. Building a model from the simplest possible might provide insights into just how much complexity is supported by the data and may simplify the number of assumptions that need to be made.

Selectivity

An important element of the SS3 approach is the need to model selectivity. The selectivity curves filter the length composition of the underlying population to explain the observed fleet specific length compositions. For big skate, although dome shaped selectivity was allowed in the model, asymptotic selectivity was the emergent shape, and in the absence of informative data, is likely to be an appropriate form.

In the longnose assessment, the fishery was assumed to have logistic selectivity while the surveys were allowed to have domed selectivity. Fixing at least one fleet to be logistic is desirable , but it should be remembered that the selectivities are estimated relative to each other, and the fishery may have domed selectivity. The form of the selectivity function affects the biomass scale, which in these assessments is poorly determined.

Selectivity for big skate was modelled separately by gender to reflect a higher proportion of males in the catch. This departure from 50:50 sex ratio is not readily explained and may be related to differences in natural mortality.

The limited amount of data in the assessments means that the same selectivity was applied to all the fishery components. It also means this same selectivity is applied in the hindcast stock trajectory implying a constancy for more or less 100 years, which seems unlikely.

Natural Mortality, M

Natural mortality is estimated in the assessments but is informed by a prior based on a range of empirical estimation methods (e.g. Hamel, 2015). This, in effect, provides an estimate of the average annual non-fishing mortality experienced by an individual over its lifetime. In both assessments, the model appears to be able to estimate M adequately with a well-defined posterior distribution. The ability to estimate M is likely due the apparently very low fishing mortality. Most of the total mortality, Z, is explained by M and this will be informed by the recent age and length data. Since the length compositions show little annual change, and age data are sparse, it seems that the model is in effect estimating Z at an approximately steady state corresponding to the post-2004 period (when age data become available).

As the estimated fishing mortality rate is negligible, much of the stock dynamics will be driven by factors external to the fishery. Whatever the true level of M, it is likely to vary over time, and since M cannot be reliably included in the model dynamically (as there are no data to support it) the interpretation of hindcast stock trends is extremely difficult.

Weighting multinomial data

Length and conditional age compositions are modelled as multinomial distributions where sample size is a critical weighting factor in the likelihood. The problem of identifying the correct effective sample size is well known. It will be most pronounced when the actual number of samples is small because the variability in the observations will be greatest. In both assessments, sensitivity to the choice of weighting was investigated using Francis, Dirichelet-multinomial (D-M) and harmonic mean weighting. In both assessments, data weighting affected the relative weight given to the surveys compared to the composition data. D-M weighting gave less weight to the surveys, tended to fit the growth data better, and is discussed further below. For longnose skate, the D-M weighting resulted in higher estimates of M which appeared to be more realistic and was adopted in the final reference model.

Recruitment model

Both assessments use the Beverton-Holt stock-recruitment function parameterized in terms of steepness, h, and recruitment at unexploited biomass, R0. For these stocks, steepness was fixed at 0.4 and probably represents a sensible choice given the likely low productivity of skates. Plots of estimated stock recruit data indicate there are no data points in the left-hand portion of the plot that would enable steepness to be estimated.

Apparently, both species of skate lay eggs throughout the year. These are large, few in number, and subject to predation by macrofauna. This is unlike many exploited teleosts that have small planktonic eggs produced seasonally in large quantities. One might question whether skate life history is consistent with the discrete time Beverton Holt model that is typically used for such teleost fish.

The lack of any recruitment signal in the length compositions and limited age data that is confined to a few recent years meant that it was not possible to estimate recruitment deviations around the stock-recruitment function. The assessments therefore use deterministic values derived from the Beverton Holt curve and most recruitment is simulated near the plateau of this function. It means that the estimated population trajectory is largely driven by a stable age structure conditioned on a constant value of M, invariant selectivity and an assumption about steepness.

Uncertainty

Systematic sensitivity analysis which considers the principal sources of uncertainty is presented for the assessments. The analyses consider the influence of a range of modelling assumptions on the principal stock metrics.

