Independent peer review report --- Bering Sea and Aleutian Island Flatfish (yellowfin sole, rock sole and Alaska plaice) Stock Assessments

Yan Jiao

Prepared for

Center for Independent Experts

June 2018

Contents

Executive Summary

The 2018 stock assessments of three Bering Sea/Aleutian Islands (BSAI) flatfish species – yellowfin sole, northern rock sole and Alaska plaice – were reviewed by a Center for Independent Experts (CIE) review panel. The review sought to evaluate the progress of the stock assessments and to ensure that the North Pacific Fishery Management Council (NPFMC) bases its decisions on the best information available when managing these three species. The CIE Review took place at the Alaska Fisheries Science Center (AFSC), Seattle, Washington during April 16 - 18, 2018. The assessments of the stocks done by the stock assessment team were presented to the CIE review panel and the validity of the data, assessment procedures, and results were discussed. The AFSC provided all the necessary logistic support, background information, documents and further data and model configuration explorations that were requested by the review panel.

All the three stock assessments have been used for years and there are no new changes in their model structures, but data are updated to 2016 or 2017. All the models are coded in ADMB, and both likelihood and Bayesian approaches were used to estimate parameters. The 2018 peer review panel found that the stock assessments addressed all the terms of references (TORs) adequately.

There were no major weaknesses in the assessments, but improvements could be made with regards to exploring reasons or patterns of selectivity changes over time, and better incorporating uncertainty from parameters of natural mortality and catchability.

It is more desirable to compile the reports based on a model selection framework and to incorporate model selection uncertainty in the future. Measurements in selecting models can be model goodness-of-fit, model prediction ability, model robustness and fisheries-specific measurements, such as retrospective error (Jiao et al. 2012; Gelman et al. 2014; Hooten and Hobbs 2015).

All the three assessments are considered to represent the best scientific information available for the three species, and are suitable as bases for management advice. All the three stocks are found not overfished and not experiencing overfishing.

1. BACKGROUND

This report reviews the 2018 stock assessments of three flatfish species distributed in BSAI at the request of the Center for Independent Experts (CIE). The three species that underwent review are BSAI yellowfin sole, northern rock sole and Alaska plaice. The assessment team includes Drs.Thomas Wilderbuer, Daniel Nichol, and James Ianelli. The CIE panel includes Drs. Joseph Powers, Matthew Cieri, and Yan Jiao. According to the CIE Statement of Work, "Having these assessments vetted by an independent expert review panel is a valuable part of the AFSC's review process." As a review panel member, I was provided with draft stock assessment reports and web access to relevant files and documents (Appendix 1) and participated in the Stock Assessment Review Meeting. Extra documents and model runs were provided during the review upon request from the CIE peer review panel (Appendices 1 and 4).

Prior to the presentations on the stock assessments of the three species under review, an overview of the fisheries in BSAI and their management processes and management tiers were introduced by Dr. Wilderbuer. Presentations on the bottom survey trawl conducted in BSAI, the observer program, age determination, and the fisheries management and council processes were provided also (Appendix 1). These surveys, observer programs and biological data validation provided crucial data used in the stock assessments of these three flatfish stocks.

For the three species/stocks under review, the current stock assessment model used, alternative model runs explored, the biological references points estimated, and the recommended Tier for management were presented by the stock assessment team (see Agenda in Appendix 2). Discussions on the quality of the data, the appropriateness of the model equations and error structures, key parameters, and the estimation algorithms were made throughout the review.

The quality of the data for each species was discussed and comments and suggestions are listed under TORs. During the review meeting, the stock assessment team was always available when required for further discussion, additional data exploration and clarification, and clarification of how the Alaska flatfish TORs were addressed. Recommendations from NPFMC SSC reports and minutes of Plan Teams of GOA and BSAI groundfish were both reviewed to determine the extent to which they had been addressed.

As a CIE reviewer, my duty was to evaluate the stock assessments of the 3 species with respect to the TORs, which are attached in Annex 2 of the Appendix 2. This report provided the findings of the independent review that is undertaken by me in accordance with the CIE Statement of Work (SOW).

2. ROLE OF INDIVIDUAL REVIEWER IN THE REVIEW ACTIVITIES

My role as a CIE independent reviewer was to conduct an impartial and independent peer review in accordance with the SOW and the predefined TORs herein.

About two weeks before the meeting, assessment documents and supporting materials were made available to the review panel via email and a http website by Dr. Wilderbuer. I read all the documents that I received prior to the review.

The CIE peer review meeting was held at the AFSC, Seattle, Washington, from April 16-18, 2018. The meeting followed the "tentative agenda (Appendix 2)" of the CIE review. The meeting was open and was organized constructively. On the morning of April 16 before the meeting, the assessment review committee met with Dr. Wilderbuer and the stock assessment team to discuss the meeting agenda, reporting requirements, and meeting logistics. During the meeting, all documents were accessible online.

Presentations were given during the review according to the agenda to provide the CIE reviewers the background information on the data used in the stock assessment models, the model development, biological reference points, and the management tiers. I was actively involved in the discussion during the presentations by 1) listening to the presentations carefully, making notes on the points that are not included or not clearly stated in the documents provided prior to the meeting; 2) asking questions for clarification on the data usage and model development; 3) making comments and providing possible alternative solutions to questions arising during the meeting.

After the peer review meeting, I summarized the findings and recommendations according to the predefined TORs. This review report is formatted according to my interpretation of the required format and content described in Annex 1 of Appendix 2.

