Center for Independent Experts (CIE) Review of Gulf of Alaska rex, Dover, and flathead sole stock assessments

Independent Peer Review Report

P.L. Cordue

July 2019

Table of Contents

Executive summary

A CIE review of three stock assessments of Gulf of Alaska (GOA) flatfish stocks was conducted at the Alaska Fisheries Science Center from April 29 to May 03, 2019. The participants included three CIE reviewers, the primary assessment author, the chair of the meeting and NMFS staff who presented on relevant topics. This report is one of three independent reports produced by the CIE reviewers.

Rex, Dover, and flathead sole are all demersal species caught primarily with bottom trawl. Rex and flathead sole are most abundant in depths less than 200 m while the reported range for Dover sole is 300-1500 m. Rex sole has the highest consistent historical catches with a peak of 6000 t in 1996 and the 2010-2017 catches ranged from 1300-3700 t. Dover sole had a sharp peak in catches in 1991 and 1992 at about 9000 t but catches from 2010-2015 ranged from just 220-550 t. Catches for flathead sole have been consistently above 1000 t since 1990 with a peak of 3900 t in 2010 and a range from 2010-2017 of 1600-3900 t. The GOA stocks of these species are estimated to be at high current stock status with low exploitation rates.

The stocks are assessed using age-structured, two-sex models implemented in Stock Synthesis 3. The model structures are appropriate given the available biomass indices and composition data.

An important issue, common to all three stock assessments, is that more care could be taken in the preparation of the input data.

Common to all three assessments is the issue of the use of the 1980s trawl surveys and the use of the 1990 and 1993 surveys. The 1980s surveys should not be used (non-standard vessels and gear); the 1990 and 1993 surveys should probably be used in the base model, but a sensitivity should also be done which excludes them (they were conducted later in the year than the surveys since 1996).

For the commercial fisheries, length, age, and age at length data are all potentially susceptible to unrepresentative sampling despite what appears to be an excellent observer program. The problem arises because coverage for a particular species can be somewhat patchy from year to year. Purely by chance the sampled catch in a particular year can come from an unrepresentative part of the commercial catch. An analysis of the data to determine the drivers of fish length (or age or age at length) followed by appropriate post stratifying and scaling is needed for each stock.

In each assessment it was assumed that the trawl survey had a catchability (q) exactly equal to 1. In general, this is a poor assumption but for these species it was determined, during the review meeting, that *a priori* the q should not be that different from 1. However, the q's should be estimated as free parameters in the models with informed priors.

The most recent assessments for rex (2017), Dover (2015), and flathead sole (2017) are the only available scientific information that assess most recent stock status. As such, they represent the best available scientific information. More importantly, the assessments are qualitatively reliable in the assessment of high current stock status and low fishing mortality. This conclusion is based on the scale of the catches, the scale of the trawl survey indices, and the *a priori* knowledge of the trawl survey q's.

Background

A CIE review of three stock assessments of Gulf of Alaska (GOA) flatfish stocks was conducted at the Alaska Fisheries Science Center from April 29 to May 3 2019. The participants included three CIE reviewers, the primary assessment author, the chair of the meeting and NMFS staff who presented on relevant topics (see Appendix 3). This report is one of three independent reports produced by the CIE reviewers. There is also a summary report that contains a brief summary of the meeting and has recommendations that were supported by all three of the CIE reviewers (see Appendix 4).

Rex, Dover, and flathead sole are all demersal species caught primarily with bottom trawl. The GOA stocks are estimated to be not overfished and not experiencing overfishing. Rex and flathead sole are most abundant in depths less than 200 m, while the reported range for Dover sole is 300-1500 m (McGilliard and Palsson 2015a). Rex sole has the highest consistent historical catches with a peak of 6000 t in 1996, and the 2010-2017 catches ranged from 1300-3700 t. Dover sole had a sharp peak in catches in 1991 and 1992 at about 9000 t, but catches from 2010-2015 ranged from just 220-550 t. Catches for flathead sole have been consistently above 1000 t since 1990 with a peak of 3900 t in 2010 and a range from 2010-2017 of 1600-3900 t.

Rex sole has an extended spawning period from October to May, and they move deeper as they age (McGilliard and Palsson 2017). Dover sole also have an extended spawning period, but it is later than rex being from January to August, with a peak in May (McGilliard and Palsson 2015a). Flathead sole are known to spawn in March and April and may begin as early as January (Turnock et al. 2017). Rex and flathead sole have a moderate lifespan with natural mortality (M) assumed to be about 0.2 while Dover sole is much longer lived with an assumed M of 0.085. The values of M are based on maximum age.

Stock structure for each species is poorly known and a single stock is assumed for each species in the GOA. Biomass indices are available for each species from the GOA trawl survey (originally triennial and now biennial). The surveys also provide length and age data. Age data from the commercial fishery are only available for rex sole (length data are available for all three species). Age structured two-sex models are used in the stock assessments. They were originally implemented in purpose written code but were transitioned to Stock Synthesis 3 in 2013 (Dover and flathead) and 2015 (rex).

Review activities

Prior to the review meeting the stock assessment documents were downloaded using the provided links in the Performance Work Statement (see Appendix 2). The most recent assessment reports, for each of the three species, were read in detail and the earlier assessment reports were briefly reviewed.

The review meeting starting Monday April 29 was attended. When the Terms of Reference (TOR) were being presented, I raised an issue. As written, the TOR focused very much on the stock assessment MODEL rather than the whole stock assessment. The meeting agreed that we would interpret “stock assessment model”, in TOR 2 and 3 for each species, to mean “stock assessment”. This allows us to review the stock assessments rather than just the models (there is a technical distinction).

I also raised the issue of the requirement in the PWS for a clear statement in the executive summary of our individual reports to “specify whether the science reviewed is the best scientific information available”. It is generally the case that the stock assessments under review are the ONLY scientific information available on recent stock status. As the only scientific information available it is clearly the best, but it is also the worst. The question is redundant. An assessment should be based on the best scientific information available, but unless there are competing assessments, it is not a question of whether an assessment is the best available information but rather whether it is appropriate to be used as a basis to provide management advice.

The first day consisted mainly of presentations by NMFS scientists on programs which provide stock assessment data (see Appendix 3). The primary assessment author gave an introductory presentation on the Gulf of Alaska (GOA) ecosystem and the three flatfish species under consideration. She also gave a presentation on the development (over time) of the stock assessments for rex sole.

On Tuesday the reviewers put a brief specification together for a model run for rex sole that incorporated some suggested improvements to the base model. The primary assessment author gave a presentation on Dover sole. I noted that there were some very large positive residuals for the age at length data. There was a suggestion by the chair that the plots of the residuals were perhaps misleading as the large residuals were not where most of the data were present (e.g., older ages, shorter lengths). We agreed that there was a need to think about alternative graphs that could be produced. I also noted the need for a study of the drivers of length, age, and growth for Dover, but also for the two other species to support appropriate post stratification and scaling of composition data (before being used as an input into a stock assessment). We also discussed the need for informed priors on trawl survey catchability (q) rather than assuming q = 1.

On Wednesday the results of the requested rex sole model run were presented. It gave very similar results to the base model but had larger and more realistic estimates of uncertainty. The chair gave a presentation on a first attempt at creating informed priors for the trawl survey q's for Dover when the surveys went to different maximum depths. A specification for a Dover model run was agreed. The meeting was adjourned after the morning session as the chair and the primary assessment author had to attend another meeting from 2-4 pm.

On Thursday, there was a morning session with a flathead presentation including the results of a couple of model runs following the rex specification for improvements. Again, the point estimates were similar to the base model and there were more realistic estimates of uncertainty. The assessment author needed some time to complete the requested Dover model run so the reviewers took the opportunity to do some writing – specifically a draft summary report (not required under the TOR but it was suggested by the chair that we might produce such a report which focussed on recommendations that were agreed to by all three reviewers). The reviewers had no difficulty agreeing on a set of general recommendations that were applicable to each of the three species (and more generally applicable in some cases).

The Dover model run was presented but there had been difficulties obtaining a converged data weighting and sensible estimates of selectivity. Consequently, the model results were not useful. However, the general specification, with some modification, was still agreed to be applicable. The meeting was closed just before 5 pm on Thursday.

After the meeting, the full set of documents and presentations, that had been shared on a Google drive, were downloaded to allow further consideration of the material. There was consultation among the reviewers with regard to the summary report (given comments on the draft from the chair and the assessment author).

