Petrale Sole
Stock Assessment Review (STAR) Panel Report

National Oceanic and Atmospheric Administration
Northwest Fisheries Science Center
Auditorium and Online
2725 Montlake Boulevard E
Seattle, WA 98112

July 24-28, 2023

Participants

STAR Panel Members

John Field, National Marine Fisheries Service, Southwest Fisheries Science Center (Chair)
Kristin Marshall, National Marine Fisheries Service, Northwest Fisheries Science Center
Joseph Powers, Center for Independent Experts
Martin Cryer, Center for Independent Experts

Stock Assessment Team (STAT) Members

Ian G. Taylor, National Marine Fisheries Service Northwest Fisheries Science Center
Vladlena Gertseva, National Marine Fisheries Service Northwest Fisheries Science Center
Nick Tolimieri, National Marine Fisheries Service Northwest Fisheries Science Center

STAR Panel Advisors

Whitney Roberts, Washington Department of Fish and Wildlife,
Groundfish Management Team representative
Gerry Richter, B&G Seafoods, Groundfish Advisory Subpanel representative
Marlene A. Bellman, Pacific Fishery Management Council representative

Overview

A Stock Assessment Review (STAR) panel was convened on July 24-28, 2023, at the Northwest Fishery Science Center's Montlake Laboratory, under the Pacific Fishery Management Council's (PFMC) Terms of Reference for Groundfish Stock Assessment Review Process for 2023-2024 (PFMC, June 2022). Petrale sole is a highly desirable, high attainment stock found throughout the California Current (although infrequently encountered south of Point Conception), with a long exploitation history. The stock was estimated to be well below target level from the 1960s through the early 20th century. The stock was determined to be overfished in 2009, but found to be rebuilt in a 2015 update assessment, following a pulse of strong recruitment in the mid-2000s. The 2013 benchmark petrale sole assessment was most recently updated in 2019.

Drs. Ian Taylor and Vladlena Gertseva of the Northwest Fisheries Science Center (NWFSC) presented the 2023 petrale sole assessments to the panel, other members of the Stock Assessment Team (STAT) were present at the meeting and provided additional details on various aspects of the model. In general, model estimates and trends are very similar to those from previous assessments, although the scale of age 3+ biomass is lower, and the population is estimated to be declining as a consequence of recent below-average recruitments. However, the base model results, and those of most model sensitivities, indicate that the stock is at 33.6% of the estimated unfished equilibrium spawning output level, well above the 25% target level.

Summary of Data and Assessment Models

The 2023 petrale sole assessment was a fully integrated age-structured bench-mark assessment conducted in Stock Synthesis (SS3 Version 3.30.21.00) using catch, length, age, and index data from fishery dependent and independent sources. Natural mortality, growth and recruitment are estimated, while steepness (h) is fixed at 0.80. Sex-specific selectivity with time blocks based on important management changes and milestones are a key model feature.

The fishery is almost entirely comprised of commercial trawl effort and landings. Landings data extend to at least 1900 in California waters, and fisheries extended north to Oregon and Washington waters during the 1930s. Historical catch estimates for California and Oregon were essentially unchanged between the last benchmark and this assessment, although historical catch estimates for the Washington fishery changed substantially, to considerably lower levels, as a result of a more comprehensive catch reconstruction developed by WDFW. Discards have been low both historically and recently for most years of the fishery.

Although previous models distinguished summer fisheries from winter fisheries (in which winter fisheries targeted spawning ground), the 2023 model combined previous seasonal fisheries into a single annual fleet. This reduced the model complexity and number of parameters needed to inform separate selectivity curves, while providing results highly comparable to earlier models. However, the 2023 assessment continued the approach of having separate bottom trawl fisheries north and south of 42° N (California/Oregon border). Fisheries length and age composition data are extensive for the northern fleet, with some (generally limited) data available historically to the 1940s and 1950s, although age composition data are fairly sparse for the southern (California) fishery. A fishery-dependent CPUE index that was included in earlier models was excluded from the 2023 model, as the influence of the index was minimal and survey data are considered to be considerably more robust and reliable.

Survey data from the historical triennial bottom trawl survey, as well as the West Coast Groundfish Bottom Trawl Survey (WCGBTS) are key model inputs, with indices developed using sdmTMB. One significant change from the 2013 model was the treatment of the triennial survey index as a single time series, rather than distinct early and late time series. Another key change in the 2023 model is the removal of the fishery dependent CPUE index that was used in earlier models. Reasons for this include concerns over the hyperstability of catch rates when fishing on spawning grounds, and the fact that the nearly 20 year time series of survey data from the WCGBTS provide robust estimates of abundance. Age composition data are available only from WCGBTS, and treated as conditional-age-at-length to better inform estimates of growth internally.

An environmentally driven recruitment index, recently developed by Nick Tilimieri using data from the Copernicus Marine Environment Monitoring Service (CMEMS) and Mercator Ocean International (MOI) and following the approach used in Haltuch et al. (2020), was explored as a sensitivity in the model. Haltuch et al. (2020) had previously developed an environmental index of recruitment for petrale sole using products from a California Current Regional Ocean Modeling System (ROMS) model. However, when the ROMS model used in that analysis was updated to include years from 2011 forward, the outputs exhibited some distinct discontinuities with those estimated in the 1980–2010 period, indicating changes in scale and trend across the 2010/2011 boundary that are generally considered to be model artifacts. Ongoing analyses of alternative model products show promise for the development of a comparable index from alternative data sources, but the STAT indicated that the index would benefit from additional development and review before being included in the assessment base model.

Requests by the STAR Panel and Responses by the STAT

Request No. 1*:* Provide additional description of the new WA catch reconstruction. Provide a model run with the current base model and the previous assessment's historical catch reconstruction from Washington. Compare spawning output and depletion and provide equilibrium Maximum Sustainable Yield (MSY) for each model.

Rationale: While the new WA catch reconstruction is an improvement over the previous assessment, it is strongly influencing the estimates of spawning output. The Panel would like to better understand how the parameterization of the 2023 base model interacts with the new Washington Department of Fish and Wildlife (WDFW) historical catch reconstruction to influence the perception of spawning output.

STAT Response: The STAT provided additional information about the historical catch reconstruction from Washington, which will be summarized in the post-STAR version of the stock assessment. Briefly, two historical reports inform the landings from 1938-1969, but these reports combine Canadian catches with those from Washington (Alverson and Chatwin 1957, Ward et al. 1969). Alverson and Chatwin (1957) used various reports and data to estimate proportions of catches between Washington and British Columbia for 1948-1955. Based on their work, an average of 11.6 percent was used to apportion the catches observed from 1938 to 1969 (aside from 1948-1955, which were estimated on a year-specific basis in that study) to Washington state waters. From 1970-1980, Washington catch areas were based on fish receiving ticket data that are not available in PacFIN.

The STAT presented a model run using the base model with the 2019 historical catch reconstruction from WA, as requested. The 2019 reconstruction influenced the historic spawning output but had little effect on depletion, and a less than expected influence on the equilibrium MSY (2680 mt versus 2482 mt in the base model).

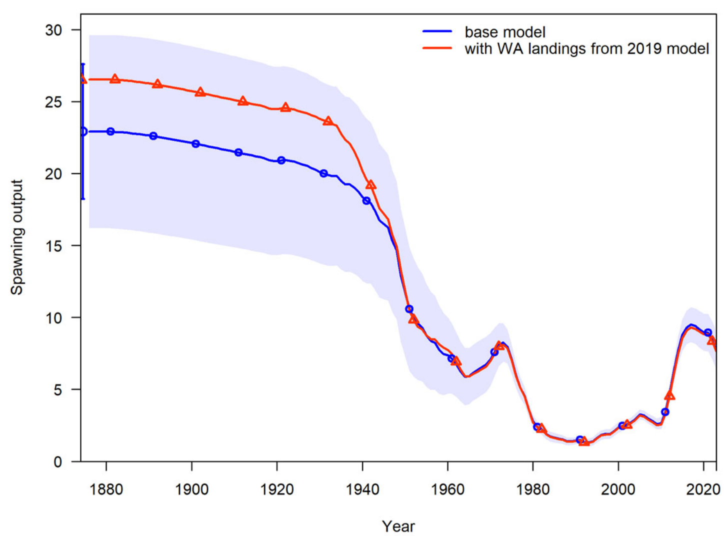


Figure1. Spawning output time series of base model run that uses the Washington historical catch reconstruction from the previous (2019) stock assessment.

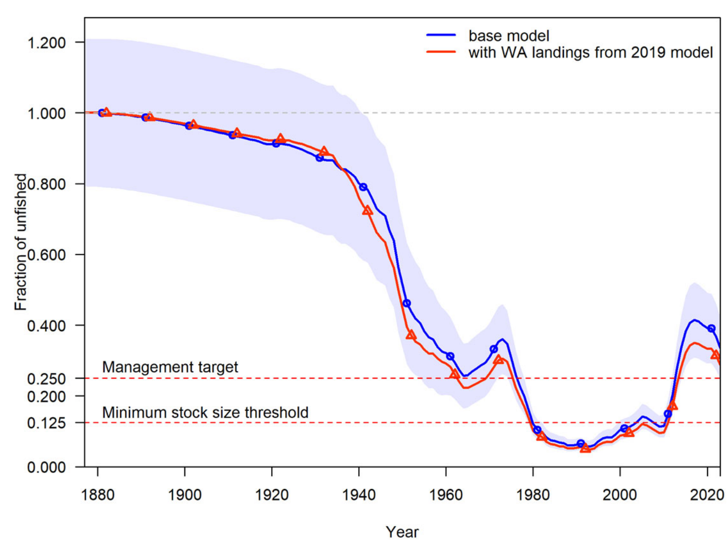


Figure 2. Depletion time series of base model run that uses the Washington historical catch reconstruction from the previous (2019) stock assessment.

Panel conclusion: Using the 2019 catch reconstruction had less influence than the Panel had expected on the spawning output, and in particular, the equilibrium MSY. Including the Washington catch reconstruction from the 2019 assessment to the current base model alone was responsible for less than half of the change in MSY. Therefore, changes in the life history dynamics (reduced M and higher steepness), and estimates of historical discards in the base model contributed to the changes in perception of the stock.

Request No. 2: Provide plots of the fits to the West Coast Groundfish Bottom Trawl Survey (WCGBTS) index in the base model and in the sensitivity run that added the environmental recruitment index (pre-STAR draft assessment Section 2.4, Appendix A).

Rationale: The Panel would like to understand how including the environmental recruitment index would influence the fit to the WCGBTS index in the most recent decade.