3. SUMMARY OF FINDINGS AND RECOMMENDATIONS FOR TORs

Below I provide the summary of findings of each TOR for each species in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

3.1 BSAI yellowfin sole TORs

3.1.1 *Evaluate the strengths and weaknesses of the assumptions made in applying the stock assessment model including how survey indices are scaled to the populations. Specifics might include:*

a. How natural mortality estimates are estimated/applied

Natural mortality (*M*) is assumed to be 0.12/year for both sexes in the assessment team recommended base model. The value of 0.12/year was from multiple sources of evidence, which include estimates of *M* by fitting to catch-at-age with Japanese pair-trawl data, by using life history of maximum age of yellowfin sole, and by the profiling over a range of *M* with fixed *q*.

The assessment team also explored alternative runs by estimating M of one sex while fixing *M* at 0.12/year for the other sex, but the assessment team felt that since the population sex ratio annually observed at the time of survey is a function of the timing of the annual spawning in adjacent inshore areas, it is questionable that providing the best fit to these observations is really fitting the population sex ratio better. So, that the assessment team recommended the use of 0.12/year for both sexes. The review team eventually agreed with the assessment team but suggested that future incorporation of uncertainty of M in the assessment can be explored.

b. Assumptions about survey "catchability"

The catchability (*q*) is modelled as a function of annual water temperature anomaly (T) at survey stations less than 100m, i.e., $q = e^{-\alpha+\beta T}$. The relationship was first examined in Wilderbuer and Nichol (2001). The results of incorporating T to estimate *q* has resulted in an improved fit to the survey biomass. The results match the observations from the bottom trawl surveys (Nichol 1998; Wilderbuer et al. 2017).

A recent research found that *q* changes when wave height and wave-induced vessel motion changes through field experiments in September of 2003 (Somerton et al. 2018). A strong temperature effect on the survey catchability has been found in the stock assessment model (Wilderbuer and Nichol 2001). The assessment team addressed the possibility that trawl sampling efficiency is more influenced by waves, but yellowfin sole availability to the survey is more influenced by temperature. This can further be interpreted as the inter- and intra- annual variations in survey catchability. The signal detected through the annual water temperature anomaly from the assessment model reflects inter-annual variation while the wave caused variation reflects intra-annual variation in catchability.

c. Application of fishery and survey age-specific schedules (maturity, body mass, selectivity)

Two studies on maturity-at-age (Nichol 1995; TenBrink and Wilderburer 2015) were conducted historically, and information from the two studies was pulled together and used in this assessment. The differences in the estimated maturity were very limited, and their influence in the estimated spawning stock size were limited also (2%).

Weight-at-age is routinely updated, both for the fishery and the survey. I support the selection of using observed weight-at-age data directly, but for age groups in some years with small sample size, smoothing or a model-based approach may be used to avoid measurement uncertainty caused by small sample size or samples from limited spatial locations.

