Lingcod
Stock Assessment Review (STAR) Panel Report

Virtual Online Meeting
July 12-16, 2021

Participants

STAR Panel Members

John Field, National Marine Fisheries Service Southwest Fisheries Science Center (Chair)
Will White, Oregon State University
Matt Cieri, Center for Independent Experts
Cathy Dichmont, Center for Independent Experts

Stock Assessment Team (STAT) Members

Kelli Johnson, National Marine Fisheries Service Northwest Fisheries Science Center
Ian Taylor, National Marine Fisheries Service Northwest Fisheries Science Center
Brian J. Langseth, National Marine Fisheries Service Northwest Fisheries Science Center
Andi Stephens, National Marine Fisheries Service Northwest Fisheries Science Center
Laurel S. Lam, Pacific States Marine Fisheries Commission, Northwest Fisheries Science Center
Melissa H. Monk, National Marine Fisheries Service Northwest Fisheries Science Center
John E. Budrick, California Department of Fish and Wildlife
Melissa A. Haltuch, National Marine Fisheries Service Northwest Fisheries Science Center

STAR Panel Advisors

Lynn Mattes, Oregon Department of Fish and Wildlife, Groundfish Management Team representative
Gerry Richter, B&G Seafoods, Groundfish Advisory Subpanel representative
Todd Phillips, Pacific Fishery Management Council representative

Overview

Summary of Data and Assessment Models

Lingcod is a valuable target species for both commercial and recreational fisheries along the U.S. West Coast. The most important change from the previous assessment is that a new boundary was selected for the north and south assessments based on new genetic results pointing to separate stocks. Although a recent genetic study (Longo et al. 2020) show an area of admixture (a mixing of previously diverged or isolated genetic lineages) between stocks centered on Northern-central California (between approximately 38.6° and 39.5° N), Cape Mendocino was selected as the boundary, due to a combination of genetic, biological (size at age) and management considerations. New data from 2017-2020 were added and all data were re-analyzed, and there was a fresh look at selectivity and factors that affect it. Furthermore, based on the previous STAR panel recommendations, the model now estimates M and h. The STAT used several different diagnostic tests during model selection to develop the draft base case.

The northern boundary of the lingcod North stock assessment stopped at the USA-Canada border, despite the absence of evidence for a clear genetic break in the stock between those regions. The assessment in Canada has not recently been updated. For future research, the need to include more northern regions should be investigated. A first step would be to compare data between Canada and USA to see whether there are sufficient data to justify a joint stock assessment. Information on the Mexican fisheries show that groundfish are not targeted and treated as a bycatch, often from more artisanal fleets. The available data is therefore not of the format and quality for use in future assessments.

It is important to note that this change in the stock boundary means that the 2021 assessment results are not directly comparable to the previous assessment. As a result, no bridging analyses between the present and 2017 assessment were provided.

Lingcod have faster dynamics than many of the groundfish on the US west coast, specifically demonstrating faster growth, earlier maturation and fewer older individuals observed in the population due to higher natural mortality (M). There are sex-specific differences in both behavior and spatial distribution, including by depth and season. Additional ageing done for the 2021 assessment made it possible for new age-based maturity curves to be developed, which shows that the maturity at age curves between males and females are fairly similar compared to the maturity at length. The core biological samples in the North and South assessment come from WCGBTS, although no survey was undertaken in 2020. Lingcod can be found shallower than the survey minimum depth of 55m.

The northern model data is characterized by containing longer time series and more data inputs. The CA recreational data were new to the northern model because of the boundary change. The southern model had fewer data inputs, and abundance indices were generally relatively short and/or did not cover the entire range of the stock (other than the WCGBTS). The south model also only has a single source of conditional age-at-length (CAAL) compositions, from the WCGBTS (sparse age data from fisheries were excluded). Commercial landings data for both assessments were updated and reconstructed further back in time compared to the previous assessment, and now start in 1889, an important improvement. These were divided into two fisheries, trawl and fixed gear. Missing data were interpolated between years. Recreational fisheries data were also updated.