STAT Response: The STAT provided the requested model fits. The environmental index suggests higher recruitment at the end of the time series, which leads to better fit to the final index value (2022) and a less steeply declining trend over the last five years. The likelihood across all years indicates a better fit to the Triennial survey and WCGBTS index with the base model.

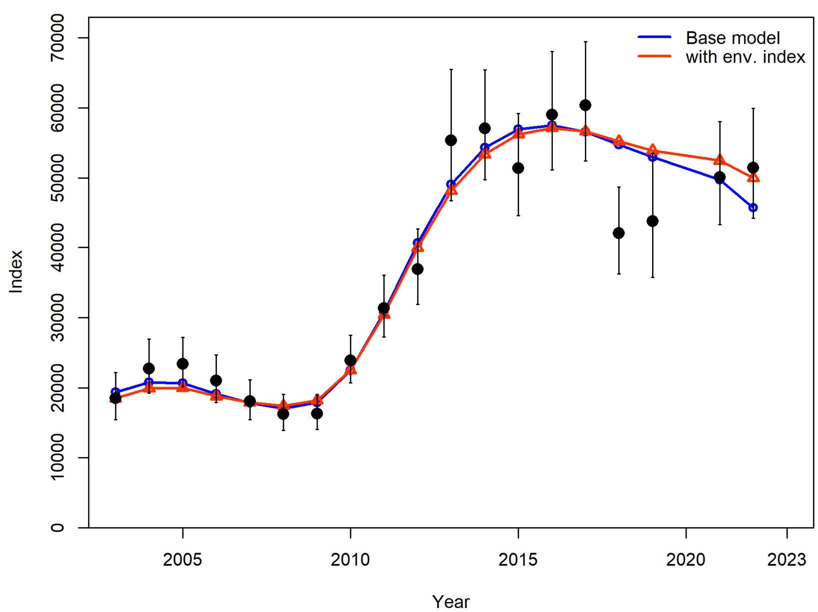


Figure 3. Model fit to the WCGBTS index from the base model and a model with the environmental index of recruitment added.

	Triennial	WCGBTS
Base	-4.98	-31.87
Env. Index	-4.64	-29.86

Panel conclusion: The Panel agrees with the STAT's summary. It makes sense that the environmental index is most influential in the most recent years of observations when there is less data informing the model. The difference in overall fits to the two survey indices between the two model runs is very small, given the change in NLL units is only 0.34 for the Triennial index and 2.01 units for the WCGBTS. The Panel sees merit to considering including this index in future assessments.

Request No. 3*:* Present standard profiles a.) with recruitment deviations turned off, and b.) using sigmaR of 0.3 to 0.8, incrementing by 0.1.

Rationale: The Panel would like to determine the impact of different values of sigmaR on the stock assessment.

STAT Response: The STAT provided the requested profile, however when sigmaR was increased to 0.8 the model did not converge. The maximum likelihood estimate is 0.49 (the base model used 0.5). A fairly broad range of simaR values are plausible (NLL change <2), from about 0.4 to 0.65.

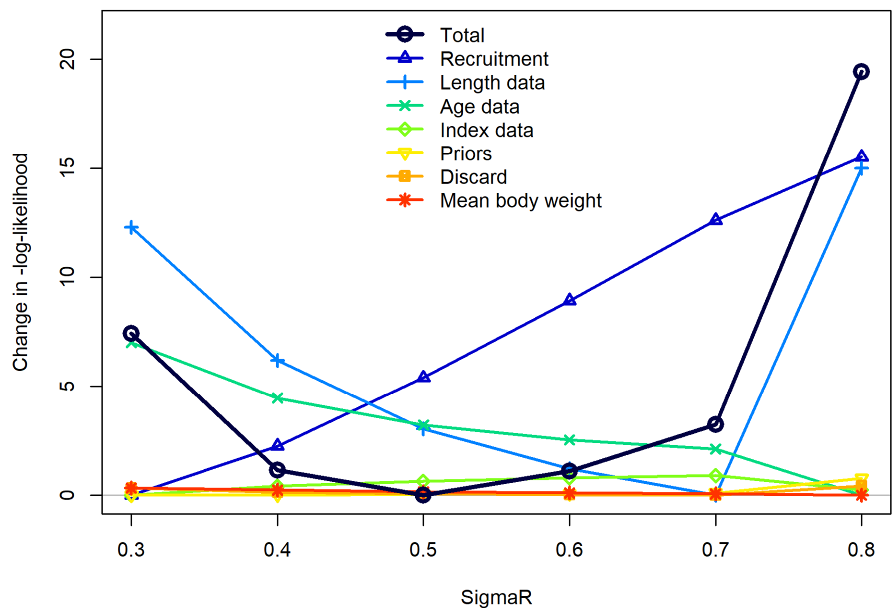


Figure 4. Likelihood profile over sigmaR showing all likelihood components.

The STAT noted that the recruitment likelihood component includes the bias adjustment parameter (Methot and Wetzel 2013), which was not adjusted with changing sigmaR values. Changing the corresponding bias adjustment values would be best practice if these runs were being explored as potential alternative parameterizations to the base model (Methot and Taylor 2011).

The model run that does not estimate recruitment deviations demonstrates the effects of changes in fishing pressure over time and the effect of recent above average recruitments. The effect of increasing sigmaR is predictably increasing variable recruitment deviations, but this results in little effect on the spawning output or depletion time series.

The tuning algorithm within r4ss provides additional information about how the specified sigmaR is performing (based on simulation analysis conducted in Methot and Taylor 2013). The STAT provided a plot of these results from each of the runs in the sigmaR profile. This provides additional support for the choice of sigmaR = 0.5 in the base model.

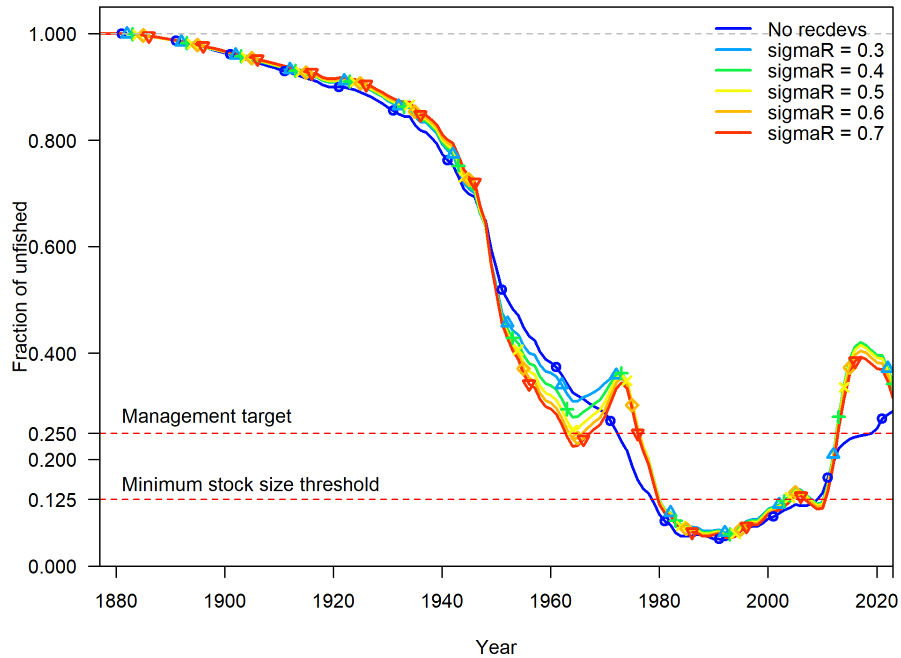


Figure 5. Depletion time series assuming sigmaR values specified in the likelihood profile. An additional run is included with no recruitment deviations estimated.

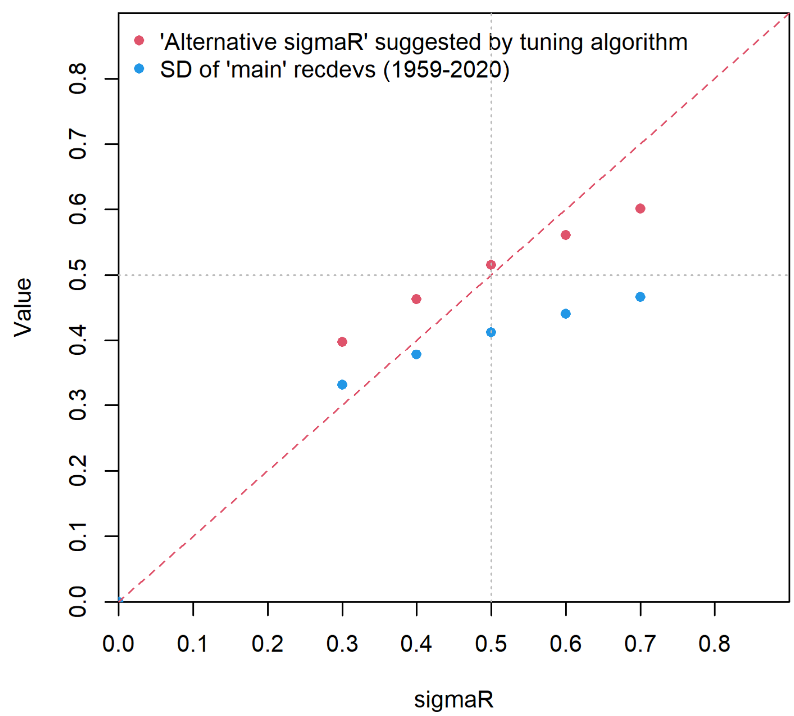


Figure 6. The relationship between the sigmaR specified in the profile and the alternative value suggested by the tuning algorithm.

Panel conclusion: The likelihood profile over sigmaR, the model run without recruitment deviations estimated, and the analysis of starting and ending value of sigmaR provided useful diagnostics. The Panel agrees with the STAT's summary and supports the base model sigmaR of 0.5.

Request No. 4*:* Provide model runs that a) eliminate the extra SD on the Triennial survey, b) eliminate the extra SD on the Triennial survey and upweight it (e.g., set lambda = 10) and c) remove the Triennial survey. If possible, show any earlier evaluations that exist for the fits of the Triennial survey to other flatfish species.

Rationale: The Panel would like to better understand how the fit to the Triennial survey is affecting overall model behavior and estimates of spawning output. Additional evaluations from other species would provide context for the Panel.

STAT Response: The STAT provided the requested model runs. Eliminating the extra SD and upweighting the Triennial survey index allowed an improved fit to 2004. This model run led to implausibly low spawning output and a less depleted stock and an implausibly abundant stock in the case of the upweighting and would likely lead to degraded fits to the length data. In contrast, removing the Triennial survey index had negligible effects on spawning output and depletion time series. Length data from the Triennial were retained in the model during this exploration, so this is not unexpected. Mean lengths from the fishery and survey data showed declines during the 1970s and 1980s.