Summary of findings

Each of the TOR for each of the species are considered below. Most of the weaknesses are common to the three stock assessments and are only explained in detail for the rex sole assessment. Also, as most of the recommendations apply to all three species, the recommendations are given jointly in a later section.

Gulf of Alaska rex sole

1. *Evaluation of the ability of the stock assessment model for GOA rex sole, with the available data, to provide parameter estimates to assess the current status of rex sole in the Gulf of Alaska.*

The 2017 base model for rex sole is a two-sex age-structured model with two areas (eastern GOA; western and central GOA combined) which have different growth parameters estimated (McGilliard and Palsson 2017). Trawl survey biomass indices are fitted separately by area with q = 1 assumed for both time series.

An existing spatial analysis of residuals for growth data showed markedly different length at age in the eastern GOA compared to the west and central GOA (Figure 1). This obvious growth difference was the basis for the two-area two growth morph model that the authors selected, and it is a useful improvement from previous models.

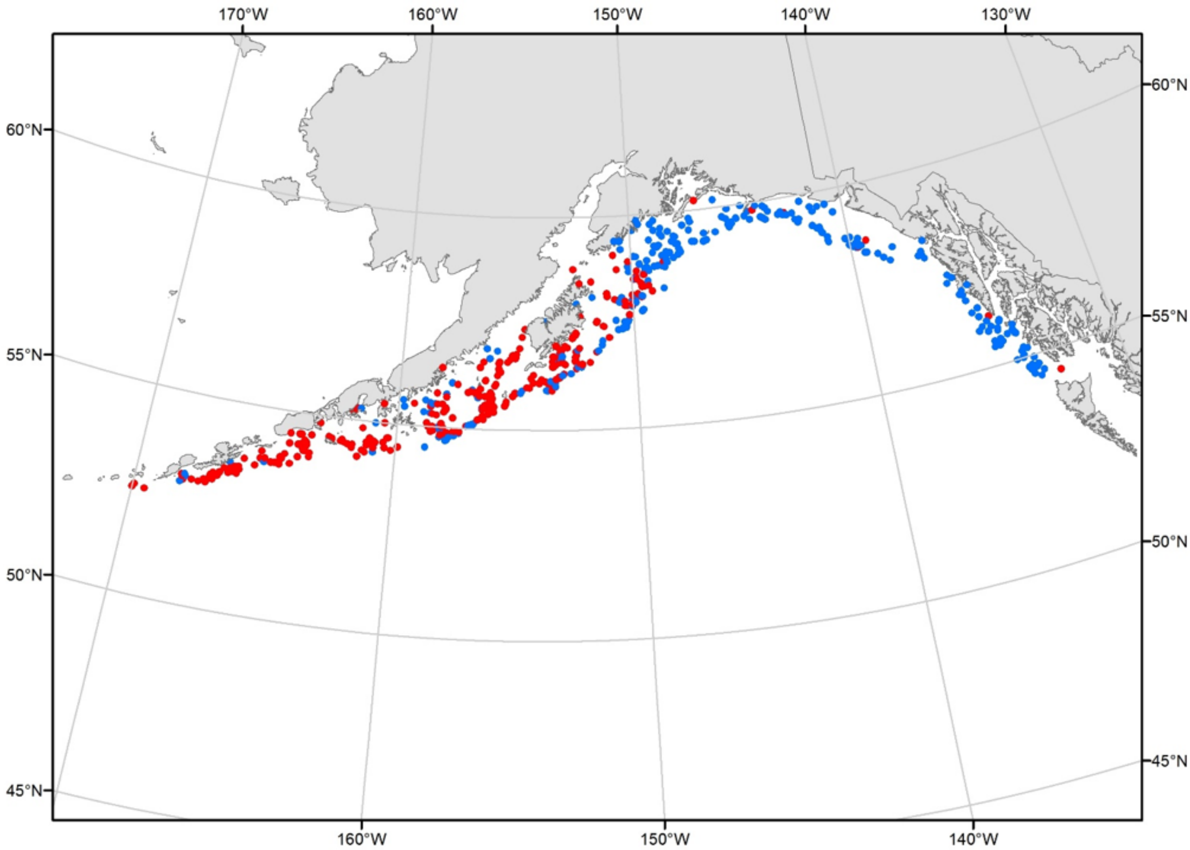

Figure 1: Rex sole residuals for von Bertalanffy growth curves fitted to age and length survey data outside of the model. Points are plotted at the location of the haul from which the fish were sampled. The blue points are more than 1 residual standard error below the curve and the red points are more than 1 residual standard error above the curve. (From McGilliard 2019a, produced by Beth Matta.)

The model is fitted to trawl survey biomass indices assuming that q = 1. This is a poor assumption in general but for this stock and this survey it is adequate. From the inshore distribution of rex sole it is apparent that a maximum survey depth of 500 m (which was achieved in every year of the survey) is adequate to cover the spatial distribution of the population. Also, an experiment with the trawl survey gear supported a vulnerability (for fully selected fish in front of the net) of about 1.2 (Sommerton et al. 2007). The mean value for an informed prior on the trawl survey q should therefore be fixed at about 1.2 (and the s.d. from the experiment was about 0.175).

The reviewers requested a modified model run: remove the 1980s trawl survey data; estimate a single trawl survey q with an informed prior (mean=1.2, s.d.=0.175); and Francis iterative reweighting for the composition data (Francis 2011, Punt 2017).

The model run requested by the reviewers (where q was estimated with an informed prior) gave similar results to the base model with the estimated q not far from the assumed value of 1 (Figure 2).

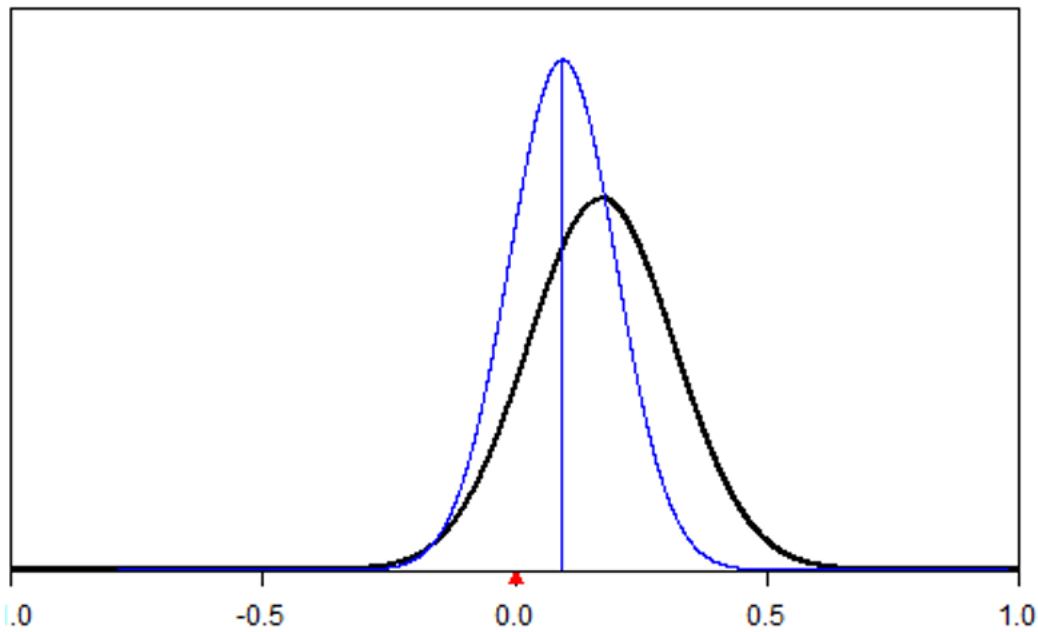

Figure 2: For the run requested by the reviewers: the prior distribution for the rex trawl survey q (black) and the estimated MPD estimate (vertical blue line) and associated approximate posterior distribution (blue). The x axis is the natural log of q (so 0.0 corresponds to a q of 1). From McGilliard 2019a.

The trawl survey biomass indices (eastern and non-eastern) are not fitted particularly well, but since they are essentially flat from 1990 to 2017 (with some short-term oscillation), it is apparent that the historical catches are having little impact on the stock (Figure 3). The scale of the stock as estimated in the model is robust because the trawl survey q is approximately equal to 1 (as estimated in the reviewers' model and as assumed in the base model).