The commercial trawl fishery index was analyzed similarly to the previous assessment by applying the vector-autoregressive spatio-temporal (VAST) model of Thorston and Barnett (2016). Only the new border changed in the analyses. The recreational abundance indices available to each stock were re-analyzed with the updated data using a consistently applied Bayesian delta-GLM modelling framework. For the Northern stock these included the Oregon nearshore index, the Washington recreational index, Oregon recreational boat survey index, Oregon at sea CPFV index, and the CA MRFSS Dockside CPFV Index. The South model recreational indices available included the MRFSS Dockside CPFV Index, the Californian CPFV Onboard Observer Survey (not used, survey also not undertaken in 2020), and the California (Deb Wilson-Vanderberg) CPFV Onboard Observer Index (1987-1998). The commercial fixed gear index was not available for the South. Both the northern and southern models used commercial trawl indices developed in the 2017 assessment (unchanged from that assessment).

Fishery independent surveys included the NWFSC West Coast Groundfish Bottom Trawl Survey (WCGBTS) and the AFSC/NWFSC West Coast Triennial Shelf Survey (also a bottom trawl survey). Indices from both surveys were developed for each model using the VAST model. An index based on the NWFSC Hook and Line Survey in the Southern California Bight was developed using a Bayesian delta-GLM model and included in the southern model. An index based on the California Cooperative Fisheries Research Survey (CCFRP), a hook and line survey that takes place both inside and outside of the California Marine Protected Area network was developed but ultimately not used in the model.

Length compositions and CAAL compositions are available for both commercial fleets; and the Washington, Oregon, and Californian recreational fleets (noting that there are no age data for the CA recreational fleets). The latter data are appropriately allocated to the northern and southern stocks. Commercial trawl length-composition data started in 1965 in the north and 1977 in the south. Data collection for the commercial FG in the north began in 1986 (with additional samples in 1971, 1980 and 1982), and began in 1992 in the south (with additional samples in 1988 and 1989). There are three sets of length compositions per fleet, female, male and unsexed and two sets of CAAL per fleet, male and female. This structure is unlike the previous draft base assessment where unsexed individuals were split to sex assuming a 50:50 ratio irrespective of length.

Discard rates were modelled for the commercial data only using data from the West Coast Groundfish Observer Program (WCGOP), which started in 2002. Observer coverage rates were relatively low in the early period but increased to nearly 100% for all the limited entry vessels since 2011. A single annual discard rate was calculated for each fleet per year. For the recreational data, catch is input as for retained catch and the estimated dead discard catches (an assumed 7% mortality rate from all discards) combined.

The previous assessment used the maximum age seen in the data as a basis for M, which was 21 for both males and females. This draft base assessment instead used the 99^{th} percentile age observed in the data, which was 18 and 13 for females and males respectively. These maximum ages are very similar for DFO Canada estimates of 20 and 14 respectively, although in Alaska maximum ages of up to 36 have been observed. The Hamel-Then prior for natural mortality was used in the assessment for each sex but applied across all ages. Prior means were 0.3 for females and 0.415 for males. This was a significant increase from the previous assessment estimates of 0.257 used for both sexes (as fixed values). The base model estimated natural mortality rates were 0.42 and 0.41 for females and males in the northern model; for the southern model the estimates were 0.17 and 0.22 for females and males respectively. The unusual differences in the estimates of natural mortality by region were discussed extensively during the week.

The 2017 STAR panel recommendations included cross-validating age-reading among laboratories and years. The STAT's investigations found that there was little ageing error – ages were similar among readers and between years.

Mean weight-at-age from WCGBT survey were analyzed to track a single age over time to investigate the possibility of time-varying growth. The younger ages are quite flat, whereas mean weights at ages 5 and 6 show signs of changes over time. However, these potential changes need to be considered in conjunction with the larger standard errors associated with mean weight. These data do not show a strong case for time-varying growth, but the possibility could not be excluded.

The STAT utilized three primary diagnostic tests while developing the base models, including likelihood profiles to highlight conflicts within the data, using the SS3 jitter function to ensure that models had converged to the true minimum likelihood, and an evaluation of the correlations among Bayesian posteriors (although this model is based on Maximum Likelihood Estimation, the STAT was able to do Bayesian analyses on the posterior estimates using MCMC in order to identify "inefficient" parameters). It is with the above diagnostic tests that it became clear that there was strong autocorrelation among several of the parameters, particularly among a number of the selectivity parameters. Several of the selectivity parameters were subsequently fixed in the base models.