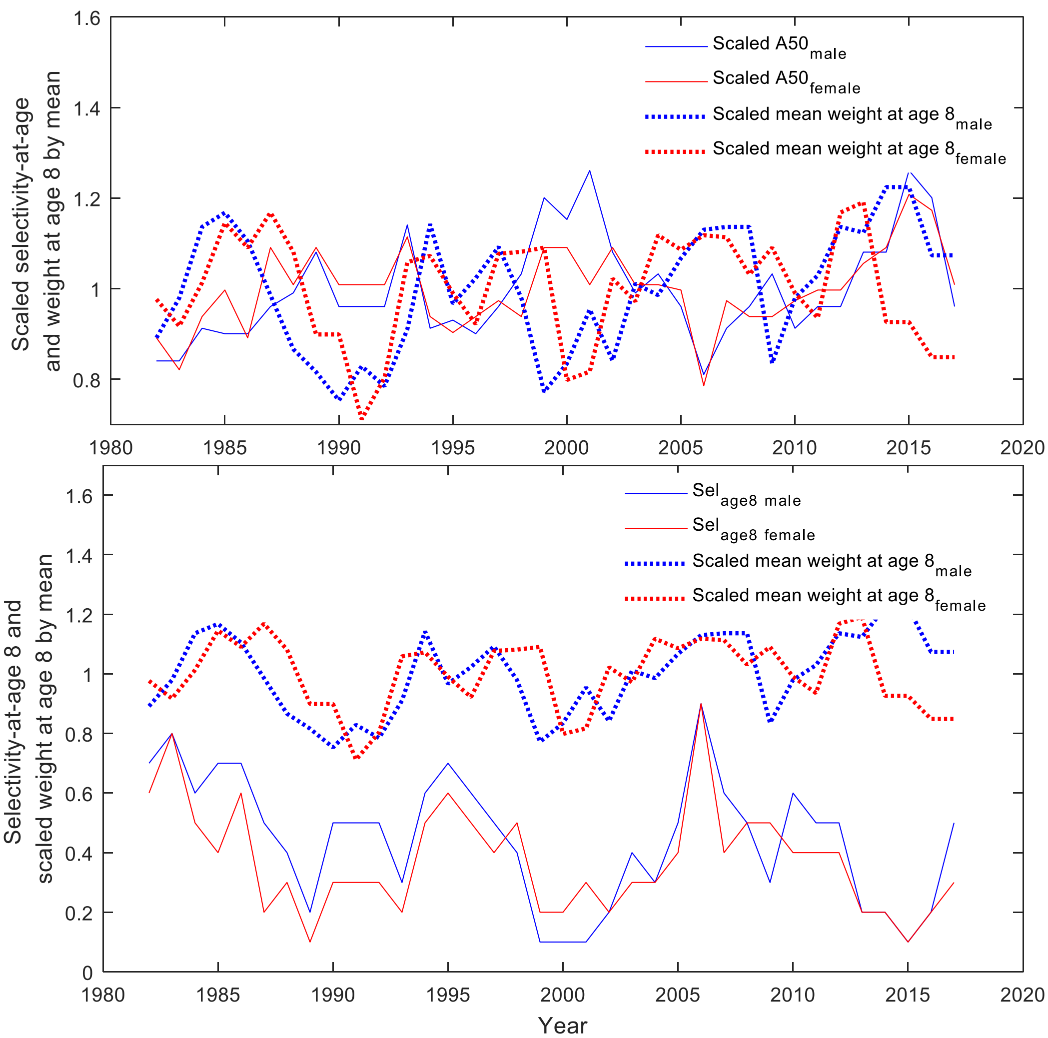

Figure 1: Estimated age of selectivity at 50%, selectivity of age 8 and mean weight at age 8 scaled by the average of the values between 1982 and 2017 (Values are digitized from Tables 4.9 and 4.16).

Selectivity for both fishery and survey is modelled as a logistic function of age with survey selectivity not changing over time, but fishery selectivity varying every year. Further exploration on the reasons or patterns of the changes in selectivity should help 1) understanding the fish/fishing behavior; and 2) better estimation of both BRPs and future stock projections. Whether there is a consistent change in both selectivity pattern and growth pattern is worth an exploration in the future. A preliminary visual diagnostic found that age for selectivity at 50% increases and selectivity at age 8 decreases when yellowfin sole grow smaller (Figure 1). Comparison between age versus size-based selectivity seems worthwhile and exploration on whether changes in selectivity-at-age is because of growth variation is necessary in the future.

d. The application (or lack thereof) of a stock-recruitment relationship (and associated parameter estimates)

A Ricker model with a two separated period approach was used in the base assessment, and the stock-recruitment relationship from the period of 1978-2012 was used to derive reference point and population projections. Yellowfin sole demonstrated productivity changes over time, the strategy used by the assessment team is found reasonable.

3.1.2 *Evaluate the stock assessment approach used focusing specifically on how fisheries and survey data are compiled and used to assess the stock status relative to stated management objectives under the Bering Sea and Aleutian Islands Fishery Management Plan (FMP) and the Magnuson-Stevens Act requirements. Elements should consider:*

a. The FMP "Tier" designation

The yellowfin sole stock assessment provided reliable estimates of B_{MSY} and probability density function of F_{MSY}, and the projected spawning biomass is higher than B_{MSY}, so that the FMP Tier designation for yellowfin sole is Tier 1a. The review panel agrees with the assessment team and SSC.

b. Fishing rate estimation relative to overfishing definitions

Estimated full selection fishing mortality is 0.08-0.09/year in recent 5 years, much smaller than F_{MSY} which is estimated to be 0.12/year; the reported catch is lower than ABC and OFL (about 40-50%), so overfishing is not occurring. The review panel feels that the conclusion is reasonable and robust to alternative model configurations.

c. Stock status determinations relative to B_{MSY}

The estimated female spawning stock biomass is much higher than B_{MSY} for yellowfin sole, so the stock is not overfished. The review panel feels that the conclusion is reasonable and robust.

3.1.3 *Recommend how assessment data and/or models could be improved*

The survey spawning biomass is used to calibrate both the trend and scale of the population size. Catchability plays a heavy role in the estimability or the stability of the model estimated population size of this species. An extra set of model explorations by estimating *M* for one sex and estimating *q* as a function of T demonstrated strong retrospective pattern (Figure 2). This indicted that the scale of the population size will largely benefit from any extra information that should infer either *q* or *N*. A retrospective analysis of *q* will also help us to understand the stability of the scale of the estimated *q* over time. An informative prior of the scale of *q* from field experiment adjusted by T each year, i.e., central tendency calibrated by informative priors but variation among years adjusted by T, should help if scale of the *q* is not stable.

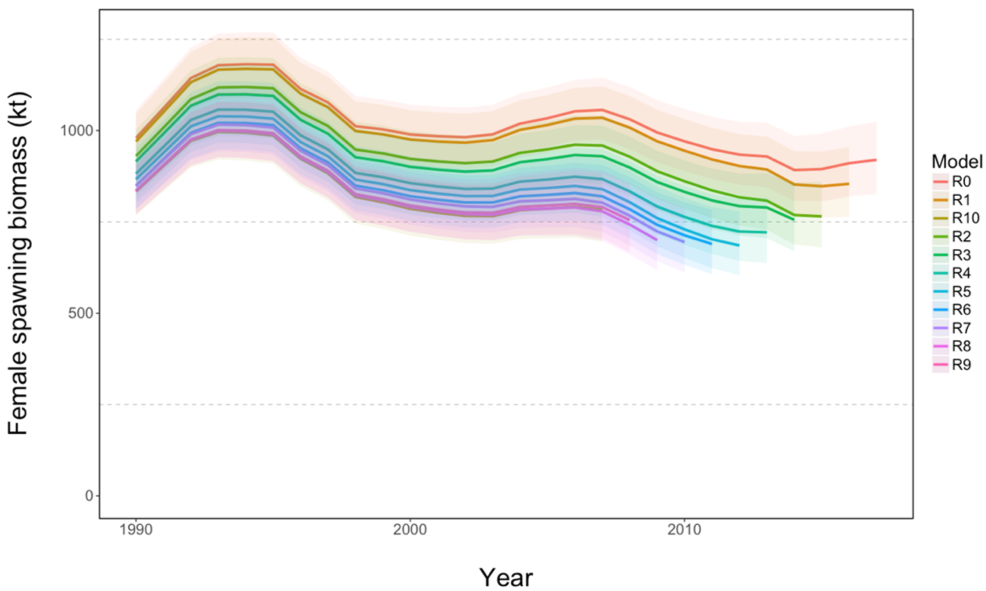

Figure 2: Retrospective analysis when model configuration is with fix M=0.12 but estimating *q* ~ *f*(*T*). (Figure is from the responses to the review request)