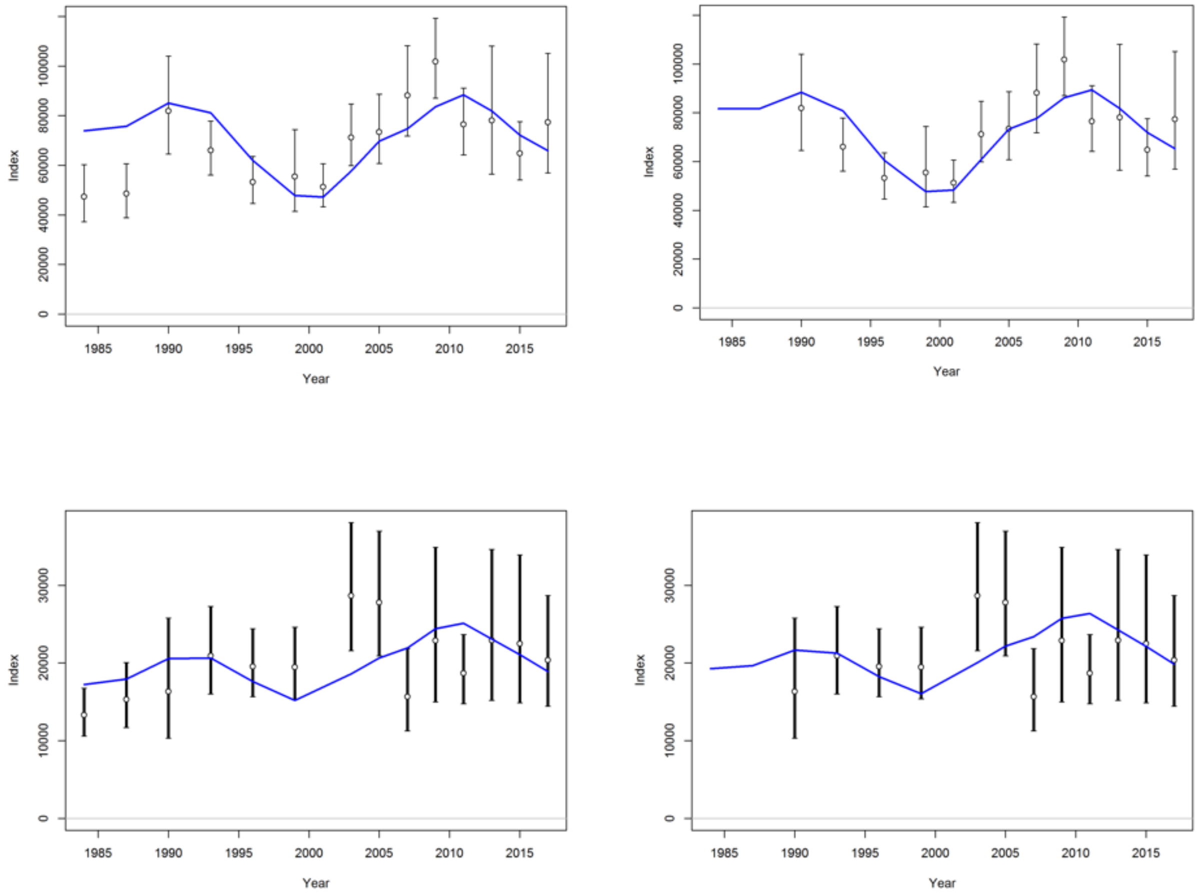

Figure 3: The MPD fits to the rex sole non-eastern (top) and eastern (bottom) trawl survey indices. The plots on the left are for the base model in 2017 and the plots on the right are for the reviewers' run. From McGilliard 2019a.

The base model assessment from 2017 estimates high stock status and low fishing mortality. These estimates are qualitatively reliable. Some improvements to the assessment model will change the exact estimates but will not change the main assessment conclusions.

2. Evaluation of the strengths and weaknesses in the stock assessment model for GOA rex sole

The rex sole stock assessment is the strongest of the three assessments that were reviewed. In comparison to the other two species rex sole has some age data from the commercial fishery which allows more reliable estimates of fishing selectivity to be made. It was the introduction of this age data into the assessment that allowed the stock assessment to move from Tier 5 to Tier 3a as the F based reference points were previously very unreliable (McGilliard and Palsson 2017).

The use of a two-sex, age structured model is appropriate given the available data and knowledge of the biology. The use of two areas is appropriate given the demonstrated growth differences by area. The availability of a consistent time series of trawl surveys since 1990 is a particular strength (shared by all of the reviewed stock assessments).

The assessment does have some weaknesses, but they are not nearly substantial enough to compromise the reliability of the main assessment results.

Preparation of input data
An important issue, common to all three stocks, is that more care could have been taken in the preparation of the input data.

The time series of trawl surveys cannot be considered consistent until 1990. The two surveys in the 1980s used a different survey design (station occupation timing and order) with substantially different vessels and trawl gear (Palsson 2019). These surveys should not have been used in the base model. There is also the question of whether the 1990 and 1993 surveys should be included in the base model. The timing of surveys is later than the remainder of the time series and station occupation was also a bit different. In general, they should probably stay in the base model but should certainly be removed in a sensitivity run. There is also the question of whether the biomass time series has a constant q from year to year. It appears that this species (and also Dover and flathead) could potentially be migrating during the period of the trawl survey. The timing of this migration could change somewhat from year to year and consequently the q (and the selectivity) could be somewhat variable. An analysis of gonad stages of fish examined from the survey catches may be informative as to the level of potential variation if there is a spawning related migration overlapping with the timing of the survey. It may be sensible to add some “process error” to the trawl survey coefficients of variation (CVs).

For the fishery, length, age, and age at length data are all potentially susceptible to unrepresentative sampling despite what appears to be an excellent observer program. The problem arises because coverage for a particular species can be somewhat patchy from year to year. Purely by chance the sampled catch in a particular year can come from an unrepresentative part of the commercial catch.

For example, suppose that the length of fish increases with depth. It may be that in one year a lot of the catch is caught in relatively deep water but by chance, for that species, it happens that most of the samples come from trawl catches that were caught in relatively shallow water. This is a well-known problem and the solution involves an initial analysis of data to determine what factors are important in determining fish length (or age or age at length) and then post stratifying the fishery data according to those factors. And, finally, scaling the data appropriately (by fish numbers from the sample to the haul and then from each haul up to the stratum; and then summing across strata). If one of the factors that drives the measurement of interest is not available for some samples (e.g., because of port sampling, where depth is not available) it may be that those samples cannot be used (or should only be used in a sensitivity).

Use of raw length frequencies is inappropriate in a stock assessment. Best practice does require some attempt to scale the data (even if post stratification is not needed – which can only be shown by a careful formal analysis).

Trawl survey q
As already stated, it is unnecessary and poor practice to assume that a relative biomass time series (such as from a trawl survey) has a q exactly equal to 1. In the case of rex sole, it is not a bad assumption, but that was only determined after there was an analysis (at the review meeting) to show that *a priori* the q should not be that different from 1 (the mean of the prior being 1.2 with a relatively low standard deviation).

Estimation of recruitment deviations
For rex sole, it appeared that there were a number of early recruitment deviations that were being estimated despite not being observed in the age data. It is generally not a good idea to give the model lots of free recruitment deviations that are not supported by age data (e.g., they can be used to simply fit noise in biomass indices).

3. *Recommendations for improvements to the assessment model.*

The recommendations to improve the rex sole stock assessment are almost identical to those for the other two species and are presented jointly in the Recommendations section.

Gulf of Alaska Dover sole (Deepwater flatfish)

1. *Evaluation of the ability of the stock assessment model for GOA Dover sole, with the available data, to provide science advice to inform the management of Dover sole in the Gulf of Alaska.*

The Dover sole assessment was only transitioned to SS3 in 2015 and no more recent assessments have been produced. As such, there has been less investigation of the growth data and associated poor model fits. A two-sex, age-structured, single area model is fitted to the data with a single set of growth parameters. The trawl survey biomass indices are modified by a random effects model to fill in gaps in strata which were not surveyed in some years (the maximum depths surveyed vary from 500 m, to 700 m, and 1000 m and in 2001 the eastern strata were not surveyed). As in the other assessments q = 1 is assumed. No fishery age data are available, so the estimated fishery selectivity (especially given issues with growth) must be considered uncertain.