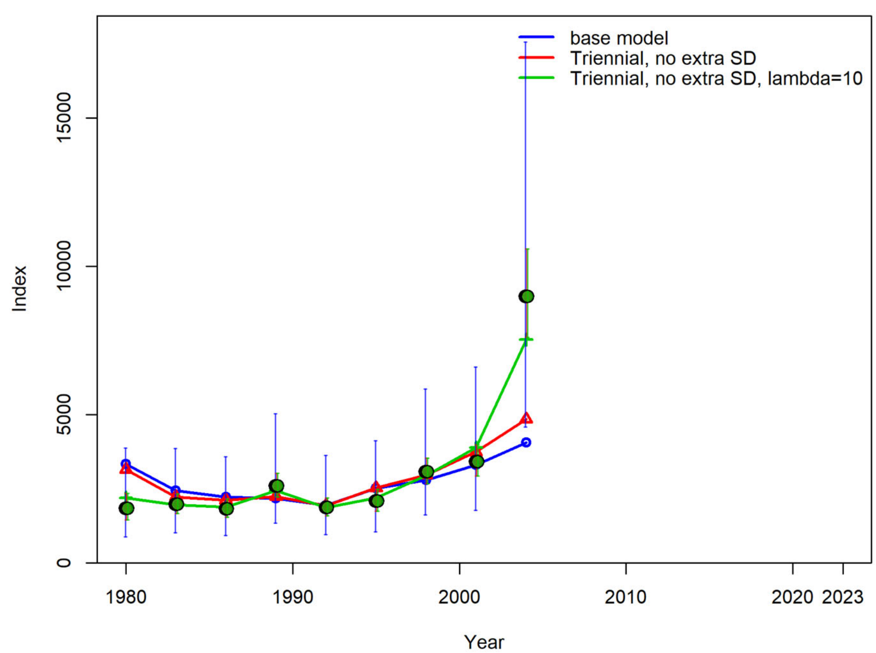


Figure 7. Model fits to the Triennial survey index when the index is upweighted.

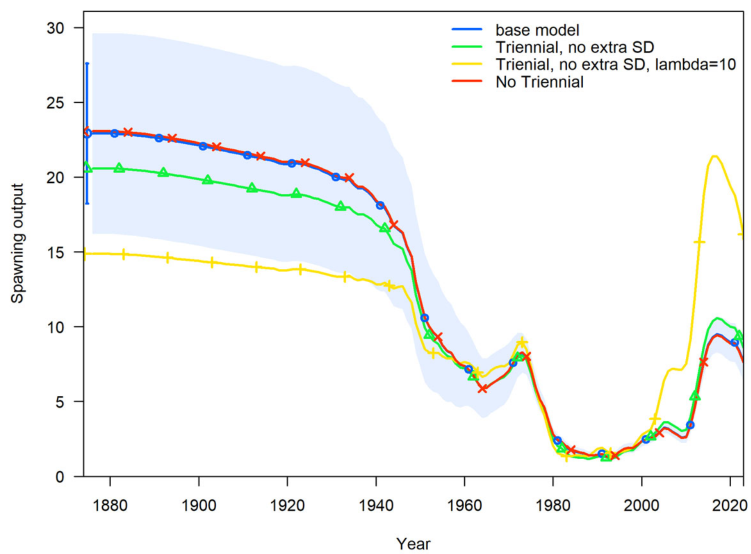


Figure 8. Spawning output time series for sensitivity runs that upweight and remove the Triennial survey index.

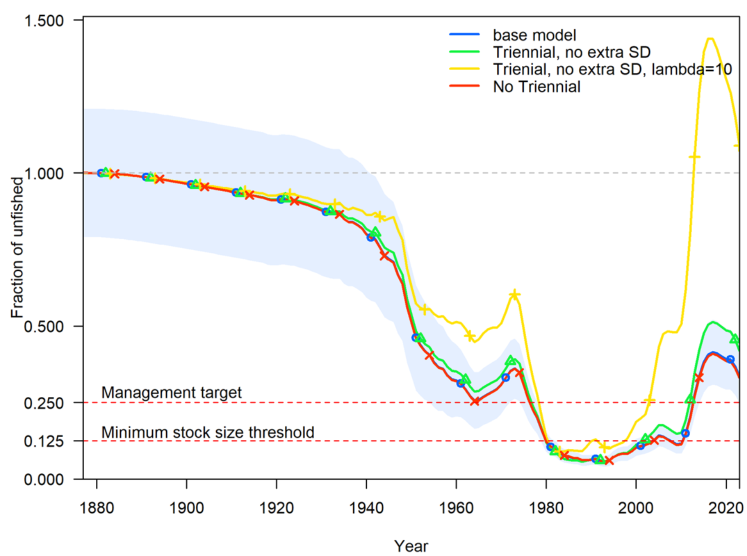
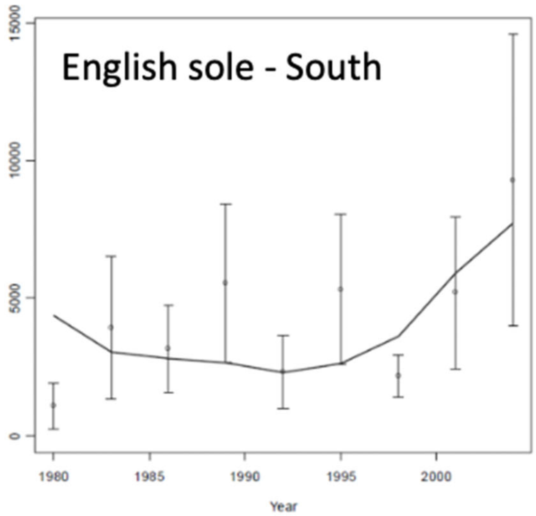
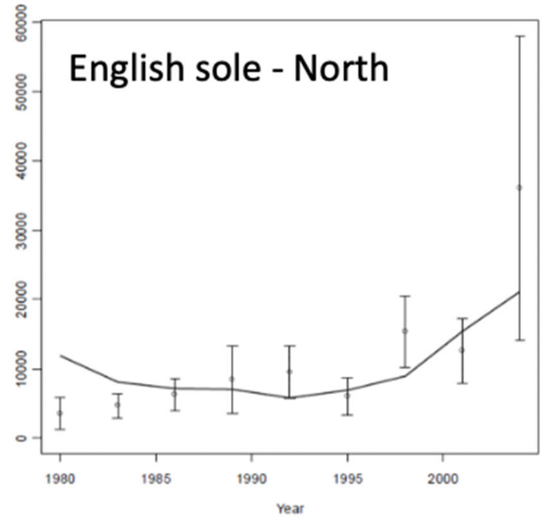
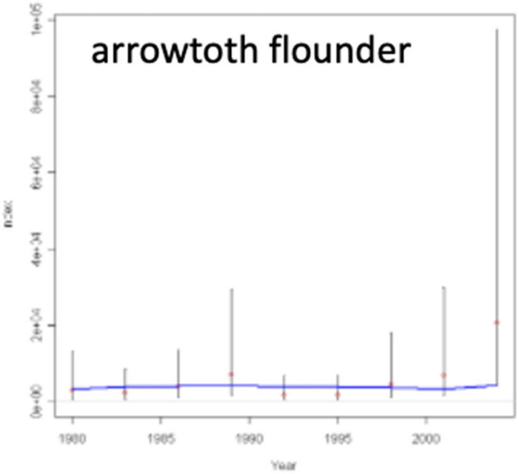


Figure 9. Depletion time series for sensitivity runs that upweight and remove the Triennial survey index.

The STAT provided fits to the Triennial survey from assessments of other flatfish species and two skates (English sole, arrowtooth flounder, Dover sole, rex sole, longnose skate and big skate). All species demonstrated relatively poor fits in 1980 and 2004. They were more successful at fitting the 2004 datapoint by adding extra SD on the survey index (all species), separating the Triennial survey into early and late periods (rex sole), and/or utilizing a time-varying catchability (q) parameter (Dover sole, longnose skate, big skate). Plausible explanations exist for the 1980 poor fits when the survey was still in its early years and trawling performance may have been poorer.







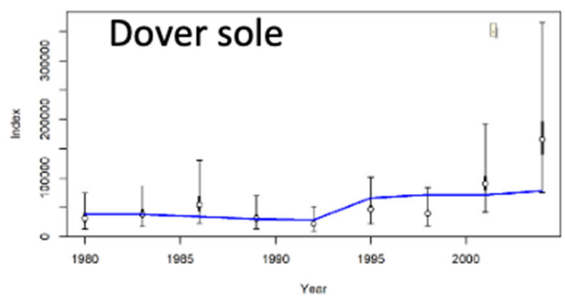


Figure 93: Fit to the Triennial survey index of abundance. Thicker lines indicate input uncertainty before the addition of estimated additional uncertainty parameter.

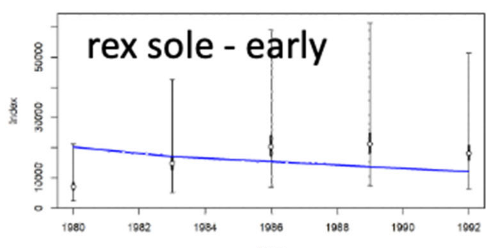


Figure 23: Fit to index data for the early period of the AFSC/NWFSC West Coast Triennial Shelf Survey (Triennial Survey). Lines indicate 95% uncertainty interval around index values. Thicker lines indicate input uncertainty before addition of estimated ad uncertainty.

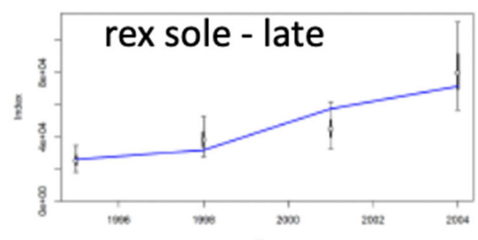
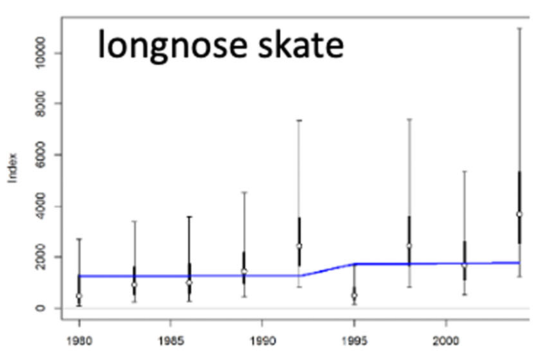


Figure 24: Fit to index data for the late period of the AFSC/NWFSC West Coast Triennial Shelf Survey (Triennial Survey). Lines indicate 95% uncertainty interval around index values. Thicker lines indicate input uncertainty before addition of estimated additional uncertainty.