I support the research effort to explore changes of selectivity over time. Also, how different selectivity configurations improve model goodness-of-fit and influence stock assessment results (in the estimated population size and fishing mortality) needs to be quantified in the stock assessment document. Comparison between age versus size-based selectivity seems worthwhile and exploration on whether selectivity-at-age variation matches growth variation is necessary in the future. Model configuration with survey selectivity varying over time is suggested to be explored to compare the similarity and differences between the survey and the fishery. Such a comparison may help to understand whether the varying selectivity is because of fishing behaviour or policy changes, or because of growth changes, etc. Further exploration on the reasons or patterns of the changes in selectivity should also help to better estimate both BRPs and future stock projections.

Natural mortality *M* is fixed in the recommended base model, which is considered appropriate in estimating the overall population size and fishing mortality. Integrating uncertainty of *M* in the future stock assessment sis still worthwhile as more data including informative priors become available. Bayesian parameter estimation does not guarantee global maximum or minimum, and in many cases we look for biologically meaningful estimates. Biological meaningful priors for many of the key parameters, such as *M*, selectivity-at-age, and stock size from the survey combining together with the selection of MCMC algorithms often help convergence of the model to biologically meaningful estimates.

Although model comparison and selection are not immediately necessary, it is desirable to use a systematic framework and set of metrics to select the most appropriate model configuration from a series of model runs. Measurements in selecting models can be model goodness-of-fit, model prediction ability, model robustness and fisheries-specific measurements such as retrospective error (Jiao et al. 2012; Gelman et al. 2014; Hooten and Hobbs 2015).

I consider the assessment represents the best scientific information available for the stock assessment of BSAI yellowfin sole although improvements or adjustments in model structure development are possible. The review panel considered the BSAI yellowfin sole assessment sufficient to provide the basis for the management of this fishery.

3.2 Northern rock sole TORs

3.2.1 *Evaluate the strengths and weaknesses of the assumptions made in applying the stock assessment model including how survey indices are scaled to the populations. Specifics might include:*

a. How natural mortality estimates are estimated/applied

Natural mortality *M* is fixed to be 0.15/year for both sexes in the assessment team recommended base model. The value of *M* used for rock sole in other areas was assumed to be 0.2/year, which is based on the longevity of rock sole. For northern rock sole, alternative *M* values were evaluated, and the best point estimate of *M* = 0.18/year when *q* is fixed at 1, 0.159 and 0.19 for males and females when *q* is fixed at 1.5/year.

The assessment team also explored alternative runs by estimating *M* of one sex while fixing *M* at the other sex, with or without fixing *q*. The review team encouraged further exploration and incorporation of *M* parameter uncertainty in the future. However, the overall population status and fishery status are robust to all the alternative runs of *M* configuration.

b. Assumptions about survey "catchability"

Rock sole survey trawl catchability (*q*) was estimated at 1.4 from field experiments (Somerton and Munro 2001). Three configurations of *q* were explored by the assessment team: 1) fixed *q* = 1.5; 2) modelled *q* as a function of annual water temperature anomaly (T) at survey stations less than 100m, i.e., $q = e^{-\alpha+\beta T}$; and 3) estimate *q* but using estimated *q* from Somerton and Munro (2001) as prior distribution $N(1.4, 0.056^2)$. Both the second and third configurations resulted much higher mean *q* (1.98 – 2.22) than the estimate from the field experiments. Unlike the yellowfin sole, the relationship between survey biomass and T was found not significant for northern rock sole. The assessment team recommended the base model to use fixed *q* =0.15, which is close to the field experiment recommended *q* (0.14, Somerton and Munro 2001). Although the differences are limited, better rationale is needed to address the use of 0.15 versus 0.14.

c. Application of fishery and survey age-specific schedules (maturity, body mass, selectivity)

Two studies on maturity-at-age (Stark 2012) were conducted historically based on anatomical scan and histological analysis, and results from Stark (2012) based on histological analysis were used in this assessment. The differences in the estimated maturity were very limited for 50% maturity-at-age, but the differences between maturity curves are substantial.

Weight-at-age was estimated based on sex specific length-at-age and weight-at-length, then three year running average was used to develop sex specific weight-at-age matrix. Uncertainties from both submodels are not considered in the modelling framework.

Life history traits have been found to be largely influenced by both biotic and abiotic factors. For northern rock sole, length-at-age has been observed to vary across years. Examination on these life history traits every few years is suggested, if not every year.

Selectivity for both fishery and survey was modelled as a logistic function of age. Survey selectivity is assumed not to change over time, but fishery selectivity is assumed to vary every year. Further exploration on the reasons or patterns of the changes in selectivity should help to understand the fish/fishing behavior and to better estimate both BRPs and future stock projections. Whether selectivity variation pattern matches the variation pattern of growth is worth an exploration in the future. Comparison between age versus size-based selectivity seems worthwhile, and exploration on whether changes in selectivity-at-age are due to growth variation is necessary in the future.

d. The application (or lack thereof) of a stock-recruitment relationship (and associated parameter estimates)

Recruitment each year was estimated from the assessment model. A Ricker model was then fitted to the estimated recruitment and spawner biomass time series. Three models using three time period of recruitment and spawner biomass estimates were explored by the assessment team. The stock-recruitment relationship from the full time period was used to derive reference point and population projections. Northern rock sole demonstrated productivity changes over time, and the time period of recruitment to be used largely influence reference point estimates. The strategy used by the assessment team was found reasonable after exploring plots of recruitment per spawner, and estimates of stock recruitment relationships for different time periods. Future evaluation of factors influencing recruitment or the productivity of the population is recommended.