The results of these sensitivities for longnose skate provides a clear indication of where the main issues occur. They show that the biomass related quantities (SB0, SSB_2019, MSY_SPR) are most sensitive to selectivity, and suggest greatest uncertainty in the estimate of the scale of the biomass. Reference points based on absolute biomass will therefore be highly uncertain, though ratio values such as Bratio_2019 are more robust. However, the latter is sensitive to the highly informative prior on q for the WCGBT survey and will significantly influence estimates of depletion. Growth estimates are likely to be dependent on the age data which is sparse and limited to a few years. Testing the sensitivity of the model to these data was explored during the meeting when D-M weighting was used instead of Francis weighting. This gave a better fit to the growth data and was adopted for a revised base model.

Likelihood profiles for the longnose assessment also demonstrate R0 (effectively a scaling parameter) to be driven largely by the priors as is the case with the WCBGT survey q and underline the ambiguity of biomass scale in the assessment.

The scale of big skate biomass is particularly sensitive to the prior on survey catchability of the WCGBT survey and the assessment is unstable without an informative prior. This is indicative of a general problem of scaling in the assessment, and in common with the longnose skate shows that the absolute scale of the biomass is poorly determined. During the meeting, a new prior for q was developed which accounted for the survey area coverage and was used to specify a new base model.

As noted earlier, the catch data in all assessments are assumed exact or estimated with high precision. This is likely to be a necessary assumption for model convergence though it is clearly unrealistic. For longnose skate, a sensitivity test was performed with the catch halved or doubled and this shows little change from the base model. As the catch represents a small part of the total mortality, the assessment is unlikely to be significantly affected by catch estimation errors. In the big skate case, a range of different assumptions regarding catch were investigated especially in relation to estimating historical discards. While alternative assumptions made substantial changes to estimates of instantaneous fishing mortality (F), the magnitude of F is so small that this will not make a great difference to biomass estimates that are largely driven by scaling parameters such as survey catchability and steepness.

Retrospective runs did not reveal any major problems as data are sequentially removed from the assessments. However, the very restricted recent data makes such an analysis of limited importance, since the removal of recent data quickly exhausts the age data upon which much of the assessments depend.

Jitter analyses suggest that the longnose model converged on the lowest negative log-likelihood. In the case of big skate, about half the negative loglikelihoods returned from jittered runs were larger than the base case and appear to be related to low starting values of log(R0).

Stock status

For both stocks, the limited amount of data means that the specific estimated stock status is subject to high uncertainty while it seems very likely the stock is not over-fished or experiencing over-fishing. Estimated stock trends suggest that both stocks have declined since the notional start of exploitation. This decline is almost entirely dependent on assumptions of constancy in biological quantities such as growth, maturity and M. Depletion estimates are likely to be robust but need to be tested relative to alternative assumptions about steepness. These estimates only quantify current status relative to a theoretical unexploited state and should not be seen as depletion relative to biomass in 1916 which is unknown.

Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.

The two main areas of concern are the biomass scale and the choice of data weighting method. The biomass scaling issue was addressed through the formulation of the prior on q for the WCGBT survey in relation to big skate. However, this remains a major source of uncertainty and is addressed for both stocks in the decision tables by using the q prior to define states of nature.

The choice of Francis or D-M weighting makes a difference to the growth estimates. It was resolved during the meeting in relation to longnose skate by adopting D-M weighting which gave a better fit to the conditional age data with a more realistic estimate of M. This makes sense given the inability of the model to fit the survey data regardless of weighting method. The issue was less important for big skate where the cessation growth model was fit well regardless of weighting, and Francis weighting was retained. However, the problem remains and will require further investigation in the future.

Determine whether the science reviewed is considered to be the best scientific information available.

The principal limitation in these assessments is the available data. Catch data pre-1995 are regarded as uncertain. There is a shortage of length frequency and age data. While there are a number of fishery independent surveys they do not, unfortunately, appear to assist in estimating stock trends though they may help with scaling, provided swept area values are reliable. With these limitations in mind, the analyses are of a very high standard making use of state-of-the-art analytical methods. I would judge the science to be the best available.

Stock Synthesis is now a well-established modelling framework and is well suited to the type and quantity of data available for assessment. It is, however, very complex both in the form of the objective function and the multiplicity of configuration options which can obscure what it is actually doing. By their nature, stock assessment models are over-parameterized and SS3 is no exception. With relatively uninformative data as in these assessments, the model is not well anchored and a wide variety of possible interpretations of scale are possible.

When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.