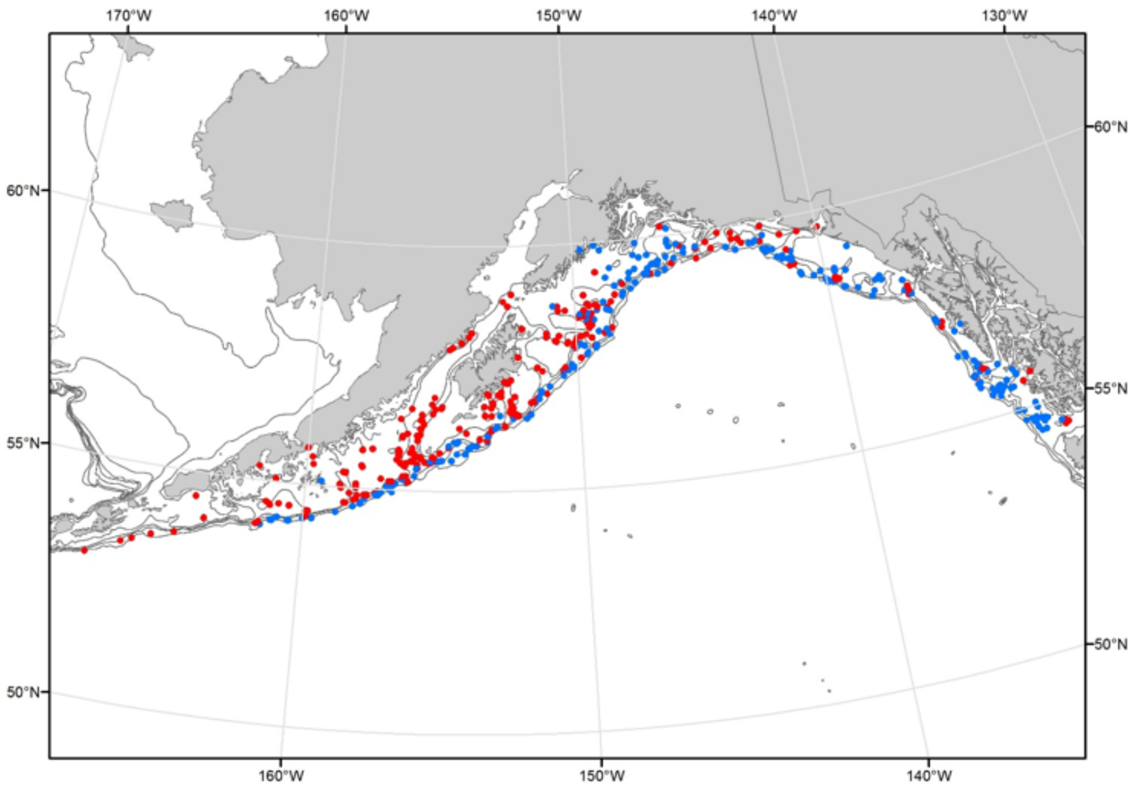

Figure 4: Dover sole residuals for von Bertalanffy growth curves fitted to age and length survey data outside of the model. Points are plotted at the location of the haul from which the fish were sampled. The blue points are more than 1 residual standard error below the curve, and the red points are more than 1 residual standard error above the curve. (From McGilliard 2019b, produced by Beth Matta).

As for rex sole, there is a spatial pattern for the growth residuals with the shorter fish at age showing a tendency to be in the east and a dominance in the deeper water (Figure 4). Since Dover move deeper as they age, the tendency to find shorter fish at age in the deep is probably just a reflection that the older cohorts are the short animals. This is largely confirmed by an examination of the length at age data by area and depth (Figure 5).

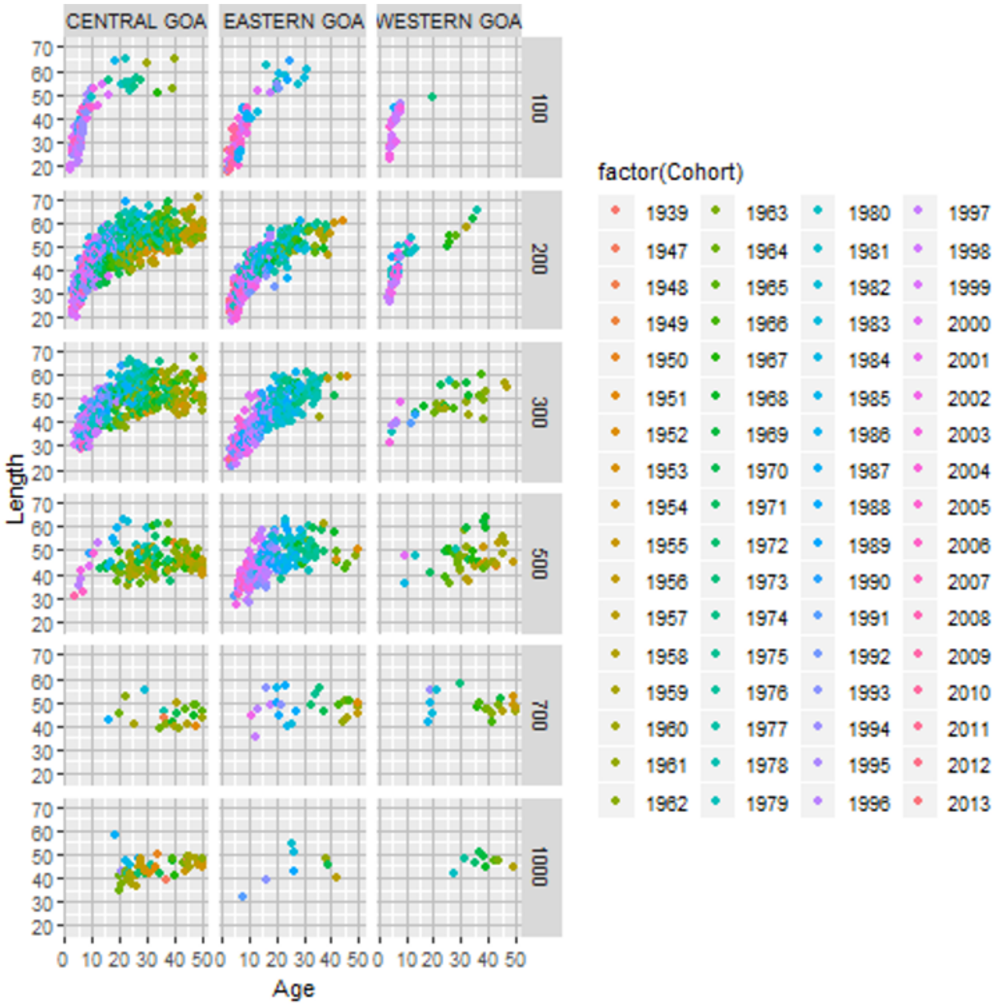

Figure 5: Female Dover sole length at age data by area (West, Central, East GOA), cohort, and depth. From McGilliard 2019b.

As for rex sole, the question of whether the assumption of q = 1 is adequate or not has to be answered by an evaluation of the available information. The experimental study on herding and selectivity has a mean and standard deviation for Dover that may be similar to rex (Sommerton et al. 2007). However, Dover have a much deeper distribution than rex, which is why the trawl survey biomass estimates had been modified to account for missing biomass in years when the survey had only extended to depths of 500 m or 700 m. But the question of how much biomass might be deeper than 1000 m had not been addressed.

At the review meeting the chair undertook a brief study to examine the proportion of Dover biomass by depth. He examined the survey CPUE by depth and also considered West Coast Dover sole survey biomass estimates by depth (the West Coast survey extends beyond 1000 m). He concluded, and the reviewers agreed, that there was very little biomass beyond 1000 m (though this should be checked with any historical commercial data that may be available). This conclusion meant that for survey years when the depth extended to 1000 m that the mean of the prior on the Dover survey q should be equal to 1.2 (as for rex). Using the survey biomass proportions by depth gave a mean for the q prior on surveys to 700 m of 1.17, and for surveys to 500 m of 1.05.

The reviewers specified a modified Dover sole model run: remove the 1980s survey data; remove the 2001 survey data (it only went to 500 m and did not survey the east); estimate three survey q's for surveys extending to 1000 m, 700 m, and 500 m with informed priors respectively of N(1.2, 0.175), N(1.17, 0.175), N(1.05, 0.175); allow the survey selectivities to be domed (on the right hand side as well as the left); and use Francis iterative reweighting.