3.2.2 *Evaluate the stock assessment approach used focusing specifically on how fisheries and survey data are compiled and used to assess the stock status relative to stated management objectives under the Bering Sea and Aleutian Islands Fishery Management Plan (FMP) and the Magnuson-Stevens Act requirements. Elements should consider:*

a. The FMP "Tier" designation

Northern rock sole stock assessment provided reliable estimates of B_{MSY} and probability density function of F_{MSY}, and the projected spawning biomass is well above B_{MSY}, so that the FMP Tier designation for northern rock sole is Tier 1a. The review panel agrees with the assessment team and SSC.

b. Fishing rate estimation relative to overfishing definitions

Estimated full selection fishing mortality is 0.05/year in the four most recent assessment years (2013-2016), much smaller than F_{MSY} which is estimated to be 0.16/year; the reported catch is much lower than ABC and OFL (about 30%), so overfishing is not occurring. The review panel feel that the conclusion is reasonable and robust across all the model scenarios explored.

c. Stock status determinations relative to B_{MSY}

The estimated female spawning stock biomass is more than two times of B_{MSY} for northern rock sole, so the stock is not overfished. The review panel agree that the conclusion is reasonable and robust.

3.2.3 *Recommend how assessment data and/or models could be improved*

The model configuration and data used for northern rock sole are very close to those for yellowfin sole, so the recommendations for both species are similar.

The survey spawning biomass is used to calibrate both the trend and scale of the population size. Catchability plays a heavy role in the estimability or the stability of the model estimated population size of this species. For northern rock sole, extra informative of *q* from field experiment helped the assessment, but values should be used appropriately. Uncertainty of *q* may be incorporated into the assessment later as further model configuration exploration.

I support the research effort to explore changes of selectivity over time. Also, how different selectivity configurations improve model goodness-of-fit and influence stock assessment results (in the estimated population size and fishing mortality) needs to be quantified in the stock assessment document. Comparison between age versus size-based selectivity seems worthwhile and an exploration on whether changes in selectivity-at-age over time matches changes in growth is necessary in the future. Model configuration with survey selectivity varying over time is suggested to be explored to compare the similarity and differences between the survey and the fishery. Although a figure as figure 1 is not created for northern rock sole, visual comparison between patterns of selectivity and growth suggested that selectivity variation is correlated with growth changes. Such a comparison may help to understand whether the varying selectivity is because of fishing behaviour or policy changes, or because of growth changes, etc. Further exploration on the reasons or patterns of the changes in selectivity should help to better estimate both BRPs and future stock projections.

Natural mortality *M* is fixed in the recommended base model, which is considered appropriate in estimating the overall population size and fishing mortality. Integrating uncertainty of *M* in the future stock assessment is still worthwhile as more data including informative priors become available.

Biological meaningful priors for many of the key parameters, such as *M*, selectivity-at-age, and stock size from the survey often help convergence of the model to biologically meaningful estimates.

I consider the assessment represents the best scientific information available for the stock assessment of northern rock sole, although improvements or adjustments in model structure development are possible.

The review panel considered the northern rock sole assessment sufficient to provide the basis for the management of this fishery.

3.3 Alaska Plaice TORs

3.3.1 *Evaluate the strengths and weaknesses of the assumptions made in applying the stock assessment model including how survey indices are scaled to the populations. Specifics might include:*

a. How natural mortality estimates are estimated/applied

Natural mortality (*M*) is fixed at 0.13/year for both sexes in the assessment team recommended base model. In the past assessments *M* was fixed at 0.25 (Wilderburer and Walters 1997). The value of 0.13/year was from multiple sources of evidence. The estimated *M* based on life history traits ranges from 0.08-0.12 for males and 0.08-0.29 for females. The best point estimates by profiling over a range of *M* are 0.13 for both male and female Alaska plaice.

The review team eventually agrees with the assessment team about the selection of fixed *M* but suggest that future incorporation of uncertainty of *M* in the assessment can be further explored.

b. Assumptions about survey "catchability"

The catchability (*q*) is fixed at 1.2 for Alaska plaice. There are no specific field experiments on the *q* of Alaska Plaice. Somerton and Munro (2001) found that among the seven species studied, herding effect exists and the mean herding effect from the seven species resulted in a bridle efficiency of 0.234. So, the value of 1.2 is an approximation of the mean catchablility of all the studied species. Bottom water temperature, once doubted to be a factor influencing *q* as was found for yellowfin sole in this area. Past assessments did not find testable relationship between bottom water temperature and *q* (Spencer et al. 2004). This has been explained by the reality that Alaska plaice are a "cold loving" species with anti-freeze protein.

c. Application of fishery and survey age-specific schedules (maturity, body mass, selectivity)

Two studies on maturity-at-age (Zhang 1987, TenBrink and Wilderburer 2015) were conducted historically and results from TenBrink and Wilderburer (2015) were used in this assessment. The differences in the estimated maturity were very limited, and their influence on the estimated spawning stock size were limited also (about 4%).

Weight-at-age was estimated based on sex specific length-at-age and weight-at-length. Uncertainty from both submodels is not considered in the modelling framework.

Life history traits have been found largely influenced by both biotic and abiotic factors. For Alaska plaice, biological sampling on its growth variation seems limited. Examination on these life history traits every few years is suggested, if not every year.