Data

At present there is a large investment of analytic effort into somewhat limited data. More resources devoted to data collection would be highly beneficial. Priority should be given to the collection and processing of more age samples from both stocks. This needs to be maintained to create a long time series of observations.

Modelling approach

The use of SS3 allows highly complex and parameter rich models to be developed and the assessment models used in the assessments reviewed fall into this category. In general, while exploring complex models is undoubtedly useful, there should a systematic attempt to reduce complexity by critically examining the precision and posterior distributions of the parameters as well as their correlations. This would help in identifying redundancy and may help in improving model stability and predicative power.

The assessment models chosen are one way of incorporating a variety of different data into a comprehensive analysis. However, there may be some merit in applying other simpler approaches to gain insights into the information content of the data and resolve conflicting signals. There are two aspects, somewhat connected, and applicable to both assessments that need to be considered:

1. The model is unable to track to WCGBT survey because natural mortality is assumed constant as is selectivity, while recruitment is deterministic with almost no annual variability. Since fishing mortality is negligible, this means the underlying modeled population shows little variation and the indices seen in the survey are treated by the model as white noise. However, this survey shows autocorrelation between the annual biomass values which strongly suggests it is tracking population change. While the WCGBT survey is probably the best information available on population change, it is in effect being discarded through model mis-specification.
2. For both stocks, using D-M weighting provided a better fit to the growth data than Francis weighting, because these data receive higher weight relative to the survey data. Thus, there is the rather unsatisfactory situation that the survey data (i.e., the annual biomass data) have a direct effect on the estimated growth, yet it is clear the model cannot fit the abundance data due to modelling assumptions. Hence, model mis-specification will affect the growth estimates to a degree dependent on the choice of weighting scheme.

Although the qualitative effects of the weighting method were the same for each stock, Francis weighting was retained for the big skate assessment while D-M weighting was chosen for longnose skate. While the STAT leaders gave plausible reasons for their preference, one cannot escape the feeling that the choice is nevertheless somewhat arbitrary and might in the future be resolved by modelling the survey data more realistically.

In view of these problems, there is value in trying to make more effective use of the WCGBT survey data by analyzing it separately to avoid some of the issues above. This could be done using any number of approaches where a simpler population dynamics model that avoids a full age structure could be applied to the survey data. There does not seem to be any need to hindcast the model back to 1916 since the WCGBT data provide an absolute abundance estimate from 2003, and it would be quite feasible to run the assessment forward from the late 1990s when catches are best known. Not only are the early catch data highly unreliable, but the assumptions of invariance in the hindcast are very unlikely to be correct, making the 1916 estimate of questionable value.

Since the abundance data are not adequately fitted in the SS3 assessment, it is probably better to estimate growth external to model from the conditional length at age and length composition data.

Natural mortality, M, is the principal source of mortality for these two stocks since fishing mortality is negligible. In the assessments presented, M is assumed constant over time and age (or size). Given its importance in determining the death rate, it would be desirable to try to model M more convincingly. Estimating temporal trends would be difficult, but a more realistic treatment of mortality by size should be possible using, for example, the Lorenzen meta-analysis (Lorenzen 1996). M could then be estimated within the assessment model as a scaling constant on a chosen size dependent function.

Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

The review was conducted in a constructive manner and the STAT teams were helpful and extremely responsive to the requests from the Panel for additional analyses with all the essential runs being completed during the meeting.

Many of the issues discussed have been referred to in earlier sections of this report. These included:

- Identifying the q prior assumption for the WCGBT survey that better captured existing knowledge
- The choice of weighting method in SS3

Towards the end of the meeting, there were discussions on the states of nature for decision tables. The q prior was chosen to inform the main axis of uncertainty relating to biomass scale.

Overall, there was effective engagement from all members of the Panel, the STATs and the Panel advisors. This led to improvements in the configuration of the base models.

Recommendations for future assessments are discussed in the next section.

Conclusions and Recommendations

The assessments of big and longnose skate represent the best science available given the existing data and the SS3 software available to the STATs. The analyses were thorough and considerable work had gone into making good use of data from a variety of sources. The limited amount of age data and lack of fit to fishery independent abundance indices, especially the WCGBT, means that despite the elegance of the assessments, it is difficult to have confidence in the estimated stock trends. Since the estimated catch represents a very small fraction of the stock biomass, the conclusion that neither stock is overfished or experiencing over-fishing seems robust, however.