The model run was not successful as the iterative reweighting did not converge (the procedure needs to be checked), the survey selectivities were not sensible, and consequently the estimated q's were not sensible. It was realised that there was no need to split the trawl time series into three components as 1.2 and 1.17 are almost the same. Therefore, the recommendation is to split the trawl surveys into two time series, those years that survey only to 500 m, and for the years that survey to 700 m or 1000 m to just use the estimates to 700 m (and use the N(1.17, 0.175) prior on that trawl q).

The fit of the 2015 base model to the trawl indices is not great, but it will be improved when the 1980s indices are removed (Figure 6). The scale of the biomass indices, combined with the information that q is unlikely to be very different from 1, gives an assurance that the spike in catches in 1990 and 1991 (of about 9000 t each year) is of no concern and that the stock is currently at high stock status and is lightly fished.

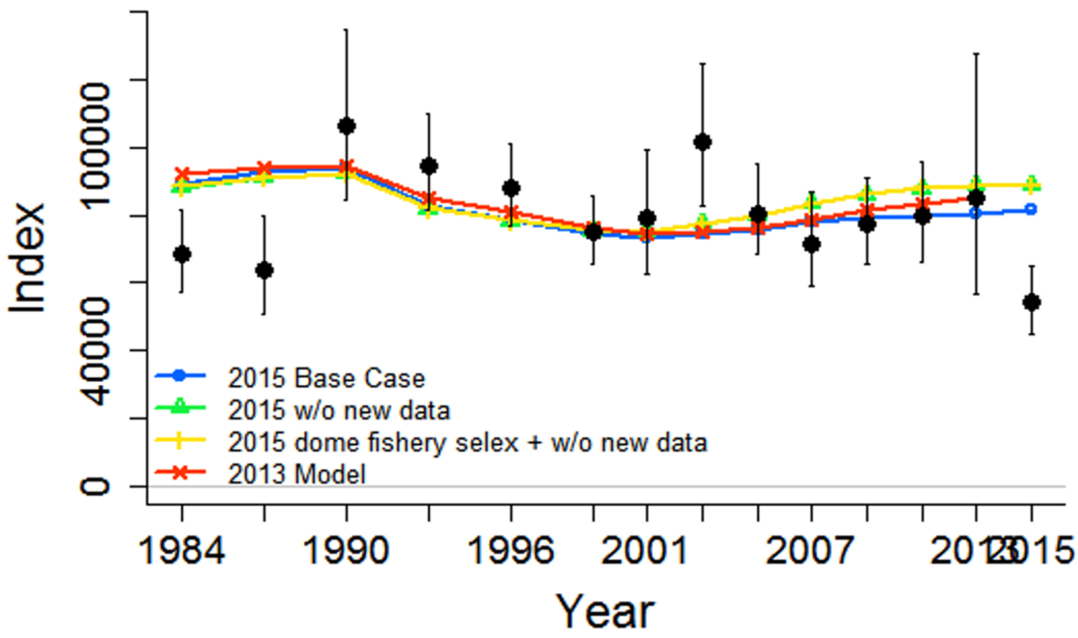

Figure 6: MPD fits to the GOA trawl survey biomass indices for Dover sole (with approximately 95% confidence intervals). Fits are shown for the 2015 base model, two 2015 sensitivities, and the 2013 base model. From McGilliard 2019b.

The base model assessment from 2015 estimates high stock status and low fishing mortality. These estimates are qualitatively reliable. Some improvements to the assessment model will change the exact estimates but will not change the main assessment conclusions.

2. Evaluation of the strengths and weaknesses in the stock assessment model for GOA Dover sole

The Dover sole stock assessment is the most problematic of the three assessments that were reviewed. It was only transitioned to SS3 in 2015, and there has been less time to investigate problems of poor fits that may be associated with spatial and temporal changes in growth.

The use of a two-sex, age structured model is appropriate given the available data and knowledge of the biology. The availability of a consistent time series of trawl surveys since 1990 is a particular strength (shared by all of the reviewed stock assessments).

The assessment does have some weaknesses, but they are not nearly substantial enough to compromise the reliability of the main assessment results.

Preparation of input data

An important issue, common to all three stocks, is that more care could have been taken in the preparation of the input data.

Common to all three assessments is the issue of the use of the 1980s surveys and the use of the 1990 and 1993 surveys. The 1980s surveys should not be used; the 1990 and 1993 surveys should probably be used in the base model, but a sensitivity should also be done that excludes them (see the discussion under rex sole).

Specific to Dover sole is the issue of the maximum depth covered by each survey. This issue was considered in the assessment, but the approach was to fill in the gaps due to different levels of survey coverage (using a random effects model). This is not the best approach – if a stratum wasn't surveyed in a particular year then, that should just be accepted. As a general principle, "making up data" by any means, no matter how sophisticated, is not good practice. The recommended approach is to split the surveys into two time series, those that went only to 500 m, and those that went to 700 m or 1000 m (and to obtain a consistent time series by only using the estimates to 700 m and accounting for the "lost" biomass in the informed prior on q).

As for the other stocks, care in the preparation of commercial composition data is required. At this stage, only length data are available from the fishery, but there needs to be an analysis of the drivers of length (using commercial and survey data) and consequently appropriate post stratification and scaling.

Trawl survey q

As already stated, it is unnecessary and poor practice to assume that a relative biomass time series (such as from a trawl survey) has a q exactly equal to 1. In the case of Dover sole it is not a bad assumption, but that was only determined after there was an analysis (at the review meeting) to show that *a priori* the q should not be that different from 1. The informed priors for the two trawl q's should be carefully developed, and the q's should be estimated as free parameters.

Estimation of recruitment deviations

For Dover sole it appeared that the recruitment deviations that were being estimated were supported by the age data. It would be a good idea to adopt a formal rule for which deviations to estimates based on; for example, observing a cohort at least three times in the age data (but not just at "very young" or "very old" ages).

3. Recommendations for improvements to the assessment model.

The recommendations to improve the Dover sole stock assessment are almost identical to those for the other two species and are presented jointly in the Recommendations section.

Gulf of Alaska flathead sole

1. Evaluation of the ability of the stock assessment model for GOA flathead sole, with the available data, to provide parameter estimates to assess the current status of flathead sole in the Gulf of Alaska

The 2017 base model for flathead sole is a two-sex, age-structured model, with a single set of growth parameters estimated (Turnock et al. 2017). Trawl survey biomass indices are fitted with q = 1 assumed. There is no issue with the variation in maximum depths for the trawl surveys as flathead has an inshore distribution. No fishery age data are available, so the estimated fishery selectivity should be considered uncertain.

A spatial analysis of residuals for growth data showed a spatial pattern with areas of predominately positive and negative residuals (Figure 7). The spatial pattern is not as obvious as it is for rex sole but there may be some benefit in further exploration of this pattern.