Selectivity for both fishery and survey is modelled as a logistic function of age and does not change over time. If the species do not change growth obviously among years and do not change distributions according to climate regimes, then time varying selectivity-at-age is not a concern.

d. The application (or lack thereof) of a stock-recruitment relationship (and associated parameter estimates)

Recruitment of age 3 fish were estimated and explored during the review. A clear stock-recruitment pattern or model is difficult to identify. Alaska plaice demonstrated productivity changes over time, the strategy used by the assessment team was found reasonable. Because of lacking a reasonable stock-recruitment model, reference points derived from spawner-per-recruitment analysis and equilibrium recruitment were used to assess the fishery and population status. F_{OFL} is set to be $F_{35\%}$; F_{ABC} is set to be $F_{40\%}$; $B_{40\%}$ is used as B_{MSY} proxy.

3.3.2 *Evaluate the stock assessment approach used focusing specifically on how fisheries and survey data are compiled and used to assess the stock status relative to stated management objectives under the Bering Sea and Aleutian Islands Fishery Management Plan (FMP) and the Magnuson-Stevens Act requirements. Elements should consider:*

a. The FMP "Tier" designation

Alaska plaice stock assessment provided reliable estimates of $B_{40\%}$, $F_{40\%}$ and $F_{35\%}$, and the projected spawning biomass is 1.5 times of $B_{40\%}$ in 2018, so that the FMP Tier designation for Alaska plaice is Tier 3a. The review panel agrees with the assessment team.

b. Fishing rate estimation relative to overfishing definitions

Estimated full selection fishing mortality is 0.04-0.06/year in recent 5 years, much smaller than $F_{40\%}$ (0.124/year) and $F_{35\%}$ (0.149/year); the reported catch is lower than ABC and OFL (about 40-50%), so overfishing is not occurring. The review panel feels that the conclusion is reasonable and robust.

c. Stock status determinations relative to B_{MSY}

The estimated female spawning stock biomass (191,460 tonnes) is much higher than $B_{40\%}$ (126,900 tonnes) for Alaska plaice, so the stock is not overfished. The review panel feel that the conclusion is reasonable and robust.

3.3.3 *Recommend how assessment data and/or models could be improved*

The survey spawning biomass is used to calibrate both the trend and scale of the population size. Catchability plays a heavy role in the estimability or the stability of the model estimated population size of this species. The current catchability is fixed and borrowed information from other flatfishes. A direct field experiment should be very useful if possible in this case.

Natural mortality *M* is also fixed in the recommended base model, which is considered appropriate in estimating the overall population size and fishing mortality. Integrating uncertainty of *M* and *q* in the future stock assessment is still worthwhile as more data including informative priors become available.

Biological data collection for this species may be collected every few years if not every year to detect the potential changes in life history traits.

How Alaska plaice in the northern Bering Sea interact with the fishes in the currently assessed area, such as stock structure, mixing rate, dynamics characteristics, etc., may be studied over time.

Effort to study the factors influencing recruitment dynamics may be considered over time to better understand the productivity of Alaska plaice.

Projection of the population currently is done to meet the SSC's need and council's need. It may be done by considering the uncertainty of the current population size, the uncertainty in key parameters such as selectivity, growth, maturity, etc. The projections to include both alternative fishing mortality levels and the state of nature may be added from the scientific view of uncertainty evaluation.

I consider the assessment represents the best scientific information available for the stock assessment of Alaska plaice, although improvements or adjustments in model structure development are possible. The review panel considered the Alaska plaice assessment sufficient to provide the basis for the management of this fishery.

4. SUGGESTIONS FOR IMPROVEMENTS OF NMFS REVIEW PROCESS

The current review process is very well designed. This specific review done in AFSC was exceptionally organized both in the conduct of the meeting and in written and oral presentations of the assessment. I have no further recommendations about the review process.

5. ACKNOWLEDGMENTS

I would like to thank all the Stock Assessment Team members contributing to the meeting for their informative presentations on the stock assessments of Alaska flatfish and for providing helpful responses to the review panel's questions. Many thanks also to Bering Sea survey groups, and the Age and Growth Program and other scientists at the meeting for their presentations and their contribution to the discussions throughout the meeting. I also would like to thank Ms. Susan Robinson from Ocean Peace Inc., who contributed to the explanation of the A80 and TAC, etc. Special thanks also go to other members of the review panel, Drs. Joseph Powers and Matthew Cieri for productive discussions on the assessments.

6. REFERENCES

Gelman, A., Huang, J., and Vehtari, A. 2014. Understanding predictive information criteria for Bayesian models. Statistics and Computing. 24:997-1016.

Hooten, M.B., and Hobbs, N.T. 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85: 3-28.

Jiao, Y., Smith, E., O'Reilly, R., and Orth, D. 2012. Modeling nonstationary natural mortality in catch-at-age models: an example using the Atlantic weakfish (*Cynoscion regalis*) fishery. ICES Journal of Marine Science 69:105-118.

Nichol, D.R. 1995. Spawning and maturation of female yellowfin sole in the eastern Bering Sea. In Proceedings of the international flatfish symposium, October 1994, Anchorage, Alaska, p. 35-50. Univ. Alaska, Alaska Sea Grant Rep. 95-04.

Nichol, D.R. 1998. Annual and between sex variability of yellowfin sole, *Pleuronectes asper*, spring-summer distributions in the eastern Bering Sea. Fish. Bull., U.S. 96: 547-561.

Somerton, D., Weinberg, K., Munro, P., Rugolo, L. and Wilderbuer, T.K. 2018. The effects of wave-induced vessel motion on the geometry of a bottom survey trawl and the herding of yellowfin sole. Fish. Bull. 116:21–33.