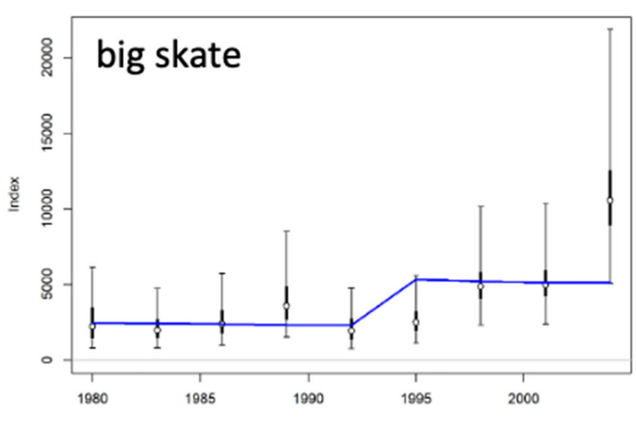


Figure 10. Stock assessment model fits to the Triennial survey index for four flatfish and two skate species.

Panel conclusion: The Panel agrees that upweighting the Triennial survey leads to better fits to 2004 but implausible spawning output and relative stock status. Comparisons with other species are helpful to support the conclusion that 1980 and 2004 are difficult to fit for other assessments beyond petrale sole. Plausible explanations exist for poor fits in 1980 while the survey configurations were new and trawling performance may have been lower. Future work could explore the 1980 and 2004 survey years and investigate bottom contact time and trawl start and end points to explore factors contributing to poor fits and potentially biased indices in those years.

Request No. 5*:* Provide a model run that shifts the start of the most recent commercial selectivity time block from 2018 to 2017.

Rationale: The Groundfish Management Team (GMT) representative pointed out the shift in selectivity blocks in 2017 would be more consistent with changes in management. The Panel would like to better understand how changing this selectivity block could affect the pearson residuals of fits to the age and length compositions.

STAT Response: Shifting the start of the last time block on selectivity from 2018 to 2017 resulted in a worse fit to the length comps (3 NLL units), no change in the fit to the age data, and negligible impacts on spawning output and depletion. Given that the block was implemented to fit a change in length compositions observed first in 2018, and the slightly improved fit with the block starting in 2018, the STAT prefers to keep the block as specified in the base model.

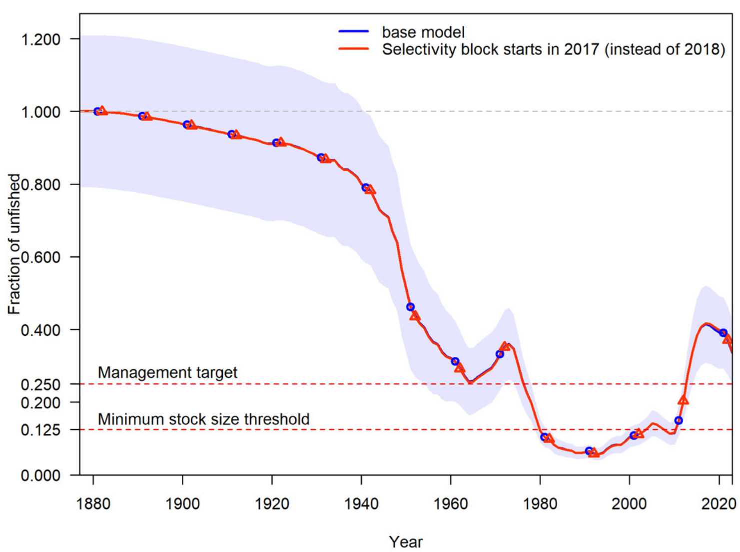


Figure 11. Depletion time series for sensitivity run with the start of the most recent time block on selectivity starting in 2017 compared with the base model where the block starts in 2018.

Panel conclusion: The Panel agrees that there is negligible difference between the two model runs and supports the STAT's preference to keep the block as specified in the base model given the slightly improved fit using the selectivity block starting in 2018.

Request No. 6: Provide a model run using the environmental recruitment index and re-estimates steepness (*h*).

Rationale: The panel would like to better understand how the addition of the environmental recruitment index would affect the relationship between estimates of recruitment deviations and steepness.

STAT Response: Estimating steepness after including the environmental index still results in h=1. As in the base model, when steepness is estimated to be higher, M is estimated to be considerably lower (M=0.1) than the base model. This low value of M would correspond with a maximum age of 50 based on the Hamel prior relationship between maximum age and M.

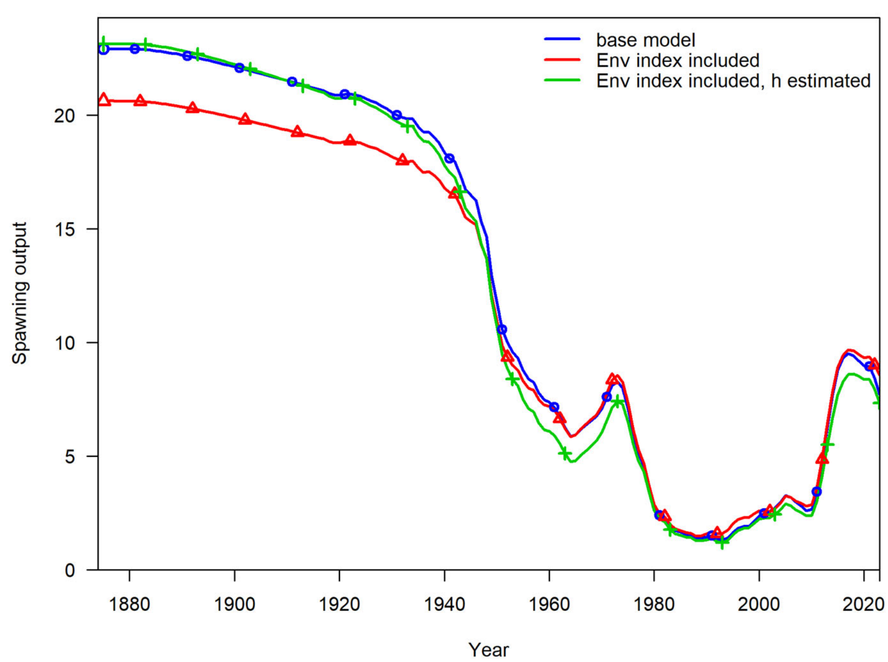


Figure 12. Spawning output time series from requested sensitivity run with the environmental index added and steepness estimated compared with the base model and sensitivity to the addition of the environmental index provided in the pre-STAR assessment report.

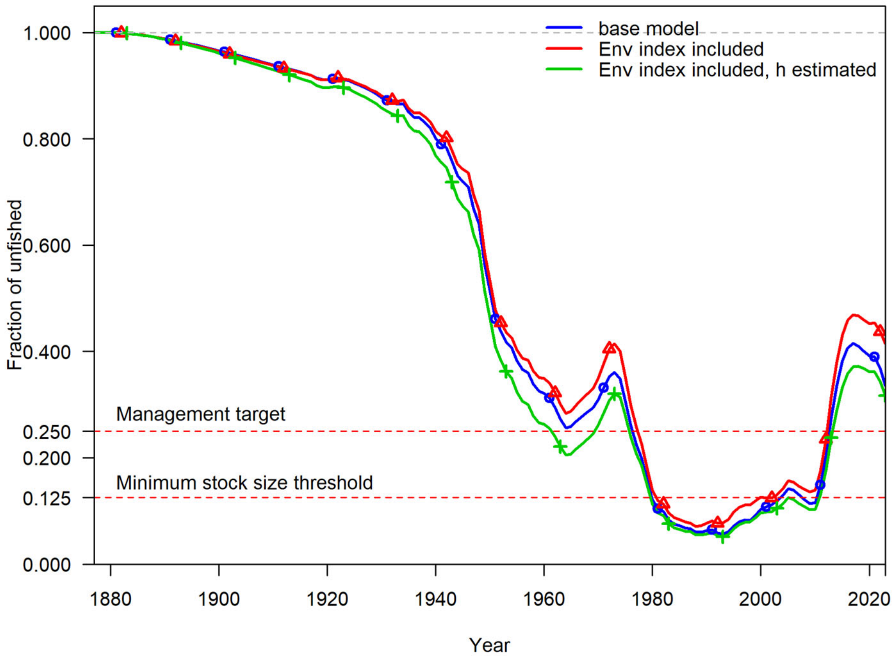


Figure 13. Depletion time series from requested sensitivity run with the environmental index added and steepness estimated compared with the base model and sensitivity to the addition of the environmental index provided in the pre-STAR assessment report.

Panel conclusion: Re-estimating steepness after adding the environmental index did not prevent the base model from estimating implausibly high steepness. This sensitivity suggests that the estimate of steepness may be insensitive to the addition of an environmental index as data informing age-0 recruitment deviations. Future research could investigate using the environmental index to inform time-varying parameters of the stock recruitment relationship (h or R0), however this is not possible in the current version of Stock Synthesis.

Request No. 7: Provide model runs that run the new base model with a) the sex-specific M estimates from the actual 2019 assessment from the corrected version of Table 24, b) the steepness value from the 2019 assessment, and c) combine a) and b). Compare spawning output and depletion and provide equilibrium Maximum Sustainable Yield (MSY) for each model.

Rationale: The Panel would like to better understand how the parameterization of the 2023 base model contributes to the Panel's perception of overall productivity. The STAT discovered an error in Table 24 in the pre-STAR report in the reporting of parameter values from the 2019 stock assessment. This will be corrected in the post-STAR assessment document.

STAT Response: Increasing sex-specific M from the base model to the values from the 2019 assessment caused historic spawning biomass to be lower, as did changing both M and steepness. None of these resulted in a substantial change to equilibrium MSY, which varied by less than 30 mt for all of these sensitivities.