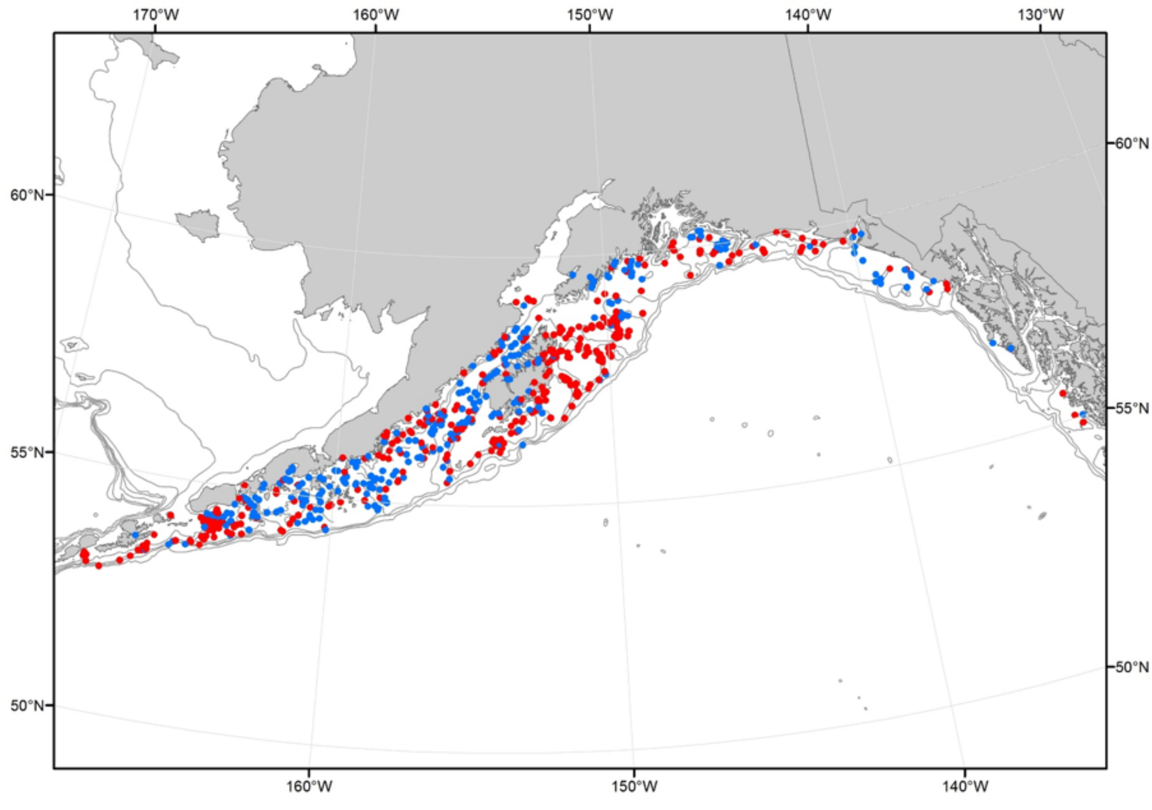

Figure 7: Flathead sole residuals for von Bertalanffy growth curves fitted to age and length survey data outside of the model. Points are plotted at the location of the haul from which the fish were sampled. The blue points are more than 1

residual standard error below the curve and the red points are more than 1 residual standard error above the curve. (From McGilliard and Turnock 2019, produced by Beth Matta).

The model is fitted to trawl survey biomass indices assuming that q = 1. This is a poor assumption in general, but for this stock and this survey it is adequate. From the inshore distribution of flathead sole it is apparent that a maximum survey depth of 500 m (which was achieved in every year of the survey) is adequate to cover the spatial distribution of the population. Also, an experiment with the trawl survey gear supported a vulnerability (for fully selected fish in front of the net) of about 1.2 (Sommerton et al. 2007). The mean value for an informed prior on the trawl survey q should therefore be fixed at about 1.2 (and the s.d. from the experiment was about 0.175).

The reviewers requested a modified model run: remove the 1980s trawl survey data; estimate a single trawl survey q with an informed prior (mean=1.2, s.d.=0.175); and Francis iterative reweighting for the composition data (Francis 2011, Punt 2017).

The model run requested by the reviewers (where q was estimated with an informed prior) gave similar results to the base model with an estimated q of approximately 1.3, which is not too different from the assumed value of 1 (Figure 8).

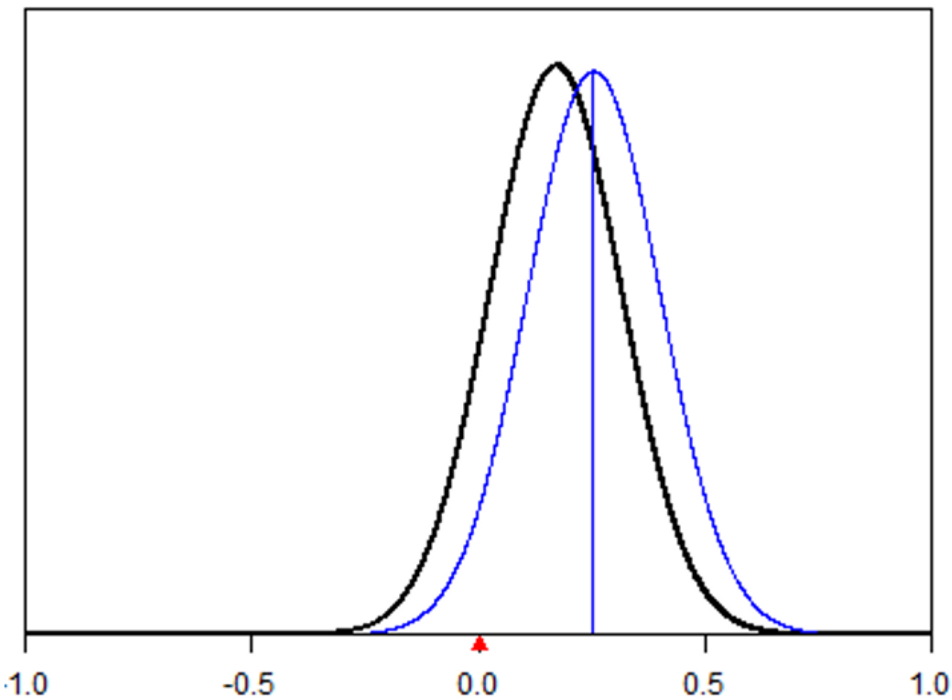

Figure 8: For the run requested by the reviewers: the prior distribution for the flathead trawl survey q (black) and the estimated MPD estimate (vertical blue line) and associated approximate posterior distribution (blue). The x axis is the natural log of q (so 0.0 corresponds to a q of 1). From McGilliard and Turnock 2019.

The reviewers' run gave a lower estimated spawning biomass (associated with the higher q), higher fishing mortality, and had a similar fit to the trawl indices (Figure 9). The estimated recruitment deviations are similar (but the early recruitment deviations, not observed in the age data, were not estimated) (Figure 9). The estimated level of uncertainty is much larger than the 2017 base model (Figure 9). The scale of the stock as estimated in the model is robust because the trawl survey q is not too much higher than 1 (as estimated in the reviewers' model and as assumed in the base model).

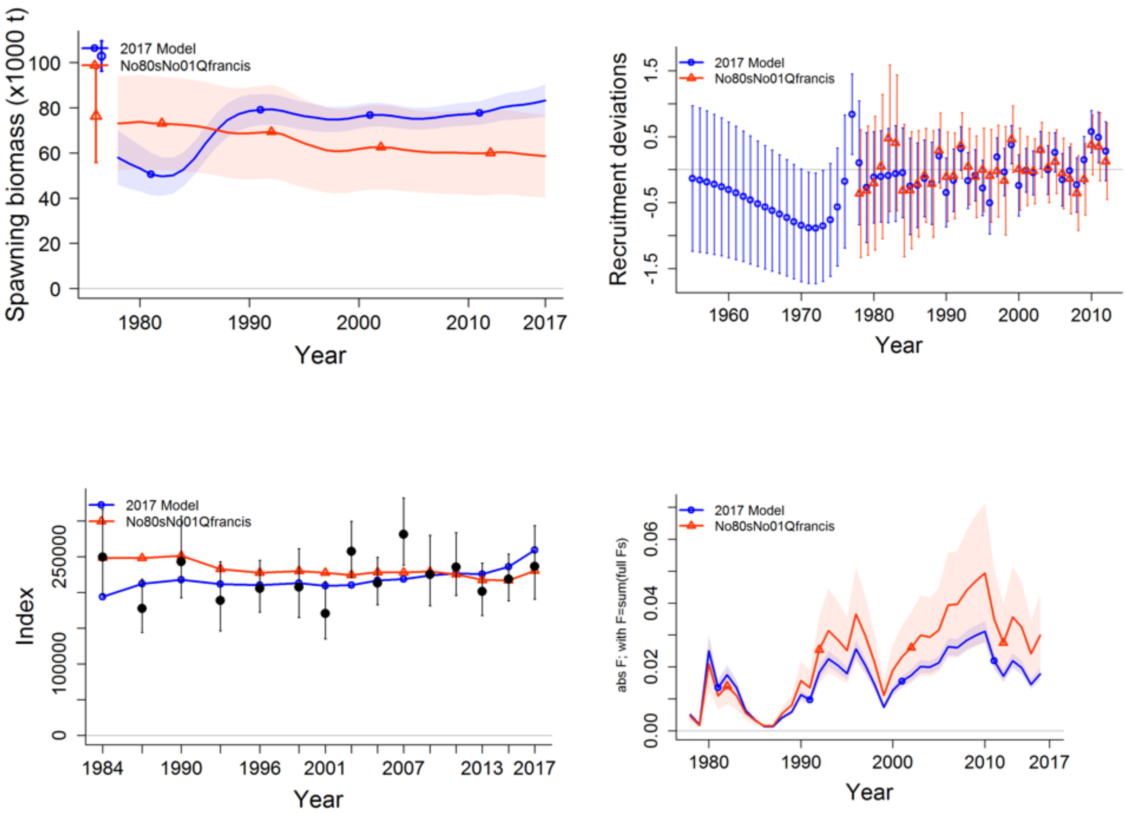

Figure 9: A comparison of estimates and results for the flathead base model (2017) and the reviewer's run (No80sNo01QFrancis): spawning stock biomass (with approximate 95% credibility intervals), recruitment deviations, fits to the trawl survey indices, and fishing mortality. From McGilliard and Turnock 2019.