Requests by the STAR Panel and Responses by the STAT

Request No. 1: Explore freeing up the width parameter on sex specific selectivity curves, including offsets, for both North trawl and south fixed gear. Do not have this time varying. This for both North and South. Provide likelihood profiles on M for females 0.18 to 0.58 in steps of 0.05.

Rationale: The Panel noted that there is tension between age and length. There is a depth gradient by sex that would suggest sex specific selectivity. It may help understand the tension between length and age composition data.

STAT Response: The revised analysis did not lead to large differences in selectivity curves. However allowing sex-specific offsets in selectivity did somewhat relax the tension between the age and length compositions in their influence on estimating M, as reflected in a flatter likelihood profile on the age data. Yet the model also produces some biologically questionable results: a maximum selectivity for females in the commercial fleet that is only 20% that of males, and a natural mortality rate for males that was ⅔ greater than females in the North. This modeling option also led to very strong influence of the Oregon recreational data, which could not be easily explained.

Panel Conclusions: The panel concluded that pursuing sex-specific selectivity offsets was a valuable avenue to further explore how to resolve tension between data sources in the model (particularly in the North) but that this option as implemented here could not be supported as an addition to the base model because of the problematic.

Request No. 2: Show the model run and the fit to the DWV index, when that index has now influence in the model. Compare that to the fit that includes the index. Provide model comparison figures. This is only for the southern stock.

Rationale: The DWV index had unusually high sensitivity without a clear reason why.

STAT Response: Removing either the index alone or both index and associated length data had similar effects, which was to produce a current estimate of stock status that is greater than B0. With those data removed, fits to other indices did not change substantially, while the projected fit to the DWV index was much flatter and did not capture a sharp decline in the late 1980s that is present in that index.

Panel Conclusions: Removing the dataset led to substantial increases in the current estimate of stock status. This appears to be due to a signal of a sharply declining stock in the late 1980s/early 1990s that does not appear in other datasets from that period. The influence of that period is apparent when the DWV index is excluded from the likelihood calculation; the revised model fits the later portion of the DWV index well but does not track the late 1980s decline. The panel discussed the spatial coverage and representativeness of the DWV index (as described in Monk et al. 2016 technical report) and concluded that it was a robust dataset that should remain in the model, and that its high influence on the outcome was not likely to be spurious.

Request No. 3: Run the southern model with the main recruitment deviations starting in 1955. Produce model comparison results and dynamic B_0 for both models. If the early recruitment deviations are unusually too high (comparable to what we see in the base model at present, turn off the early recruitment deviations for a 3rd run if time allows.

Rationale: There is an unexplained and implausible run-up of recdevs in the pre-data period that may be a model artifact.

STAT Response: Starting main recruitment deviations in 1955 did not produce a large change to the model fit, nor did it eliminate the questionable early deviations. Eliminating early recruitment deviations did not produce a large change in later model fit but did produce a higher estimate of B_0. As an additional analysis, completely excluding the recreational length data from 1959-1972 did eliminate the questionable recruitment deviations from the 1950s, and did not lead to a large change in current estimated stock status.

Panel Conclusions: Changes to the way recruitment deviations were handled did not resolve the concern about unusual early recruitment events and excursions well above B_0. Removing pre-1975 length composition data did eliminate the questionable recruitment patterns in the 1950s but it is undesirable to completely exclude those data. Those pre-1975 recreational length data were entirely from the Monterey Bay area, so an alternative solution is to add a selectivity block for that time period for that fleet. This will be the rationale for Request 6, below.

Request No. 4: Provide removals (i.e., total dead discards and total retained catch) by fleet and disposition by year for both North and South. Stat discretion on display of data-figure or table.

Rationale: This is not currently displayed in the draft reports as a single informational graphic.

STAT Response: The figures have been provided. Of note in these data is a steady decline in landings in the trawl fleets in both regions in the late 1990s and early 2000s, as well as an extremely high (~90%) discard rate in the North. This likely reflects management changes during that period, when the stock was in a rebuilding phase.