Should managers attach importance to these stocks, then I would **recommend that priority be given to the collection and processing of more age samples from both stocks. This needs to be maintained to create a coherent time series of observations. It is of particular importance to maintain fishery independent surveys to calibrate estimates of biomass.**

In common with many other assessments in this region, early catch estimates are subject to considerable uncertainty. The assumption that catches are exact and treated as fixed in the model is probably necessary but clearly unrealistic. Sensitivity to this problem needs to be adequately investigated. I **recommend that a demanding sensitivity analysis is performed where plausible alternative catch streams are generated stochastically**. The practice of halving or doubling the catch as a sensitivity test is not very demanding and is unlikely to the probe the nature of the uncertainty in the data.

Natural mortality is by far the largest component of total mortality in these stocks and will drive much of the stock dynamics. I was not entirely convinced that modelling M as a constant value was the best approach. I **recommend that the way M is modelled and estimated be reviewed. Consideration should be given to modelling M by size and scaling it to a mean value given by an empirical method.** This could avoid the need to model M by gender and would capture some of its annual variation.

The WCGBT survey appears to offer more information on stock dynamics than is apparent from the SS3 assessments. **It seems to be under-utilised and more focus should be placed on analysis of the survey to better understand the population dynamics of the stocks. Simple stage-based or projection matrix models may be options to explore.**

Since both species of skate appear to lay eggs throughout the year, one might question whether skate life history is consistent with the discrete time Beverton Holt model that is typically used for teleost fish. **Some investigation of the possible importance to continuous recruitment should be undertaken and whether discrete time models are appropriate.**

I recognize that SS3 is a powerful, useful and appropriate tool for the assessment of these stocks. However, thought needs to be given to the appropriate level of model complexity to ensure that the final base model fitted to the data also has the appropriate forecasting properties. I would **recommend that a procedure is developed to identify the most parsimonious model using an information statistic and the parameter correlation matrix.**

SS3 provides an impressive range of diagnostics to aid model development. In its present implementation, it does not appear to provide realistic posterior distributions of the estimated parameters. This is something of a limitation as it hinders identifying problematic model fits and understanding the relative contribution of priors and data to the estimates. I **recommend that SS3 be updated to provide full parameter posterior distributions.**

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

Draft assessment documents and supporting material were made available on the Pacific Fisheries Management Council ftp site two weeks in advance of the meeting. This is realistically the minimum advance time to review the assessments adequately. The principal documents are voluminous and take time to digest. As always, more time would be appreciated and would lead to more considered interventions at the review meeting. Understandably, there is a compromise to be struck between the completion of assessment documents and time available for review. Perhaps the two-week period is the best that can be achieved.

The meeting itself was constructive and productive with effective and excellent co-operation from the STAT teams. Meeting facilities were good, and the local staff provided great support to the reviewers.

References

Hamel, O. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. ICES Journal of Marine Science 72: 62-69.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49, 627–647.

Appendix 1: Bibliography of materials provided for review

The following materials were made available in the PFMC ftp site before and during the meeting. They can be found at ftp://ftp.pcouncil.org/pub/!2019%20GF%20STAR%20Panels/STAR%20Panel%202%20-%20Skates/

<u>Pre-STAR assessment reports</u>

Gertseva, V. Matson, S., Taylor, I. Bizzarro, J, Wallace, J. 2019. Stock assessment of the longnose skate (*Beringraja rhina*) in state and Federal waters off California, Oregon and Washington. Pacific Fishery Management Council, Portland, OR. Available from http://www.pcouncil.org/groundfish/stock-assessments/

Taylor, I.G., Gertseva, V., Stephens, A. and Bizzarro, J. Status of Big Skate (Beringraja binoculata) Off the U.S. West Coast, 2019. Pacific Fishery Management Council, Portland, OR. Available from http://www.pcouncil.org/groundfish/stock-assessments/

<u>Background</u>

Anon. Report of the Skates Catch Reconstruction Workshop, 27 March 2019.

Cordue, P.L. Report on Sablefish and Longnose Skate STAR Panel May 7-11, 2007.

Haist, V. Review of West Coast Groundfish Stock Assessments: sablefish and longnose skate, STAR Panel, May 7-11, 2007.