The base model assessment from 2017 estimates high stock status and low fishing mortality. These estimates are qualitatively reliable. Some improvements to the assessment model will change the exact estimates but will not change the main assessment conclusions.

2. *Evaluation of the strengths and weaknesses in the stock assessment model for GOA flathead sole.*

The use of a two-sex, age-structured model is appropriate given the available data and knowledge of the biology. The availability of a consistent time series of trawl surveys since 1990 is a particular strength (shared by all of the reviewed stock assessments).

The assessment does have some weaknesses, but they are not nearly substantial enough to compromise the reliability of the main assessment results.

Preparation of input data

An important issue, common to all three stocks, is that more care could have been taken in the preparation of the input data.

Common to all three assessments is the issue of the use of the 1980s surveys and the use of the 1990 and 1993 surveys. The 1980s surveys should not be used; the 1990 and 1993 surveys should probably be used in the base model, but a sensitivity should also be done which excludes them (see the discussion under rex sole).

As for the other stocks, care in the preparation of commercial composition data is required. At this stage, only length data are available from the fishery, but there needs to be an analysis of the drivers of length (using commercial and survey data) and consequently appropriate post stratification and scaling.

Trawl survey q

As already stated, it is unnecessary and poor practice to assume that a relative biomass time series (such as from a trawl survey) has a q exactly equal to 1. In the case of flathead sole, it is not a bad assumption, but that was only determined after there was an analysis (at the review meeting) to show that *a priori* the q should not be that different from 1. The informed priors for the trawl q should be carefully developed and the q should be estimated as a free parameter.

Estimation of recruitment deviations

For flathead sole it appeared that early recruitment deviations were being estimated despite not being observed in the age data. It would be a good idea to adopt a formal rule for which recruitment deviations to estimate based on; for example, observing a cohort at least three times in the age data (but not just at "very young" or "very old" ages).

3. *Recommendations for improvements to the assessment model.*

The recommendations to improve the flathead sole stock assessment are almost identical to those for the other two species and are presented jointly in the Recommendations section.

Recommendations

This is a set of recommendations for each of the stock assessments, which mirror the identified weaknesses. There is an additional recommendation for the observer program.

- Trawl survey biomass indices need to be consistent across time:
    - Exclude the trawl survey indices in the 1980s, as they used very different vessels, gear, and station occupation. The composition data should also be excluded or fitted with its own selectivity.
    - Carefully consider the use of the 1990 and 1993 trawl survey indices in the base model. For some species it may be considered that the later timing of these surveys means that they are not comparable to the surveys from 1996 onwards.
- Commercial composition data need to be appropriately post stratified and scaled.
- Trawl survey q's should be estimated as free parameters with informed priors.
- Recruitment deviations should only be estimated for cohorts which are observed in the age data being fitted.
- Investigate the possible use of age- *and* length-based selectivity curves (as available in SS3).
- Consider estimating growth outside the model and focusing the stock assessment on fitting biomass indices and age frequencies rather than vast quantities of conditional age at length data. The simplicity of the approach will, in most cases, outweigh the minor biases in the growth curves.

- The observer program is excellent, but has a potential flaw for bycatch species as these may rarely be the predominant species in an individual trawl catch (and thus not be sampled). This could result in a lack of data for some stocks and unrepresentative sampling for other stocks (e.g., if most of the annual catch is from small catches that are not sampled, but the species is predominant and sampled when schooling juveniles are caught).
    - Conduct a study to determine which (if any) bycatch species/stocks are not being adequately sampled by the observer program.

Conclusions

The three stock assessments reviewed shared similar strengths and weaknesses. The availability of a consistent time series of trawl survey data since the 1990s is a strength. For rex sole, the availability of age data from the commercial fishery allows more reliable estimation of the commercial selectivity than for the other two stocks where only length data are available.

The assessments are all qualitatively reliable in the assessment of high current stock status and low fishing mortality. The assessments can be improved with more careful data preparation and the estimation of the trawl survey q's with informed priors.

The AFSC review process differs from those in other jurisdictions in that there is no pressure to review a stock assessment that will shortly be used to provide management advice. The absence of that pressure and tension encourages a more relaxed and collegial atmosphere than in the other processes (e.g., STAR panels, SEDARs, SARCs).

References

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

Punt, A.E. 2017. Some insights into data weighting in integrated stock assessments. *Fisheries Research 192: 52-65.*

Somerton, D.A.; Munro, P.T.; Weingerg, K.L. 2007. Whole-gear efficiency of a benthic survey trawl for flatfish. *Fish. Bull. 105*: 278-291.

Appendix 1: Bibliography of materials provided for review

This appendix lists the main presentations and documents that were provided for review. A large number of background documents and extracts were also made available.

Presentations

Faunce, C.H.; Cahalan, J. 2019. North Pacific observer sampling design. Gulf of Alaska rex, Dover, and flathead sole. 29 April 2019. 28 slides.

McGilliard, C. 2019a. CIE review for GOA rex, Dover, and flathead sole. April 2019, 104 slides (revised).

McGilliard, C. 2019b. GOA Dover sole. April 2019, 85 slides (revised).

McGilliard, C.; Turnock, J. 2019. GOA flathead. 72 slides (revised).

Matta, B. 2019. Age determination of Gulf of Alaska flatfish at the AFSC. 29 April 2019. 44 slides.

Palsson, W. 2019. Gulf of Alaska bottom trawl survey. 56 slides.

Documents

McGilliard, C.R. and Palsson, W. 2015a. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 563-624. NPFMC.

McGilliard, C.R. and Palsson, W. 2015b. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 751-808. NPFMC.

McGilliard, C.R. and Palsson, W. 2017. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 657-742. NPFMC.

McGilliard, C.R., Palsson, W., and Stockhausen, W. 2015. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 625-674. NPFMC.

McGilliard, C.R., Palsson, W., Stockhausen, W., and Ianelli, J. 2013. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 403-536. NPFMC.

McGilliard, C.R., Palsson, W., Stockhausen, W., and Ianelli, J. 2013. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 612-756. NPFMC.

Stockhausen, W., Wilkins, M.E., Martin, M.H. 2011. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 547-628. NPFMC.

Stockhausen, W., Wilkins, M.E., Martin, M.H. 2011. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 629-690. NPFMC.

Stockhausen, W., Wilkins, M.E., and Martin, M.H. 2011. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 753-820. NPFMC.

Turnock, B.J., McGilliard, C.R. and Palsson, W., J. 2017. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 841-912. NPFMC.

Appendix 2: Performance Work Statement

National Oceanic and Atmospheric Administration (NOAA)

National Marine Fisheries Service (NMFS)

Center for Independent Experts (CIE) Program

External Independent Peer Review

Gulf of Alaska flatfish - Dover sole, rex sole, and flathead sole

Background
The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.
Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).

Further information on the CIE program may be obtained from www.ciereviews.org.

Scope
The stock assessments for Gulf of Alaska Dover sole, rex sole, and flathead sole provide the scientific basis for the management advice considered and implemented by the North Pacific Fisheries Management Council. An independent review of these integrated stock assessments is requested by the Alaska Fisheries Science Center's (AFSC) Resource Ecology and Fisheries Management Division (REFM). The goal of this review will be to ensure that the stock assessments represent the best available science to date and that any deficiencies are identified and addressed. The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (TORs) of the peer review are listed in **Annex 2**. Lastly, the tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements
NMFS requires three (3) reviewers to conduct an impartial and independent peer review in accordance with the PWS, OMB guidelines, and the TORs below. The reviewers shall have a working

knowledge and recent experience in the application of stock assessment methods in general and in Stock Synthesis in particular.