Panel Conclusions: The immediate onset of very high discards in the 1990s, coupled with the absence of discards (e.g., of smaller fish) in prior years is very suspect. It is likely that the model is overestimating the discard mortality in that period. There is not a clear path to resolving this issue for the current assessment, but future assessments should explore a reconstruction of historical (pre-1999) discards, or consider allowing a ramp-up of discard rates to the 1999 period. It would be valuable for future versions of Stock Synthesis to allow separate weighting of discarded and retained catches.

Request No. 5: Update the northern model to reflect the error in the WA recreational mortality estimates. Provide a figure of WA dead discard catch and retained recreational catch.

Rationale: An error was noted in the report regarding the WA recreational mortality estimates.

STAT Response: This error did increase the estimate of WA recreational landings by a considerable amount (50+ tons a year in many years), so this was an important correction. A figure of the relative landed catch compared to estimated dead discards suggested that the total dead discards represent a fairly modest fraction of the total recreational catch. The corrected catch model trajectory resulted in very little change in the overall model result.

Panel Conclusions: There was agreement among the STAT and the Panel that this change was made and reflects the new base North model (day 2 North base).

Request No. 6: Add time varying selectivity for the CA rec fishery in the South model. Add a block for 1959 to 1972. Provide the model results (biomass, recruitments, and recdevs) and the estimated selectivity curves. Provide information on how the fit to the data improved.

Rationale: STAR panel understands the 1959 to 1972 data encompasses a small amount of CA coastline and may not be representative of the entire area

STAT Response: The STAT divided the early period of the fishery into two blocks, one from the model start to 1972, and a second block from 1973-1982 (there was an existing time block beginning in 1982). Adding the additional block had very little effect on the model results (spawning output, recruitment estimates). The new selectivity curves suggest that the very earliest years had selectivity shifted slightly to the right, relative to the 1972-1981 selectivity. The likelihood values suggest little change in fit. In discussing the data more, and in recognition that the length data were only collected for a small portion of the fleet, and that the recruitment pattern included strong autocorrelation and an "unsettling" dynamic B_0 pattern (request 3), there was additional discussion about removing the early length frequency data. Without those data, changes in key parameters included a shift in female M to 0.17, male M to 0.23; steepness increased slightly to 0.55.

Panel Conclusions: Removing these from the model seemed to be the most satisfying solution for both the STAT and the Panel. The Panel recommends that future research would be appropriate to compare the regional differences between more contemporary data with the region in which these historical data were collected, or to otherwise to discern if there could be a more appropriate way to use the data in future models. This would become the new South base model (pending request #9).

Request No 7. Perform a model run, for the status quo Base model only, for the North without the fishery age data and rerun with Dirichlet Multinomial (DM) weighting. Provide initial results and comparison of Francis vs Dirichlet weights.

Rationale: Dirichlet weighting seemed to resolve some of the tension between age and length data, and estimated the natural mortality rate more comparable to the South model.

STAT Response: The change in weighting did not appear to be as great as the Panel expected. The DM weighting did lead to fully weighting most age composition data, which led to a scaling down of the model spawning output and a slight decrease in the estimate of M. Without fishery dependent age data biomass scales up (with or without DM) and the estimated M is smaller still (from 0.41 in the base, to 0.37 to the base with DM, to 0.27 to the base with DM and no fishery age data). Without age data and with DM weighting, a very strong year class shows up in 1962 that appeared to be potentially spurious, suggesting that age data were not the only data creating interesting signals in recruitment. The natural mortality estimate in the northern model declined substantially without age data.

Panel Conclusions: The take home message from this analysis seems to be that there will continue to be tension between the length and the ages at this stage in the model. The DM method puts much more weight on both length and age relative to the indices. There was general agreement that, at this stage, recommending a switch to DM weighting was infeasible**.**

Request No 8. A) Remove all fixed gear (FG) age composition data. B) Remove fixed gear age composition data for the years 1999-2011. C) Do A and B with and without offsets for sex-based selectivity for all fleets. Requests are only for the North model. Provide the model comparisons slides and table(s) with likelihoods and key parameter outputs, selectivity outputs, etc. The STAT will include the updated results from Day 2, Request 5.

Rationale: These model runs are requested to better understand whether management changes or unbalanced periods of age sampling data in the associated fisheries are contributing disproportionally to tension in the model. There is an observed disconnect between the fixed gear length and Conditional Age-at-Length (CAAL) in the 2000s that are more likely to reflect fleet behavior.