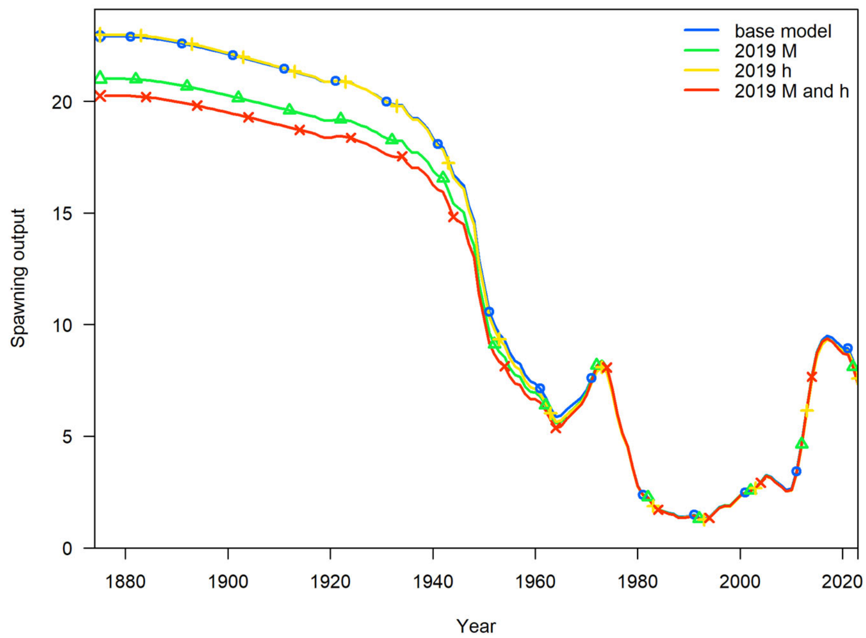


Figure 14. Spawning output time series for sensitivity runs requested using values for M and h from the 2019 stock assessment.

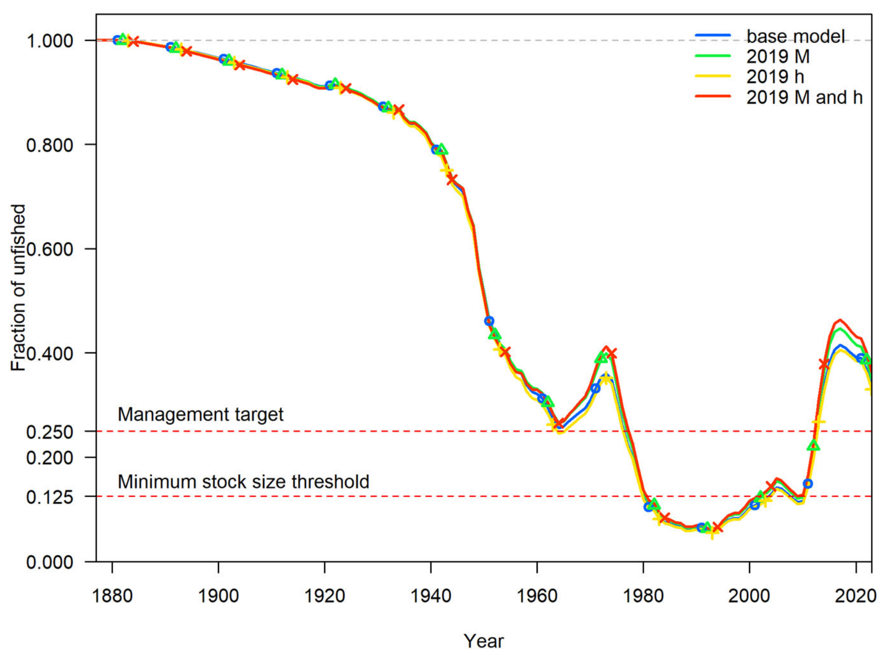


Figure 15. Depletion time series for sensitivity runs requested using values for M and h from the 2019 stock assessment.

Panel Conclusion: These runs document the influence of shifting the productivity parameters from the 2019 assessment to the current base model. Together with the run from Request 1 (using the 2019 assessment's historical catch reconstruction for Washington), the Panel concludes that changes in the historical catch were more influential than the changes in the population dynamics parameters on the perception of stock status.

Request No. 8: Provide estimates of equilibrium maximum sustainable yield for each step of the bridging analysis.

Rationale: The Panel would like to understand how each of the changes from the 2019 stock assessment contribute to the perception of productivity.

STAT Response: Updating catches had substantially more influence on equilibrium MSY than any other changes during the bridging analysis. About half of the change can be attributed to the changes to WA landings (a consistent, but larger change than that seen previously in Request No. 1). The remaining changes in the updated catches are due to updates made to discards in this assessment.

Table 1: The effects of the steps of the bridging analysis presented in the pre-STAR assessment report on equilibrium MSY.

Model run	Dead_Catch_MSY
2019 model	3,157
updated catches	2,691
no CPUE	2,688
single triennial	2,679
updated indices	2,654
updated comps	2,652
switched Ninput	2,641
annual model	2,656
re-weighted	2,408
blocks updated	2,523
SR updated	2,475
MG updated	2,485
maturity updated	2,501
fecundity updated	2,482

In addition to the modified catch histories for Washington state, the new assessment includes a more comprehensive amount of information on discards. In the 2019 model, discards were specified for four fleets and the most recent time block of the retention curve in two South fleets was only informed by a single year of data from the observer program (2002). The current base model aggregates the discards to two fleets and changes the timing of the block on retention asymptote for South fleets to rely on multiple years to inform historical discards, which led to lower estimated historical discard rates (and thus declines in total discard estimates). The post-STAR assessment document will include the total discards plots (not included here) to provide a more comprehensive summary of changes in total discard estimates between the two assessments.

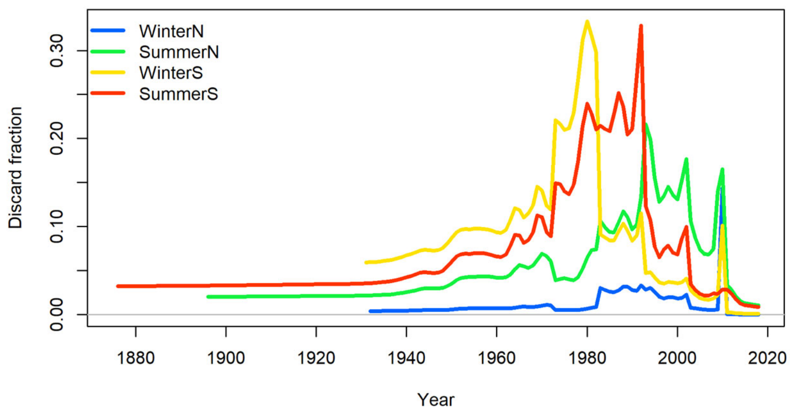


Figure 16. Discard fractions by fleet through time from the 2019 assessment.

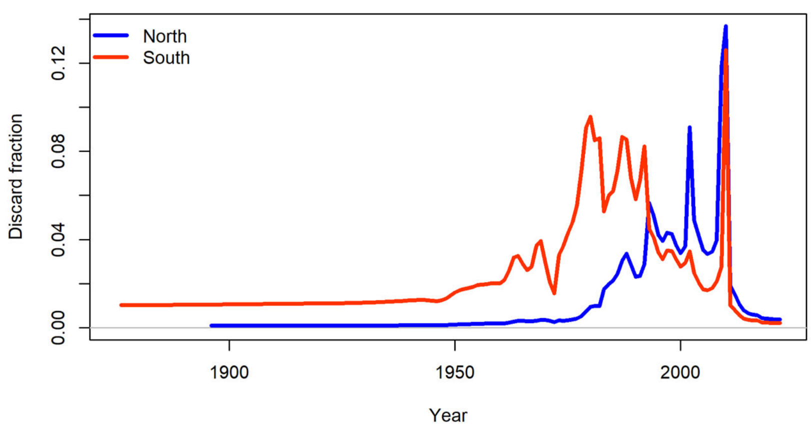


Figure 17. Discard fractions by fleet through time used in the current base model.

Panel Conclusion: The Panel agrees with the STAT's characterization of the contributing factors to the change to equilibrium MSY from the 2019 assessment to the base model. The current assessment has a more comprehensive analysis of discards and is an improvement over the previous assessment. Multiple contributing factors are reinforcing perceptions of stock productivity and the scale of the population. From this response, along with Request Nos. 1 and 7, the Panel concludes more of the changes in spawning output were attributed to catches than parameter changes (M and h). The most influential factors contributing to updated catches were the new WA catch reconstruction and changes to discard estimates.

Request No 9. Provide model runs that upweight (lambda = 10) a) ages from the south, b) lengths and ages from the south and c) ages from the north, and d) lengths and ages from the north.

Rationale: The Panel would like to better understand the influence of the recent observations of older than expected fish in the south, as well as a general sense of differences in dynamics regionally.

STAT Response: Upweighting the age data revealed tradeoffs between better fitting the age data with lower estimates of M (0.11 to 0.13, compared to 0.142 in the base model). The M profile in the assessment report shows that the age data are better fit with lower M values. The STAT hypothesizes that more older than expected fish in the age composition data are occurring because fish are older than they have been in a long time, which does not align with the M specified in the assessment that is required to balance all of the likelihood components. Spatial or temporal changes in growth could also explain this lack of fit, and these are explored in later requests.

The STAT provided diagnostic plots of fits to conditional age at length from the WCGBTS. These support the hypothesis that the recent apparent lack of fit may be best explained by undersampling older fish or lack of fit to the Von-Bertalanffy growth curve. Future research could further explore these differences in age data in recent years by fitting alternative growth curves (with more parameters).

The STAT also evaluated the model's sensitivity to removing the most recent 5 years of age data from the South (not requested). This had negligible influence on the spawning output or depletion time series, and as expected, improved the fits of the remaining age compositions in the South.