<u>Presentations</u>

Big_Skate_2019_STAR_Panel_presentation.pptx

LSKT_Skate_2019_STAR Panel_presentation.pptx

In addition to the materials listed above, further documents were made available during the review. These can be found in the ftp site listed above.

Appendix 2: Statement of Work

Performance Work Statement (PWS)

National Oceanic and Atmospheric Administration (NOAA)

National Marine Fisheries Service (NMFS)

Center for Independent Experts (CIE) Program

External Independent Peer Review

Stock Assessment Review (STAR) Panel 2

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 National Marine Fisheries Service and the Pacific Fishery Management Council will hold stock assessment review (STAR) panels and potentially one mop-up panel (if needed), to evaluate and review benchmark assessments of Pacific coast groundfish stocks. The goals and objectives of the groundfish STAR process are to:

1) ensure that stock assessments represent the best scientific information available and facilitate the use of this information by the Council to adopt Overfishing Limits (OFLs), Acceptable Biological Catches (ABCs), Annual Catch Limits (ACLs), harvest guidelines (HGs), and annual catch targets (ACTs);
2) meet the mandates of the Magnuson-Stevens Fisheries Conservation and Management Act (MSA) and other legal requirements;
3) follow a detailed calendar and fulfill explicit responsibilities for all participants to produce required reports and outcomes;
4) provide an independent review of stock assessments;
5) increase understanding and acceptance of stock assessments and peer reviews by all members of the Council family;
6) identify research needed to improve assessments, reviews, and fishery management in the future; and
7) use assessment and review resources effectively and efficiently.

Benchmark stock assessments will be conducted and reviewed for Longnose and Big skates. Both stocks were identified as strong candidates for assessment during the Pacific coast groundfish regional stock assessment prioritization process, with Longnose skate being ranked as second of all considered stocks. This analysis was based on the national stock assessment prioritization framework (http://www.st.nmfs.noaa.gov/Assets/stock/documents/PrioritizingFishStockAssessments_FinalWeb.pdf.

Longnose skate was last assessed as a benchmark assessment in 2007. The spawning stock biomass was estimated to be at 66 percent of its unfished biomass at the start of 2007. Based on that assessment, a constant catch strategy (OY = 1,349 mt) was implemented in 2009 based on a 50 percent increase in the average 2004-2006 landings and discard mortality. The constant catch strategy was revised in 2013 by implementing an Annual Catch Limit (ACL) of 2,000 mt to provide greater access to the stock and to limit disruption of current fisheries. This level of harvest was projected to maintain the population at a healthy level as projected in the 10-year forecast for longnose skate in the 2007 assessment (Gertseva and Schirripa 2008). The Scientific and Statistical Committee (SSC) recommended changing the proxy F_{MSY}[1] rate for longnose skate and other elasmobranchs from a Spawning Potential Ratio (SPR) of 45 percent to an SPR of 50 percent beginning in 2015. This recommendation, driven primarily by conservation concerns for spiny dogfish, was heeded by the Council when they adopted 2017 and 2018 OFLs consistent with this lower harvest rate. The Council adopted the default harvest control rule for longnose skate by recommending a 2019 and 2020 ACL of 2,000 mt. A new assessment is extremely important to inform both current status as well future projections.

[1] F_{MSY} The fishing mortality rate that, if applied constantly, would result in maximum sustainable yield (MSY). Used as a biological reference point, F_{MSY} is the implicit fishing mortality target of many regional and national fishery management authorities and organizations.

Big skate has not been previously assessed, but is an important and growing composition of the west coast groundfish fishery. Big skate were managed in the Other Fish complex until 2015 when they were designated an Ecosystem Component (EC) species. When the Council considered designating all skates except longnose skate as EC species, the Groundfish Management Team (GMT) estimated that catches of big skate averaged 95 mt from 2007–2011 with large landings of Unspecified Skate (see Table 4-33 in the 2015-2016 Harvest Specifications and Management Measures Final Environmental Impact Statement). Subsequent analysis of Oregon port sampling data not available when the Council considered the EC designation indicated about 98 percent of the recent Unspecified Skate landings in Oregon were comprised of big skate. The GMT revised the total mortality estimates of big skate coastwide using these new data (Table 1-10). Such large landings indicates targeting of big skate has occurred and an EC designation was not warranted. Based on this evidence, the Council decided to re-designate big skate as an actively-managed species in the fishery. Big skate were managed with stock-specific harvest specifications starting in 2017.