STAT Response: These analyses were all undertaken with the new north base model. Removing the FG ages of the years 1999-2011 made no difference and results were similar to the base case. However, removing all the FG ages had a large impact on the results by increasing the spawning biomass (although there was not a large impact on the fraction of unfished trends relative to the new base model). These results highlight that the source of tension in the model is not from the period 1999-2011 but more likely the earlier data where there are fewer alternative sources of information.

When these sensitivity tests were combined with a sex selectivity offset (whether with all FG ages removed or only those between 1999-2011) then the spawning stock scale was similar to the new base model, but there was a large change to the pattern in the fraction unfished. In terms of the recruitment results, these were quite similar to the new base model and therefore highlight that most of the information on recruitment are being provided by other sources of age data and not the FG.

The changes to the likelihood and parameters table shows that the male and female M values are unchanged with the removal of all or some of the FG ages, but as seen before, sex selectivity lowers the female M and slightly raises the male M. Since the length composition data are consistent with these runs, the length composition likelihood fits are better with the removal of the FG ages, and even better with sex-specific selectivity. Removal of the FG age data, also improves the fit to the survey data whereas the female selectivity offset fits the survey data worse compared to the new base model.

An additional sensitivity run was provided by the STAT team, which is to remove all fishery ages prior to 1990 (therefore from the trawl, FG and WA recreational ages) and all prior to 2000. This was due to the above runs pointing to a more general issue with the early age data, especially in the context of the potential for time-varying growth. If this had occurred, then ages and length over a long period in a model would create model conflict. Although the removal of these data changes the absolute spawning biomass values, there is little change to the recent fraction of unfished trends compared to fairly large differences in the 1960s and 1970s. The early recruitments have been informed by the data that has been removed. All these sensitivity runs provide similar M values (above 0.4) and do not point to any major issues that should be addressed.

Early runs of the model (not provided) applied age-varying M through applying the Lorenzon M (these runs were not presented but discussed here in principle). This approach assumes that M varies over the life of a fish through various processes and assumes that M is inversely proportional to length. The early results for the North model were reasonable but not for the South model, where the results were sensitive to the reference age. However, it should be noted that the South model does not have any age data so this assumption would be hard to fully implement in the South.

Panel Conclusions: The implementation of the female offset could be improved and further refined. The data suggests that sex-based selectivity may be occurring, but the model runs with the offset included do not provide results that are entirely plausible. The panel agrees with the STAT team regarding suggestions that it is possible to improve this issue by applying the offset only to larger animals, rather than the present approach where the offset is applied to all sizes and ages. Alternatively, one could also adjust the approach to how M is being estimated since the processes of age-varying selectivity would be confounded with age-varying M.

Although not possible to suggest for this assessment, including smaller unsexed fish into the CAAL data could improve the potential conflict between growth, M, and selectivity by providing more information on growth of the smaller length and ages. Smaller, younger fish are more difficult to sex so many of these would not be in CAAL. If the STAT continues investigating age-varying processes in the future, then it would be useful to include the unsexed individuals in the CAAL data (as unsexed). This panel's conclusions turned into 2 further requests.

Request No 9. Conduct additional tuning and exploration of the southern model without the 1959-1972 length composition data to see if a reasonable new base model can be developed. Provide the model comparisons slides and table(s) with likelihoods and key parameter outputs, selectivity outputs, etc.

Rationale: Conduct additional tuning and exploration of the southern model without the 1959-1972 length composition data to see if a reasonable new base model can be developed.

STAT Response: The results of the initial test of the Southern model without the early length composition data tuned and untuned were provided. The tuned model with no early lengths required adjustment to the mean recruitment deviation period and redoing the Francis weighting (as expected). Poor behavior of the Triennial survey’s descending selectivity parameter was resolved by adding additional error to the Triennial index. The tuned model affected the overall trend and the fit to the Triennial survey index. The uncertainty of the “All lengths” (e.g., including the 1959-1972 length data) model is smaller than the runs without the early lengths. The early period is quite uncertain for the new runs, but more similar to the base case in recent years. The fraction of unfished in 2020 of the re-tuned model is at the management target, but this is lower than the other two runs. Removing the early length data resolved the STAR panel’s issue of the large autocorrelations in the recruitment deviations. In terms of likelihood components this model fits the WCGBTS index better than the draft base case which, in terms of lingcod, is a more reliable index due to its more consistent spatial coverage. None of the models fit the last year of the WCGBTS well. All the three runs have similar steepness values, but the M values are markedly different with the new tuned model being lower.