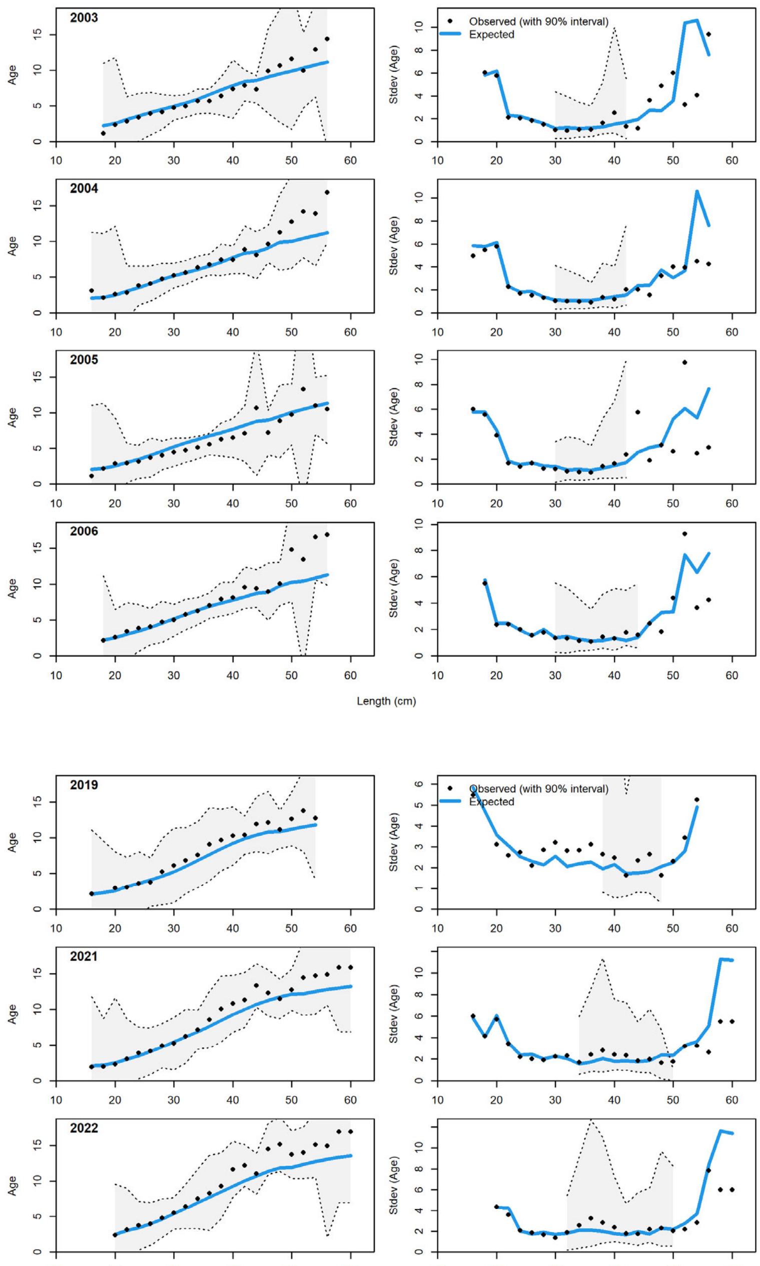


Figure 18: Conditional age at length for WCGBTS. Left column shows the model estimated mean age (blue line) +/- 90% confidence intervals (gray shaded region). Right column shows SE of mean age-at-length (obs. and exp.) with 90% CIs based on the chi-square distribution.

Panel Conclusion: The Panel agrees with the STAT's summary of the influence of the most recent age data from the South. More years of observations will help elucidate whether the most recent ages from the South are due to changes in the population or sampling. If time-varying growth is suspected, future work could investigate an empirical weight-at-age approach in this assessment as sample numbers increase. The tension between fitting ages vs lengths is relevant to many U.S. west coast groundfish stock assessments.

Declines in spawning output observed across many species during the 1970s-1990s can be attributed to fishing mortality and ecosystem variability. It would be reasonable to explore potential changes in natural mortality through time across species, as this could also contribute to the tension between the length and age data and growth.

Request No. 10: Provide model diagnostics (fit and observations of the conditional age at length data) of a run that evaluated separate growth for northern and southern areas with conditional age. For additional context, provide summary of Gertseva et al (2017) analyses of spatially varying growth for petrale sole.

Rationale: The Panel would like to better understand the spatially variable patterns inferred in the population dynamics, in particular the poor fits to the most recent years of the age data in the south.

STAT Response: An early exploratory model run separated the WCGBTS into northern and southern regions. The model had incomplete age data and estimated implausible differences in M by area and sex. However, the conditional age-at-length (CAAL) diagnostic plots are still useful to show there were no clear differences in growth between the regions.

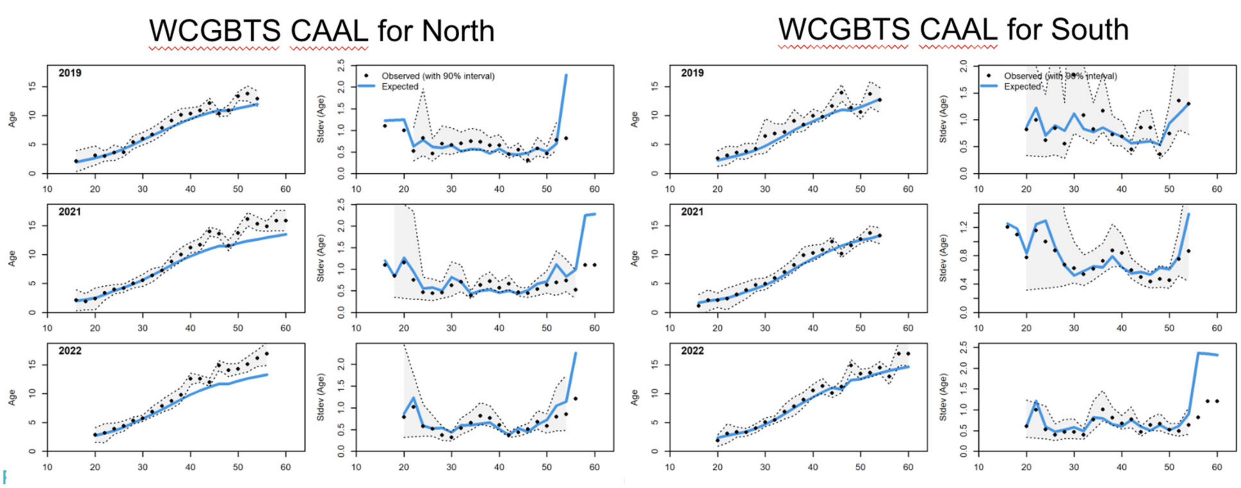


Figure 19. CAAL plots for a model run with WCGBTS split between North and South regions

Gertseva et al. (2017) showed no clear spatial trends or patterns in growth parameters by region, although faster growth was observed in the Eureka region.

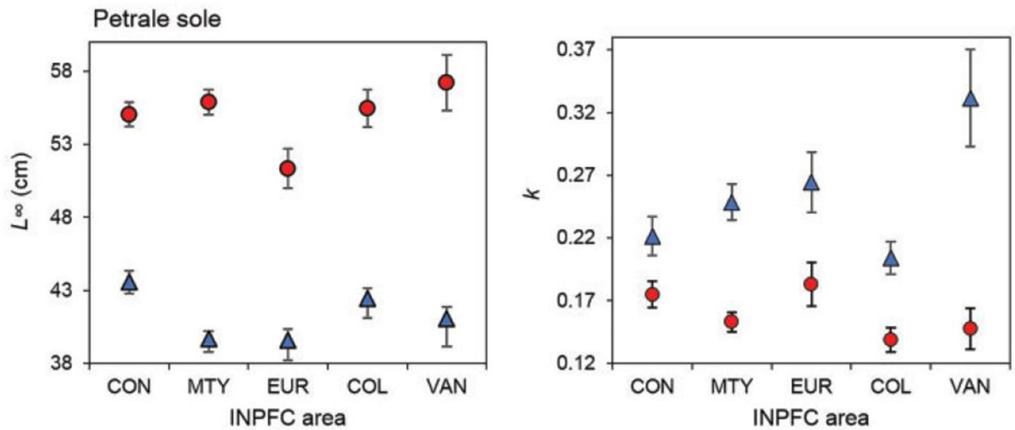


Figure 20. Regional petrale sole growth patterns, copied from Gertseva et al. (2017). Males are in blue, females are in red. Error bars represent +/-1 SD.

Panel Conclusion: The Panel agrees with the STATs summary of growth patterns and concludes there are no clear trends in growth apparent in these analyses for petrale sole that would explain the older than expected fish observed in the South. This should be investigated further in future research.

Request No. 11: Provide model results and projections using female M as the axis of uncertainty, and states of nature from the likelihood profile, using both methods described in the Terms of Reference (p12-13). Report both female and male M for each run.

Rationale: For constructing the decision table, natural mortality is likely the best axis of uncertainty. The Panel wants to ensure that the range of uncertainty resulting from the default approaches in the Terms of Reference adequately capture the uncertainty in the assessment.

STAT Response: Results and projections were provided as requested using both methods. The method bracketing uncertainty in spawning output results in states of nature with lower and higher M (0.072 and 0.219). The range of values is broader than the likelihood profile on M in the stock assessment report.