Somerton, D.A., and Munro, P. 2001. Bridle efficiency of a survey trawl for flatfish. Fish. Bull. 99:641-652.

Stark, J.W. 2009. Contrasting maturation and growth of northern rock sole in the eastern Bering Sea and Gulf of Alaska for the purpose of stock management. N. Am. J. Fish. Manag., 32:1, 93-99.

TenBrink, T.T., and Wilderbuer, T.K. 2015. Updated maturity estimates for flatfishes (Pleuronectidae) in the eastern Bering Sea, with notes on histology and implications to fisheries management. Mar. Coast. Fish.: Dynamics, Management and Ecosystem Science, 0:1-9.

Wilderbuer, T.K. and Nichol, D.G. 2017. 2001. Yellowfin sole. In Stock Assessment and Fishery Evaluation Document for Groundfish Resources in the Bering Sea/Aleutian Islands Region as Projected for 2004, chapter 4. North Pacific Fishery Management Council, P. O. Box 103136, Anchorage, Ak 99510.

Wilderbuer, T.K. and Nichol, D.G. 2017. Assessment of the northern rock sole stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions.

Wilderbuer, T.K., and Nichol, D. 2017. Assessment of the Alaska plaice stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. North Pacific Fisheries Management Council, Anchorage, AK.

Wilderbuer, T.K., Nichol, D.G., and Ianelli, J. 2017. Assessment of the yellowfin sole stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions.

Zhang, C.I. 1987. Biology and population dynamics of Alaska plaice, *Pleuronectes quadrituberculatus*, in the eastern Bering Sea. Ph. D. dissertation, University of Washington.

Appendix 1: Bibliography of materials provided for review

Draft Stock Assessments:

Aydin, K. et al. 2017. Introduction to Bering Sea/Aleutian Islands SAFE. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions.

Somerton, D., Weinberg, K., Munro, P., Rugolo, L. and Wilderbuer, T.K. 2018. The effects of wave-induced vessel motion on the geometry of a bottom survey trawl and the herding of yellowfin sole. Fish. Bull. 116:21–33.

Wilderbuer, T.K. and Nichol, D.G. 2017. Assessment of the northern rock sole stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions.

Wilderbuer, T.K., and Nichol, D. 2017. Assessment of the Alaska plaice stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. North Pacific Fisheries Management Council, Anchorage, AK.

Wilderbuer, T.K., Nichol, D.G., and Ianelli, J. 2017. Assessment of the yellowfin sole stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions.

Presented Materials:

Wilderbuer, T.K. 2018. Overview of Bering Sea flatfish fisheries and management.

Haynie, A. 2018. The amendment 80 fishery: A long-term view on management changes that have impacted this North Pacific multi-species fishery.

Concepcion, M. 2018. North Pacific observer program, Alaska Fisheries Science Center.

Britt, L. et al. 2018. Eastern Bering Sea shelf bottom trawl survey of groundfish and invertebrate resources. Groundfish Assessment Program RACE Division.

Anderl, D., and Matta, B. 2018. Age & growth program at the Alaska Fisheries Science Center.

Stram, D. 2018. North Pacific Fishery Management Council, overview of fishery management.

Wilderbuer, T.K., Nichol, D., and Ianelli, J. 2018. CIE review of the Bering Sea/Aleutian Islands yellowfin sole stock assessment.

Wilderbuer, T.K., Nichol, D., and Ianelli, J. 2018. CIE review of the Bering Sea/Aleutian Islands northern rock sole stock assessment.

Wilderbuer, T.K., and Nichol, D. 2018. CIE review of the Bering Sea/Aleutian Islands Alaska plaice stock assessment.

Appendix 2: Statement of Work for Dr. Yan Jiao

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review of
Fisheries stock assessments for yellowfin sole,
northern rock sole and Alaska plaice

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

Scope

The Alaska Fisheries Science Center's (AFSC) Resource Ecology and Fisheries Management Division (REFM) requests an independent review of the integrated stock assessments that have been developed for three Bering Sea flatfish species; yellowfin sole, northern rock sole and Alaska plaice. The fishery for these species is managed by the North Pacific Fisheries Management Council. The sum of the ABCs for these three species is 455,200 metric tons (t) in 2018, with catch levels annually set lower than the ABC due to a 2.0 million t harvest cap for all species and constraints due to Pacific halibut bycatch limits and markets. The catch limits are established using automatic differentiation (AD) Model software that uses survey abundance data and survey and fishery age and length composition data with a harvest control rule to model the status and productivity of these stocks and set quotas. Having these assessments vetted by an independent expert review panel is a valuable part of the AFSC's review process. The Terms of Reference (TORs) of the peer review and the tentative agenda of the meeting are below.

[1] http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf

Requirements

NMFS requires three reviewers to conduct an impartial and independent peer review in accordance with the SOW, OMB Guidelines, and the TORs below. The reviewers shall have working knowledge and recent experience in the application of fisheries stock assessment processes and results, including population dynamics, separable age-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management.