Panel Conclusions: The panel recommended that the tuned version of the model that removes the length composition data from 1959-1972 should be the new base case in light of the above findings. The STAT agreed. Additional exploration of the historical length frequency data should be a research need for future efforts.

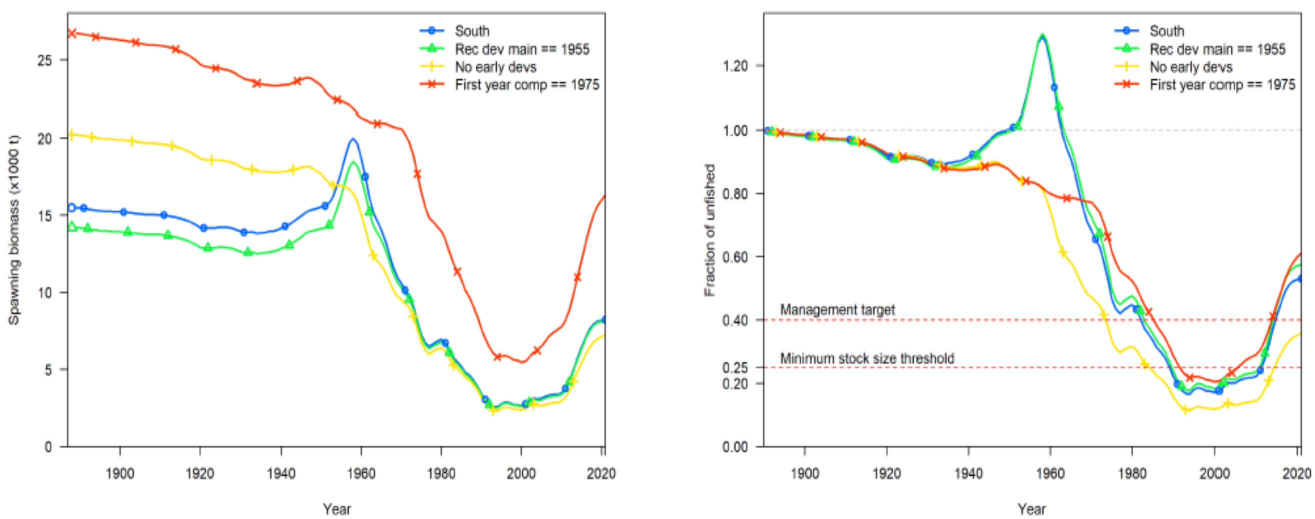


Figure 1: Initial draft base model relative to models arising from request #9.

Request No 10. Perform runs for both North and South base models with the Lorenzen mortality function. The reference age for the naturality mortality prior should be age eight (8). Provide the model comparisons slides and tables/figures with likelihoods and key parameter outputs, selectivity outputs, etc. for panel review.

Rationale: Allowing for a variable natural mortality with age for fish less than age 8 would be a more realistic estimate of natural morality for smaller and immature fish. It is well documented that natural mortality tends to be negatively correlated with body size.

STAT Response: The STAT completed the runs as requested for both North and South. These included runs in which the Lorenzen function was modeled alone as well with the sex selectivity offset in Request number 11. Results are shown in the figures and tables below.