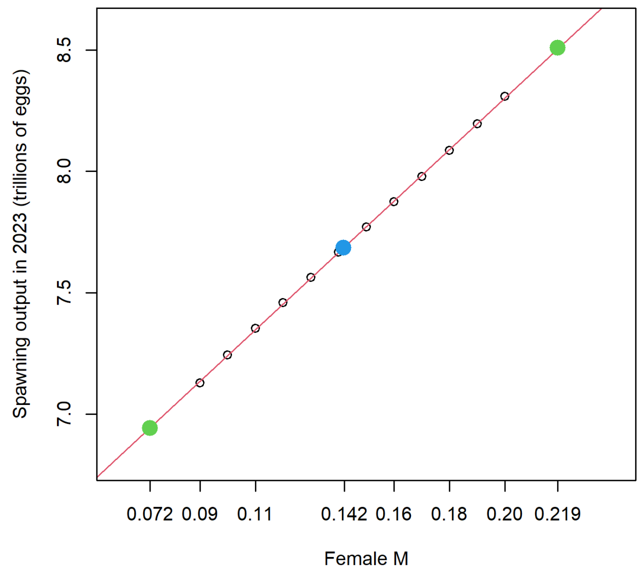
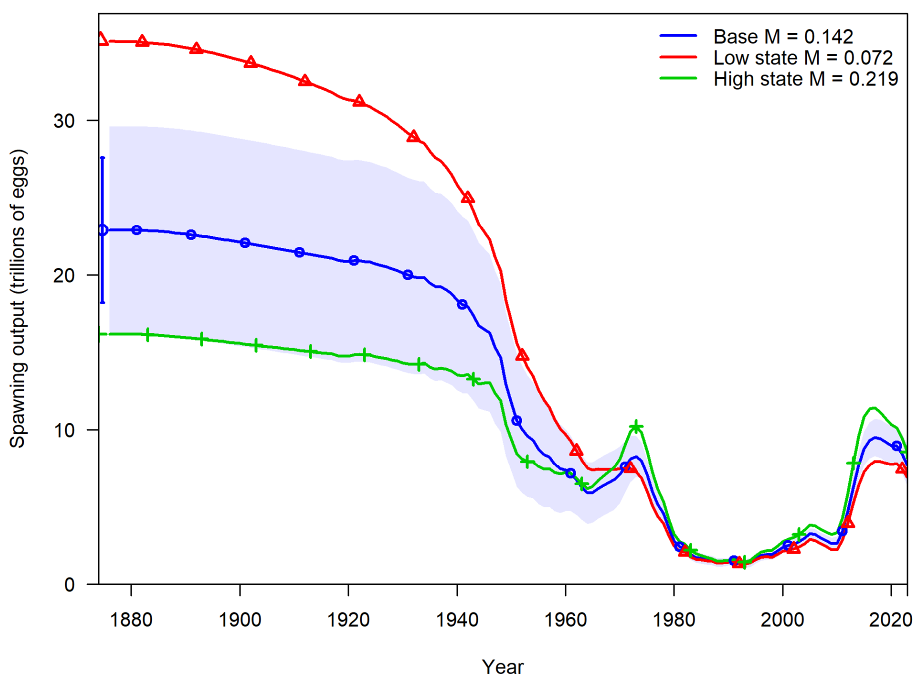
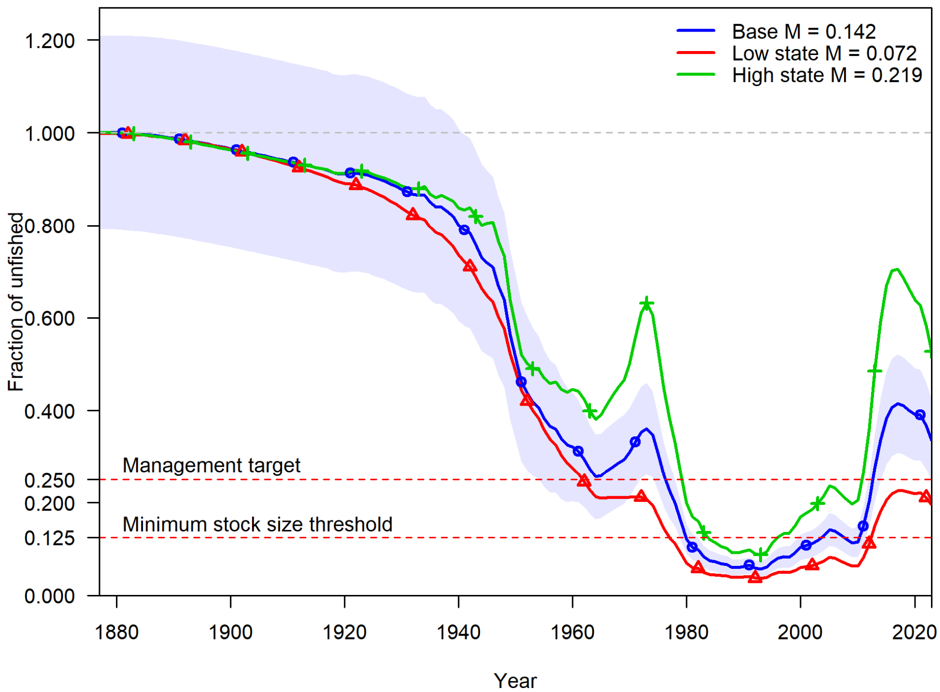


Figure 21. Female M associated with terminal year spawning output from the base model (blue dot), likelihood profile (black dots), and upper and lower bounds of 75 percent confidence intervals on spawning output (green dots).



Figures 22. Spawning output trajectories under high and low states of nature.



Figures 23: Depletion trajectories under high and low states of nature.

Panel Conclusion: The panel endorses the pre-STAR base model as the base model to recommend for informing management actions. Given the information in this model informing the likelihood profile on M, the panel concludes that the method using confidence intervals on terminal year spawning output to determine the high and low states of nature in the decision table is preferred for this assessment.

Request No. 12: Provide initial elements of a decision table based on the states of nature from Request No. 11 and the default harvest control rules, assuming a P* = 0.45. Also include an alternative of P* = 0.4. Provide a revised figure of the spawning output and depletion time series with corrected labels for M and high and low states of nature. Calculate the buffer from the OFL external to Stock Synthesis given the small discrepancies (< 1 percent) discovered in how the buffer is calculated within the software.

Rationale: These elements form an adequate range of options for the decision table that is a starting point for management consideration.

STAT Response: Projections were provided for P* = 0.45 and P*=0.4. An issue with Stock Synthesis calculations related to being near the ramp of the harvest control rule led to the output ACL not exactly matching the buffer*OFL. Therefore, projected catches were calculated using manual iterations of fixed catches in the forecast file. The projections in Table vii of the pre-STAR assessment document will be corrected in the post-STAR version of the assessment report.

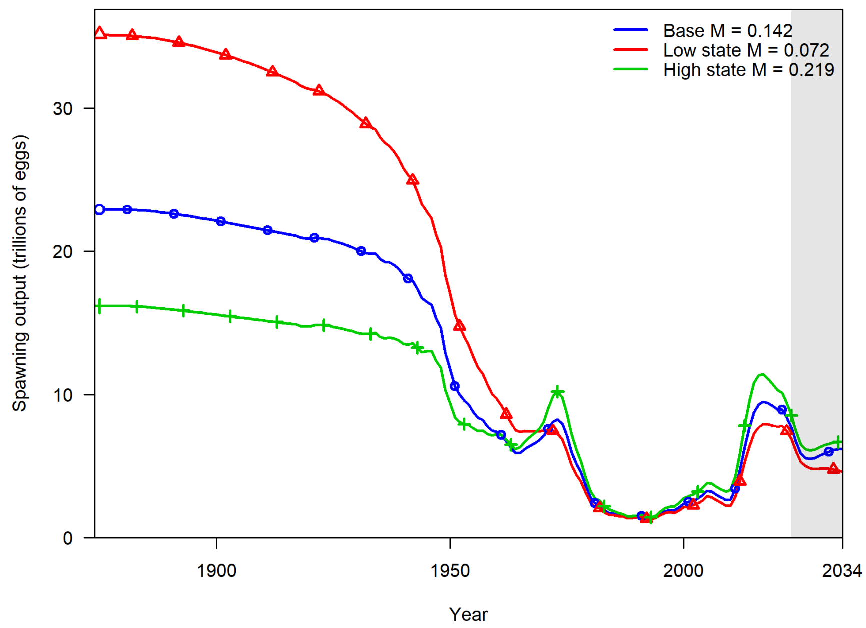


Figure 24: Spawning output for high and low states of nature, including the projection period

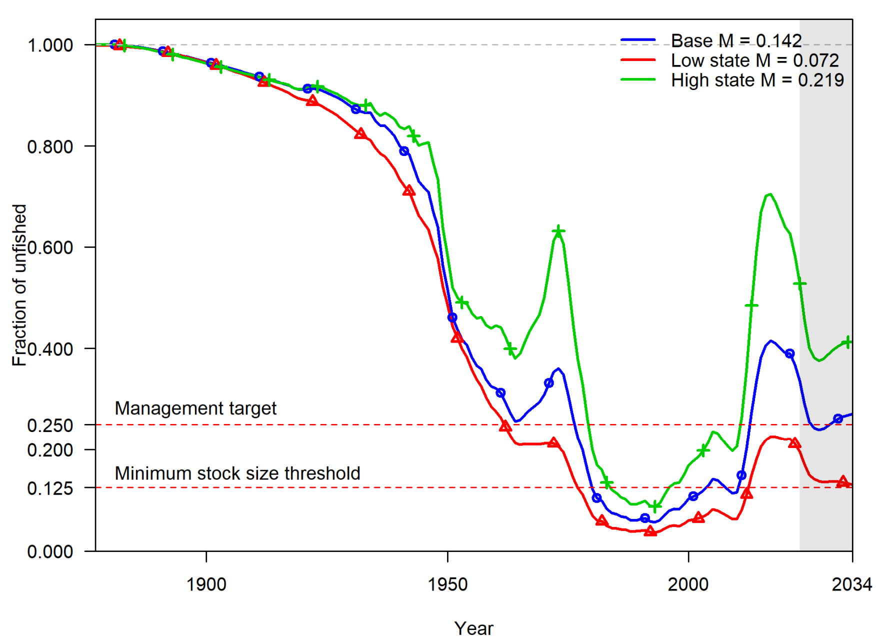


Figure 25: Depletion time series for high and low states of nature, including the projection period.

Table viii: Decision table with 10-year projections. 'Mgmt' refers to the three management scenarios (A) the default harvest control rule $P^* = 0.45$, (B) harvest control rule with a lower $P^* = 0.40$. In each case the 2023 and 2024 catches are fixed at the ACLs which have been set for that year with estimated fleet allocation provided by the GMT. The alternative states of nature ('Low', 'Base', and 'High') are provided in the columns, with Spawning Output ('Spawn', in trillions of eggs) and Fraction of unfished ('Frac') provided for each state. The colors of catch and fraction unfished are relative with lighter colors representing lower values.

Mgmt	Year	Catch	Low Spawn M=0.072	Low Frac M=0.072	Base Spawn M=0.142	Base Frac M=0.142	High Spawn M=0.219	High Frac M=0.219
A	2023	3485	6.86	0.195	7.69	0.336	8.53	0.528
	2024	3285	6.03	0.172	6.70	0.292	7.41	0.458
	2025	2354	5.27	0.15	5.85	0.255	6.49	0.401
	2026	2238	4.97	0.141	5.56	0.243	6.17	0.382
	2027	2217	4.83	0.137	5.48	0.239	6.08	0.376
	2028	2263	4.78	0.136	5.55	0.242	6.14	0.38
	2029	2334	4.79	0.136	5.69	0.248	6.28	0.388
	2030	2390	4.80	0.137	5.85	0.255	6.43	0.398
	2031	2429	4.79	0.136	5.99	0.261	6.55	0.405
	2032	2449	4.75	0.135	6.09	0.266	6.62	0.409
	2033	2460	4.68	0.133	6.16	0.269	6.67	0.412
	2034	2463	4.59	0.131	6.20	0.271	6.69	0.414
B	2023	3485	6.86	0.195	7.69	0.336	8.53	0.528
	2024	3285	6.03	0.172	6.70	0.292	7.41	0.458
	2025	2198	5.27	0.15	5.85	0.255	6.49	0.401
	2026	2117	5.05	0.144	5.63	0.246	6.24	0.386
	2027	2115	4.96	0.141	5.61	0.245	6.19	0.383
	2028	2169	4.96	0.141	5.72	0.25	6.29	0.389
	2029	2226	5.01	0.143	5.90	0.258	6.46	0.4
	2030	2279	5.07	0.144	6.09	0.266	6.63	0.41
	2031	2318	5.12	0.146	6.27	0.274	6.77	0.419
	2032	2345	5.13	0.146	6.41	0.28	6.88	0.425
	2033	2356	5.12	0.146	6.52	0.285	6.94	0.429
	2034	2360	5.08	0.145	6.60	0.288	6.98	0.432

Panel Conclusion: This provides sufficiently different biomass trajectories. These elements form an adequate range of options for the decision table that is a starting point for management consideration.