Assessments for these two stocks will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. including providing scientific basis for setting Overfishing Limits (OFLs) and Acceptable Biological Catches (ABCs) as mandated by the Magnuson-Stevens Act. The technical review will take place during a formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. 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 attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements

Two CIE reviewers will participate in the stock assessment review panel. One CIE reviewer shall conduct an impartial and independent peer review of the assessments described above and in accordance with the Performance Work Statement (PWS) and ToRs herein. Additionally, one “consistent” CIE reviewer will participate in all STAR panels held in 2019 and the PWS and ToRs for the “consistent” CIE reviewer are included in **Attachment A**.

The CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age-and size-structured models, use of Markov Chain Monte Carlo (MCMC) to develop confidence intervals, and use of Generalized Linear Models in stock assessment models. The CIE reviewer’s duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Tasks for Reviewers
The CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the PWS scheduled deadlines specified herein. The CIE reviewer shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to the STAR Panel 2 meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available (including previous stock assessments and STAR panel reports).
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

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

Contract Deliverables - Independent CIE Peer Review Reports: The CIE reviewers 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**.

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

Timeline for CIE Reviewers

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the STAR Panel 2 review meeting in scheduled in Seattle, WA during the dates of June 3-7, 2019 as specified herein, and conduct an independent peer review in accordance with the ToRs.
3) No later than June 21, 2019, each CIE reviewer shall submit their draft independent peer review report to the contractor. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each ToR in **Annex 2**

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 August 2019. The CIE reviewers' duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables

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

Within two weeks of award	Contractor selects and confirms reviewers
At least two weeks prior to the panel review meeting	Contractor provides the pre-review documents to the reviewers
June 3-7, 2019	Each reviewer participates and conducts an independent peer review during the panel review meeting
June 21, 2019	Contractor receives draft reports
July 10, 2019	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 in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; 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.

Restricted or Limited Use of Data

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

NMFS Project Contacts:
Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive
Newport, OR 97365
Stacey.Miller@noaa.gov
Phone: 541-867-0535

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112

Jim.Hastie@noaa.gov
Phone: 206-860-341

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

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

2. The main body of the reviewer's report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

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

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

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

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

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

3. The reviewer's report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review

   Appendix 2: A copy of the CIE 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

Stock Assessment Review (STAR) Panel 2

1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
3. Evaluate model assumptions, estimates, and major sources of uncertainty.
4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Annex 3: Tentative Agenda

Final Agenda to be provided two weeks prior to the meeting with draft assessments and background materials.

Stock Assessment Review (STAR) Panel 2

Longnose and Big Skate

Seattle, Washington

NWFSC

2725 Montlake Blvd, NE

Seattle, WA 98112

June 3-7, 2019

TBD

Appendix 3: Panel membership and participation

STAR Panel Advisors
Mr. Patrick Mirick, Oregon Department of Fish and Wildlife, Groundfish Management Team
Mr. John DeVore, Pacific Fishery Management Council

Technical Reviewers
David Sampson, Scientific and Statistical Committee (SSC), Panel Chair
Robin Cook, Center for Independent Experts (CIE)
Henrik Sparholt, Center for Independent Experts (CIE)
Cody Szulwalski, Alaska Fisheries Science Center (AFSC)

Panel Advisors
Jessi Doepinghaus, PFMC Groundfish Management Team (GMT)
Gerry Richter, PFMC Groundfish Advisory Subpanel (GAP)
John DeVore, PFMC, Staff Officer

Longnose Skate Stock Assessment (STAT) Team
Vladlena Gertseva, NMFS, Northwest Fisheries Science Center
Sean Matson, NMFS, West Coast Region
Ian Taylor, NMFS, Northwest Fisheries Science Center
Joseph Bizzarro, Southwest Fisheries Science Center
John Wallace, NMFS, Northwest Fisheries Science Center

Big Skate Stock Assessment (STAT) Team
Ian Taylor, NMFS, Northwest Fisheries Science Center
Vladlena Gertseva, NMFS, Northwest Fisheries Science Center
Andi Stephens, NMFS, Northwest Fisheries Science Center
Joseph Bizzarro, NMFS, Southwest Fisheries Science Center