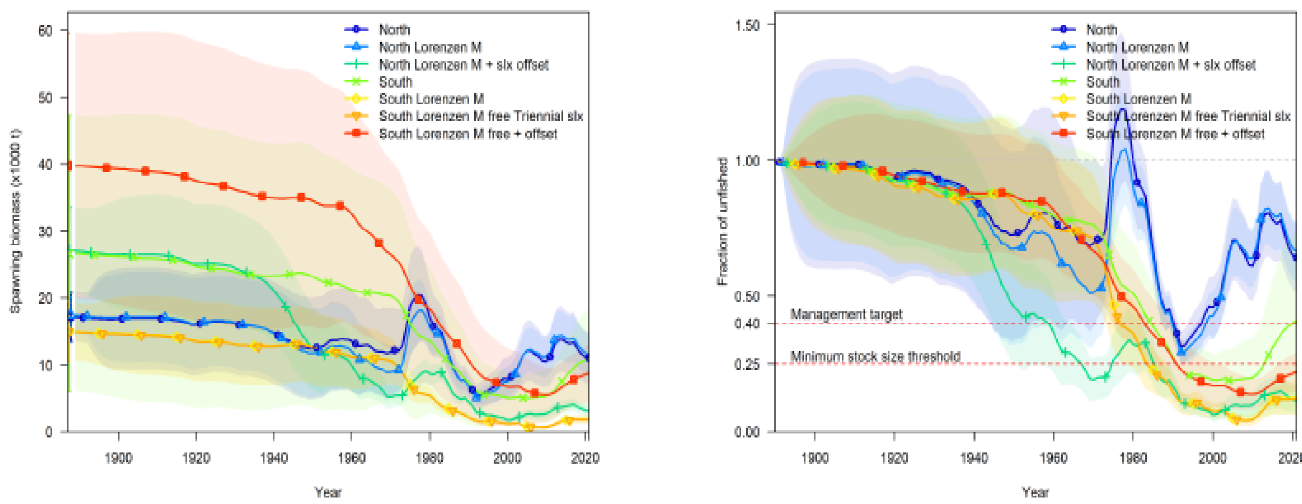


Figure 2. Comparison between using the Lorenzen mortality function, with and without sex selectivity offset, for Northern and Southern stocks compared to the base run.

Table 1. Likelihood table for the Lorenzen mortality exploration for the Northern Stock

Label	North	North.Lorenzen.M	North.Lorenzen.M.slx.offset
Diff. in likelihood from base model			
Total	0	-105.8	-421.69
Survey	0	1.54	24.62
Length comp	0	-66.84	-312.37
Age comp	0	-39.59	-141.21
Discard	0	-1.38	-2.54
Parm priors	0	-0.17	1.39
Estimates of key parameters			
Recr Virgin millions	16.73	251.83	6.28
SR BH steep	0.8	0.83	0.8
M Female	0.42	0.35	0.16
M Male	0.41	0.5	0.25
Estimates of derived quantities			
SSB Virgin 1000 mt	17.16	17.43	27.06
SSB 2021 1000 mt	11.01	11.59	3.12
Fraction unfished 2021	0.64	0.67	0.12
Fishing intensity 2020	0.26	0.3	1.31
Retained Catch MSY mt	3937.89	4359.94	1647.47
Dead Catch MSY mt	4222.53	4690.44	1686.96
Virgin age 3+ bio 1000 mt	32.69	35.1	33.76
OFL mt 2021	5084.77	4597.37	444.69

Table 2. Likelihood table for the Lorenzen mortality exploration for the Southern Stock

Label	South	South.Lorenzen.M	South.Lorenzen.M.free.Triennial.slx	South.Lorenzen.M.free.offset
Diff. in likelihood from base model				
Total	0	4.85	-2.43	338.78
Survey	0	-3.66	-3.09	12.14
Length comp	0	10.06	4.01	-34.74
Age comp	0	-7.02	-8.72	348.3
Discard	0	0.57	-0.05	1.51
Parm priors	0	-0.75	-0.77	-0.3
Estimates of key parameters				
Recr Virgin millions	2.25	4.36	4.34	3.65
SR BH steep	0.51	0.58	0.58	0.56
M Female	0.17	0.17	0.17	0.14
M Male	0.22	0.21	0.21	0.25
Estimates of derived quantities				
SSB Virgin 1000 mt	26.66	14.95	14.85	39.75
SSB 2021 1000 mt	10.88	1.88	1.8	8.77
Fraction unfished 2021	0.41	0.13	0.12	0.22
Fishing intensity 2020	0.34	0.89	0.91	0.56
Retained Catch MSY mt	843.96	901.9	902.74	762.47
Dead Catch MSY mt	849.07	905.87	906.63	766.67
Virgin age 3+ bio 1000 mt	32.8	21.91	21.81	43.38