Tasks for reviewers

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

Wilderbuer, T.K., D.G. Nichol, and J. Ianelli, 2017. Assessment of the yellowfin sole stock in the
Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. https://www.afsc.noaa.gov/REFM/stocks/assessments.htm

Wilderbuer, T.K. and D. G. Nichol. 2017. Assessment of the northern rock sole stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. https://www.afsc.noaa.gov/REFM/stocks/assessments.htm

Wilderbuer, T.K. and D. Nichol. 2017. Assessment of the Alaska plaice stock in the Bering Sea and Aleutian Islands. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. North Pacific Fisheries Management Council, Anchorage, AK. https://www.afsc.noaa.gov/REFM/stocks/assessments.htm

Introduction to Bering Sea/Aleutian Islands SAFE https://www.afsc.noaa.gov/REFM/stocks/assessments.htm

Somerton, D., K. Weinberg, P. Munro, L. Rugolo and T. Wilderbuer. 2017. The effects of wave-induced vessel motion on the geometry of a bottom survey trawl and the herding of yellowfin sole. Fish. Bull. 116:21–33 (2018). doi: 10.7755/FB.116.1.3 https://www.st.nmfs.noaa.gov/spo/FishBull/1161/somerton.pdf

2. Attend and participate in the panel review meeting
   - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
3. After the review meeting, reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus
4. Each reviewer may assist the Chair of the meeting with contributions to the summary report, if required by the TORs
5. Deliver their reports to the Government according to the specified milestone 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 40 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 and Period of Performance

The place of performance shall be at the contractor's facilities, and at the Alaska Fisheries Science Center, Seattle, Washington, USA. The period of performance shall be from the time of the award through June 1, 2018. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables

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

Within two weeks of award	Contractor selects and confirms reviewers
No later than April 12, 2018	Contractor provides the pre-review documents to the reviewers
April 16-18, 2018	Panel review meeting
May 11, 2018	Contractor receives draft reports
May 25, 2018	Contractor submits final reports to the Government

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

1. The reports shall be completed in accordance with the required formatting and content
2. The reports shall address each TOR as specified
3. The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

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 $11,000.

Restricted or Limited Use of Data

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

NMFS Project Contact

Thomas K Wilderbuer
7600 Sand Point Way NE, Bldg. 4
Seattle, WA 98115-6349
tom.wilderbuer@noaa.gov

Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.
2. The 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 Statement of Work
   Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Terms of Reference for the Peer Review

For the three assessments (Bering Sea and Aleutian Islands yellowfin sole, northern rock sole, and Alaska plaice) consider the following:

1. Evaluate the strengths and weaknesses of the assumptions made in applying the stock assessment model including how survey indices are scaled to the populations. Specifics might include:
    a. How natural mortality estimates are estimated/applied
    b. Assumptions about survey “catchability”
    c. Application of fishery and survey age-specific schedules (maturity, body mass, selectivity)
    d. The application (or lack thereof) of a stock-recruitment relationship (and associated parameter estimates)
2. Evaluate the stock assessment approach used focusing specifically on how fisheries and survey data are compiled and used to assess the stock status relative to stated management objectives under the Bering Sea and Aleutian Islands Fishery Management Plan (FMP) and the Magnuson-Stevens Act requirements. Elements should consider:
    a. The FMP “Tier” designation
    b. Fishing rate estimation relative to overfishing definitions
    c. Stock status determinations relative to B_{MSY}
3. Recommend how assessment data and/or models could be improved.

CIE Flatfish Assessment Review Agenda

NMFS Alaska Fisheries Science Center
April 16-18, 2018, room 2079
7600 Sand Point Way NE, Building 4
Seattle, Washington

Monday April 16th

9:00	Welcome and Introductions, adopt agenda, TORs	
9:15	Overview	**Tom Wilderbuer**
10:00	Bering Sea shelf trawl survey	**Dan Nichol**
	Observer Program	**Marlon Conception**
11:00	Coffee break	
11:20	Fisheries Management and Council process	**Diana Stram**
12:30	Lunch	
1:30	Age Determination	**Delsa Anderl and Beth Matta**
	BSAI Yellowfin sole	**Tom Wilderbuer and Jim Ianelli**

Tuesday April 17th

9:00	Effect of rationalization on flatfish fisheries	**Alan Haynie**
10:00	BSAI Northern rock sole	**Tom and Jim**
11:00	Coffee break	
11:20	BSAI Northern rock sole (continued)	**Tom and Jim**
12:30	Lunch	
1:30	BSAI Alaska plaice	**Tom and Jim**

Wednesday April 18th

9:00	CIE panel discussion (assessment authors will be available)
12:30	Lunch
	Continue as needed

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

CIE Reviewers

Joe Powers
Yan Jiao
Matthew Cieri

Presenters from Stock Assessment Team

Tom Wilderbuer
James Ianelli
Daniel G. Nichol

Presenters from Alaska Survey, Observer, and Age and Growth Programs

Alan Haynie, Marlon Concepcion, Lyle Britt, Delsa Anderl, Beth Matta, Diana Stram,

Other participants

Susan Robinson, Ocean Peace, Inc.

Appendix 4: A list of requests from CIE review panel

Day 1: Request or explored for extra runs for yellowfin sole stock assessment

Potential short-term requests:

- Constant fishery selectivity
- Start the model in 1982, ignore historical catches and other data
- Examine sex-specific natural mortality
- Time-varying selectivity for survey (if possible) and relationship w/ fishery. (not done during the review)

Longer term that may be explored in the future:

- size-based selectivity and potential interactions with growth changes
- Examine plus-group
- Retrospective patterns with full model for survey catchability
- Age-specific natural mortality (e.g., Lorenzen) might be considered (but since fishery ages mostly older, may not matter so much)

Extra runs named as a new base : fix M=0.12 but estimating q ~ f(T) and its retrospective pattern

Day 2: Request for extra model runs for rock sole and Alaska Plaice

Rock sole:

- Compare model 15.1 and 16. By comparing the time series of B, F, fit to survey biomass and posterior of the reference points

Alaska plaice:

- Model starting from 1982
- Explore the spawning and recruitment pattern
- Exploring the possibility of estimating either q or M in the future