Description of the Base Model and Alternative Models used to Bracket Uncertainty

The base model was unchanged from the draft base assessment model described in the summary of data and assessment models.

The alternative models chosen to bracket uncertainty are based on alternative values of female natural mortality (M), as estimated based on the likelihood profiles using the methods described in the Terms of Reference. Natural mortality values in general had the greatest influence on the perception of stock status and productivity for this model. This approach led to female M values of 0.072 (low productivity state of nature) and 0.219 (high productivity state of nature), respectively. The associated ending year depletion estimates were 0.336 (above target level of 0.25) for the base model, 0.195 for the low productivity scenario (within the precautionary zone), and 0.528 (well above target levels) for the high productivity scenario.

Forecast trajectories using the default harvest control rule and a P*=0.45 led to continued declines in abundance for all three states of nature, based on the poor recruitments estimated by the model in recent years. Under the base model, the stock dips very slightly below target levels for several years, before increasing to slightly above target levels (as forecast recruitments from the spawner recruit curve inform projections). Under the high productivity state of nature the stock remains well above target levels despite the dip in the near term future. Under the low productivity state of nature the stock approaches, but does not dip below, the MSST based on the assumption of full attainment of catches from the default harvest control rule.

The Panel and the STAT spent considerable effort trying to understand the mechanisms responsible for the change in the perception of stock productivity relative to recent assessments. It would appear that changes in catch histories, specifically the historical catches for the Washington trawl fishery and the estimated discards for all fisheries throughout the 1970s and 1990s (prior to collection of robust discard data by the WCGOP) are the most influential changes in this shift in perception. However, changes to the treatment of both natural mortality and steepness are also contributing factors, as are many more subtle changes to the model structure, as illustrated by changes to the equilibrium MSY estimates associated with reweighting procedures in the bridging analysis.

Recommended sigma value and basis of recommendation

Model estimated uncertainty values for the 2023 spawning output and OFLs were very low (0.09 and 0.14, respectively), and considered to be implausibly low given some of the assumptions of fixed parameters (e.g., steepness) in the model. Consequently, the Panel recommends using the default sigma value of 0.5 for category 1 assessments.

Technical Merits of the Assessment

The assessment team invested considerable effort to simplify the assessment model, removing unnecessary complexity, such as seasonal structure, the treatment of the triennial survey index as two distinct indices, updating a broad suite of life history and reproductive ecology parameters, and removing datasets that were consistent with other results but less influential in the model itself, such as the trawl fishery-based CPUE index.

Historical catch estimates were revised for Washington state, and remained unchanged for the other states. The WCGBTS index appears to be highly informative for this stock, and provides a robust foundation upon which to base future assessments. Compositional data are robust for the surveys and the northern fishery, and are sufficient to reliably estimate year class strength and other changes in demographic structure.

The fishery independent abundance index from the WCGBTS is highly informative for this stock, and provides a robust source of both abundance and demographic information for informing the stock assessment.

Although considerable effort was made in developing an environmental index that would improve short term recruitment estimates, the STAT decided (and the Panel agreed) that the index is not sufficiently well developed to include in the current model. However, there is considerable promise for including this or comparable indices in future assessment models.

Technical Deficiencies of the Assessment

The Panel found no technical deficiencies within the data sources or modeling approach. There remains tension between length and age data in this model, which is typical of many U.S. West Coast groundfish stock assessments.

Areas of Disagreement Regarding STAR Panel Recommendations

Among STAR Panel members (including GAP, GMT, and PFMC representatives): There were no areas of disagreement between STAR Panel members and representatives regarding STAR Panel recommendations.

Between the STAR Panel and the STAT Team: There were no areas of disagreement between STAR Panel members and the STATs regarding STAR Panel recommendations.

Management, Data, or Fishery Issues raised by the GMT or GAP Representatives During the STAR Panel Meeting

The GMT initiated Request No. 5, which asked the STAT to evaluate the sensitivity to the 2018-2022 time block (started at 2018 rather than 2017), given that the management biennium began in 2017. The STAT's response was that the rationale for a 2018-2022 time block was to allow additional flexibility to the model to fit the data given that age compositions exhibited a change in shape since 2018, just shortly after several rockfish had been rebuilt, which led to the opening of trawl spatial closures. The GMT pointed out that trawl vessels targeting petrale sole with bottom trawl gear operate very differently spatially than trawl vessels targeting rockfish, and the more likely management measure change causing differences in length selectivity is that the use of selective flatfish trawl gear was no longer required starting in 2017. In addition, with several rockfish being rebuilt, bycatch constraints on vessels' IFQ quota were reduced. Those management changes were implemented at the start of 2017, whereas the time block chosen by the STAT starts in 2018. However, the sensitivity provided by the STAT indicated that the difference of one year in the selectivity block makes very little difference in the model output, and the STAT noted that fits to data were better when the time block starts at 2018. The GMT was satisfied with the STAT's response.

Unresolved Problems and Major Uncertainties

As with many groundfish stock assessments, the natural mortality rate and the stock-recruitment relationship (particularly the use of a fixed value for steepness) remain major sources of uncertainty in understanding the stock's current and potential productivity. Although the previous assessment estimated steepness (with a prior), the model estimated value for steepness when estimated as a sensitivity in this assessment was considered to be implausible due to approaching the bounds. The likelihood profile indicated that the data were not informative over a broad range of steepness values. The decision to fix steepness at the Myers et al. (1999) prior was thought to be an appropriate approach by the panel.

The Panel noted that estimates of historical discards changed fairly substantially between the current and the 2019 stock assessment models, as a result of combining summer and winter fleets in the current model. However, the magnitude of these changes and their influence on estimates on productivity were not initially fully understood. Additional evaluation of historical discard rates and the manner in which discards are modeled for the historical period may be appropriate for future modeling efforts.

Recommendations for Future Research and Data Collection

The STAT provided a list of six research and data collection needs, summarized below. The Panel supports the STAT's list. However, there are several additional research directions the Panel wishes to support which are added at the end of this section.

The development of environmental indices that could be used to better inform estimates of recruitment and cohort strength in recent model years that are otherwise poorly informed by survey or fisheries data was very promising. However, the Panel concurs with the STAT's conclusion that more validation would be helpful prior to formally including the index into the base model, particularly with respect to better understanding and evaluating the alternative oceanographic models that are necessary for contemporary data (relative to the models used in the initial analysis) and the potential to incorporate empirical data in the place of oceanographic model outputs. The Panel encourages continued development of this and/or related indices. The SSC should discuss the most appropriate means of reviewing and incorporating such indices into future assessments.

Pending (or complementing) the potential use of an index in a future assessment model, the Panel also encourages continued consideration of a risk table to inform managers with respect to environmental trends thought to influence petrale sole productivity but not yet formally included in a model, consistent with ongoing efforts of the PFMC's Ecosystem Working Group.

The Panel agrees with the STAT that additional research into both spatial and temporal variability into productivity processes such as growth, recruitment and maturity would help identify the extent to which such processes could or should be explicitly modeled in future assessments. This should include potential density dependent effects on growth and condition associated with variable abundance levels (e.g., as a result of strong recruitment pulses) or differential fishing mortality regionally (which could result in localized depletion or greater demographic heterogeneity across space). Such processes are likely to be contributing factors to tensions among age and length composition datasets.

The Panel encourages continued discussion and research with Canadian researchers to exchange data and ideas regarding index trends, demographic structure, movement patterns, dispersal, and recruitment dynamics between U.S. and Canadian petrale sole stocks, particularly as efforts are ongoing currently to develop a new Canadian stock assessment. Efforts to explore future transboundary stock assessments are also encouraged.

The Panel agreed that exploration of the mechanisms that could explain sex-specific differences in selectivity patterns would be helpful, such as sex-specific spatial distributions or behavioral patterns. Evaluating the potential for environmental sex determination would also be helpful.

The STAT noted that the analytical solution for catchability in the WCGBTS is considerably greater than 1 in the base model. The Panel agrees with the STAT that further research into the effects of herding or other responses to survey gear would help to understand the mechanisms behind the high estimated catchability levels implied by model fits to the WCGBTS abundance indices.

Additional Future Research and Data Collection Needs discussed by the Panel are as follows:

The Panel discussed possible mechanisms for the apparent pattern of above average abundance estimates for not only petrale sole, but many other flatfish and skate species in the final year (2004) of the triennial survey. One way to evaluate whether a contributing factor to this pattern is related to vessel effects or vessel skipper fishing practices could be to evaluate the location data associated with 2004 trawls relative to those from earlier years.

There remains a paucity of age data for the southern (California) area relative to northern fleets. In addition to better informing the model, age structure data from this region would improve the ability to better evaluate spatial differences in growth, productivity and population dynamics.

If available, historical age structures (otoliths) that were surface read and not used in this assessment should be read using contemporary methods to better inform population structure within the model.

The forecast values derived from Stock Synthesis suggested some minor discrepancies when control rule buffers were applied to OFLs to arrive at ACLs, such that some ACLs were greater than the ABCs after application of the harvest control rule. Improvements to Stock Synthesis to avoid this inconsistency should be implemented.

Recommendation for whether next assessment would be a full or update assessment and basis for recommendation and category

The Panel recommends that this be considered a Category 1b assessment, as there are robust compositional data to inform year class strength and growth, and there is robust trend information from surveys. The Panel recommends the next assessment be an update, unless an environmental index or other substantive change/improvement (such as a spatially structured model) were developed that would warrant a full assessment.

Acknowledgements

The STAR Panel thanks the Stock Assessment Team, the PFMC advisory body representatives and staff, and all of the meeting participants for contributing to a productive review.

References

Alverson, D.L. and Chatwin, B.M. 1957. Results from Tagging Experiments on a Spawning Stock of Petrale Sole, *Eopsetta jordani* (Lockington). Journal of the Fisheries Research Board Canada, 14(6) 953-974.

Haltuch, M.A., Tolimieri, N., Lee, Q. and Jacox, M.G., 2020. Oceanographic drivers of petrale sole recruitment in the California Current Ecosystem. Fisheries Oceanography, 29(2), pp.122-136.

Methot Jr, R.D. and Taylor, I.G., 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68(10), pp.1744-1760.

Methot Jr, R.D. and Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142, pp.86-99.

Ward, D., Robison, R., Nye, G., Reed, D. 1969. 1969 Fisheries Statistical Report. Washington Department of Fisheries.