Tasks for Reviewers

1) Review the following background materials and reports prior to the review meeting:

Gulf of Alaska Flathead Sole

Turnock, B.J., McGilliard, C.R. and Palsson, W., J. 2017. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 841-912. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2017/GOAflathead.pdf

McGilliard, C.R. and Palsson, W., J. 2015. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 751-808. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2015/GOAflathead.pdf

McGilliard, C.R., Palsson, W., Stockhausen, W., and Ianelli, J. 2013. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 612-756. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2013/GOAflathead.pdf

Stockhausen, W., Wilkins, M.E., and Martin, M.H. 2011. Assessment of the Flathead Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 753-820. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2011/GOAflathead.pdf

Gulf of Alaska Rex Sole

McGilliard, C.R. and Palsson, W., J. 2017. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 657-742. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2017/GOArex.pdf

McGilliard, C.R., Palsson, W., and Stockhausen, W. 2015. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 625-674. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2015/GOArex.pdf

Stockhausen, W., Wilkins, M.E., Martin, M.H. 2011. Assessment of the Rex Sole Stock in the Gulf of Alaska. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 629-690. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2011/GOArex.pdf

Gulf of Alaska Dover Sole (Deepwater flatfish)

McGilliard, C.R. and Palsson, W. 2015. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 563-624. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2015/GOAdeepflat.pdf

McGilliard, C.R., Palsson, W., Stockhausen, W., and Ianelli, J. 2013. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 403-536. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2013/GOAdeepflat.pdf

Stockhausen, W., Wilkins, M.E., Martin, M.H. 2011. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 547-628. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510. https://www.afsc.noaa.gov/REFM/Docs/2011/GOAdeepflat.pdf

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

3) After the review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.

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

5) Deliver their reports to the Government according to the specified milestones dates.

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

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

Period of Performance
The period of performance shall be from the time of award through June 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
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
April 29 - May 3, 2019	Panel review meeting
May 17, 2019	Contractor receives draft reports
May 31, 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; (2) The reports shall address each TOR as specified; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $7,000.

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

Project Contact(s):

Carey McGilliard
Resource Ecology & Fisheries Management Division
NMFS| Alaska Fisheries Science Center
7600 Sand Point Way NE, Bldg. 4, Seattle, WA 98115-6349
Phone: 206-526-4696
carey.mcgilliard@noaa.gov

Annex 1: Peer Review Report Requirements

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

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

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

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

c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

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

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

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review

Appendix 2: A copy of this Performance Work Statement

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

Annex 2: Terms of Reference for the Peer Review

Gulf of Alaska Rex Sole

1. Evaluation of the ability of the stock assessment model for GOA rex sole, with the available data, to provide parameter estimates to assess the current status of rex sole in the Gulf of Alaska
2. Evaluation of the strengths and weaknesses in the stock assessment model for GOA rex sole
3. Recommendations for improvements to the assessment model.

Gulf of Alaska Dover Sole (Deepwater flatfish)

1. Evaluation of the ability of the stock assessment model for GOA Dover sole, with the available data, provide science advice to inform the management of Dover sole in the Gulf of Alaska
2. Evaluation of the strengths and weaknesses in the stock assessment model for GOA Dover sole
3. Recommendations for improvements to the assessment model.

Gulf of Alaska Flathead Sole

1. Evaluation of the ability of the stock assessment model for GOA flathead sole, with the available data, to provide parameter estimates to assess the current status of flathead sole in the Gulf of Alaska.
2. Evaluation of the strengths and weaknesses in the stock assessment model for GOA flathead sole.
3. Recommendations for improvements to the assessment model.

Annex 3: Tentative Agenda

Gulf of Alaska flatfish - Dover sole, rex sole, and flathead sole

TBD

Alaska Fisheries Science Center

7600 Sand Point Way NE

Seattle, WA 98115

April 29 - May 3, 2019

Point of contact Carey McGilliard (carey.mcgilliard@noaa.gov)

Appendix 3: Review meeting participants

Chair
Jim Ianelli, NMFS

Primary stock assessment author
Carey McGilliard, NMFS

Other presenters (NMFS)
Wayne Palsson (GOA trawl survey)
Craig Faunce (Observer program)
Jennifer Calahan (Observer program)
Beth Matta (Ageing and growth)

CIE reviewers
Patrick Cordue
Geoff Tingley
Kurt Trzcinski

Appendix 4: Summary report including recommendations agreed to by all reviewers

Summary Report

CIE Review of assessments for Gulf of Alaska rex, Dover, and flathead soles

April 29-May 3, 2019

Alaska Fisheries Science Center, Building 4, Room 2039, Seattle

Patrick Cordue, Center for Independent Experts (CIE)
Geoff Tingley, Center for Independent Experts (CIE)
Kurt Trzcinski, Center for Independent Experts (CIE)

Participants

Jim Ianelli, NMFS, chair
Carey McGilliard, NMFS, stock assessment scientist
Wayne Palsson, NMFS
Craig Faunce, NMFS
Jennifer Calahan, NMFS
Beth Matta, NMFS

Summary

A CIE review of three stock assessments of Gulf of Alaska (GOA) flatfish stocks was conducted at the Alaska Fisheries Science Center from April 29 to May 03 2019. The participants included three CIE reviewers, the primary assessment author, the chair of the meeting, and NMFS staff who presented on relevant topics.

On the first day, an introductory presentation was given on the GOA ecosystem and flatfish fisheries. Presentations on the GOA trawl survey, the observer program, and the ageing of flatfish were also given. Stock assessment presentations for the three species were given over the following days.

The stock assessments were primarily conducted by the same author who transitioned the assessments from purpose written code to Stock Synthesis 3 (SS3) in 2013 (Dover and flathead) and 2015 (rex). Subsequent assessments have primarily been refinements of the models developed in 2013 and 2015.

The assessment models and the use of data in the assessments were similar across the three assessments. Therefore, the assessments broadly shared the same strengths and weaknesses. In general, the age-structured models were appropriate given the available biological, abundance, and composition data. A particular strength of the assessments is the availability of a consistent time series of biomass estimates from the GOA trawl surveys (in particular since 1996).

The preparation of the input data can be improved in some respects. More exploratory and formal analysis of the composition data is required so that length, age, and age-at-length data can be appropriately post-stratified (if necessary) and scaled. The trawl biomass time series also needs to be treated carefully, especially for species which have a distribution below 500 m (the maximum depth of the survey in some years).

The assumption that the trawl survey biomass indices are estimates of absolute biomass (q = 1) is inappropriate for most stock assessments. It is better to estimate the “catchability” (q) and support the estimation with an informed prior (which contains the currently available information on the value of q). A first attempt at producing an informed q prior for each stock was performed during the meeting and model runs were performed with the informed priors. Although the point estimates of spawning biomass and stock status were similar to the original models the results reflected a greater and much more appropriate level of uncertainty.

The reviewers appreciated the excellent presentations by the NMFS staff, the hard work of the assessment author, and the collegial and constructive atmosphere under which the review meeting was conducted.

Main Recommendations

These recommendations address common issues found in each of the three assessments reviewed, and that may also be relevant for other assessments. These were agreed by the three CIE reviewers.

Gulf of Alaska Bottom Trawl Survey (BTS)

1. The surveys conducted in 1984 and 1987 used different vessels, a different approach and with different timing. These surveys should not be considered as part of the same timeseries as the subsequent BTS timeseries. Specifically, the biomass estimates and the composition data from these two surveys should be dropped from each of these assessments, and probably from all other assessments also.
2. The surveys in 1990 and 1993 had a different timing (later) and somewhat different survey structure. While clearly not as ‘different’ as the 1984 and 1987 surveys, there is sufficient difference that model sensitivities should be run on a species-by-species (stock-by-stock) basis that include and exclude the biomass and composition data from these two surveys.
3. Where there are gaps in survey data due to, for example, not surveying some areas in some years, these should be left as data gaps. The model structures used are more than capable of dealing with such data gaps. Data should not be created by extrapolation, interpolation or modelling to fill such gaps.

Fishery sampling

4. A more consistent, analytical and defensible approach to the scaling and stratification of fisheries data should be followed. This should meet accepted ‘best practice’ approaches, including, for example, studying the spatial and temporal patterns of length and age followed by appropriate stratification and scaling.

Modelling

5. Models should not assume that the survey q is equal to 1. Informed priors should be developed on a stock-by-stock basis.
6. Recruitment deviates should not be estimated where there is no information to inform the estimation, i.e. there has to be age data from a survey or fishery to inform the estimation process.

Observer data to support the stock assessments

7. The Observer Program delivers information to support stock assessments for a large number of groundfish stocks. On the whole, this works very well but is not the case for all stocks. With respect to this review, age data for Dover sole from the fishery are, due to the scale of the fishery and the sampling prioritisation approach of the Observer Program, insufficient to provide any recent age frequencies for use in the assessment. In addition, for some bycatch species there will be a real prospect of sampling being unrepresentative. The development of alternative Observer Program sampling strategies for low catch and bycatch fisheries to provide the required data to support the assessments should be conducted as a matter of priority.