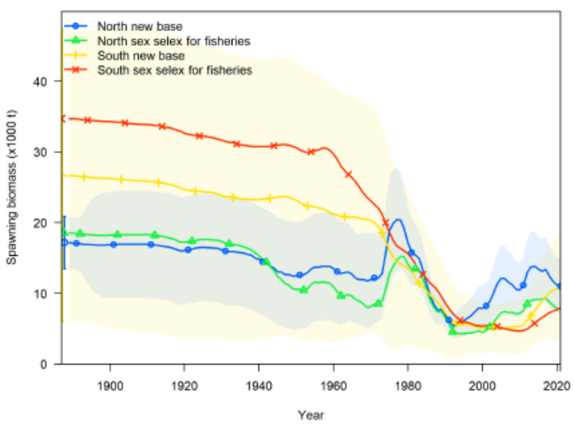
Overall, the STAT indicated that using the Lorenzen function did not dramatically change the trajectory of the SSB for either Northern or Southern stocks and had less effect than the use of a sex selectivity offset. It was noted that the South model was very sensitive to this type of natural mortality function, and gave unrealistic results, when changing the reference age of the mortality prior.

Panel Conclusions: The Panel agreed with the STAT that there was little change with the incorporation of the Lorenzen mortality function. They also recognized that in general fits to the age and length composition data improved, but were somewhat offset by the reduced fit to the indices. Overall, the change to using a sex selectivity offset dwarfed any change invoked by using the Lorenzen. The additional sensitivity of what age the prior was placed at for the South indicated that using the Lorenzen mortality function was not a fruitful approach at this juncture. While not useful for this particular workshop, natural mortality and its treatment was considered an important research recommendation.

Request No 11. Perform runs for both the North and South base models in which sex-specific selectivity offsets are implemented for fisheries but not surveys. Explore other parameterizations of the selectivity offset as time might allow. Provide the model comparisons slides and tables/figures with likelihoods and key parameter outputs, selectivity outputs, etc. for panel review.

Rationale: The STAT has observed that the selectivity differences appear to be greater for the fisheries data relative to the survey data, and has been considering alternative forms of implementing the selectivity offsets.

STAT Response: The STAT completed the response as requested. Overall, the model showed a lower scale in SSB after approximately 1940 and a lower stock size relative to management reference points (Figure 3) for the North. For the South, the difference in scale was prior to the mid 1970's, but also showed a lower stock size relative to the reference points. While the selectivity pattern has improved and the fit to the data improved substantially, the STAT was not convinced that the results led to an improved base model, particularly as the fits to the indices degraded (Figure 3). The estimated selectivity offsets were substantial, particularly in the fixed gear and recreational fisheries (Figure 4).



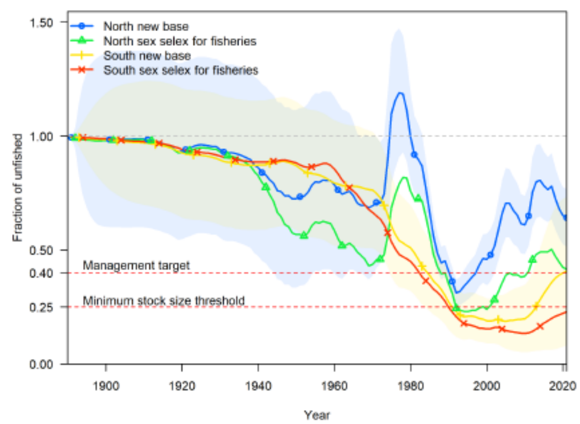


Figure 3. Spawning Stock Biomass in both Northern and Southern stocks using new base runs and comparing the use of a sex selectivity offset for males and females.

Panel Conclusions: The Panel agreed with the conclusions of the STAT. They also noted that while this change improved the overall fit in the North, it worsened it in the South. Likewise, while there was an improvement in the overall fit, and the fits to both length and age comps, it came at the expense of the survey fits to the abundance. The Panel also agreed that the selectivities in both the North and the South were better, but still rather implausible. In general, the Panel came to consensus that while this change improved the model, it was only a partial improvement. Given this result, the Panel suggested Request 12 as a way of more fully exploring sex selectivity difference between males and females in the model.

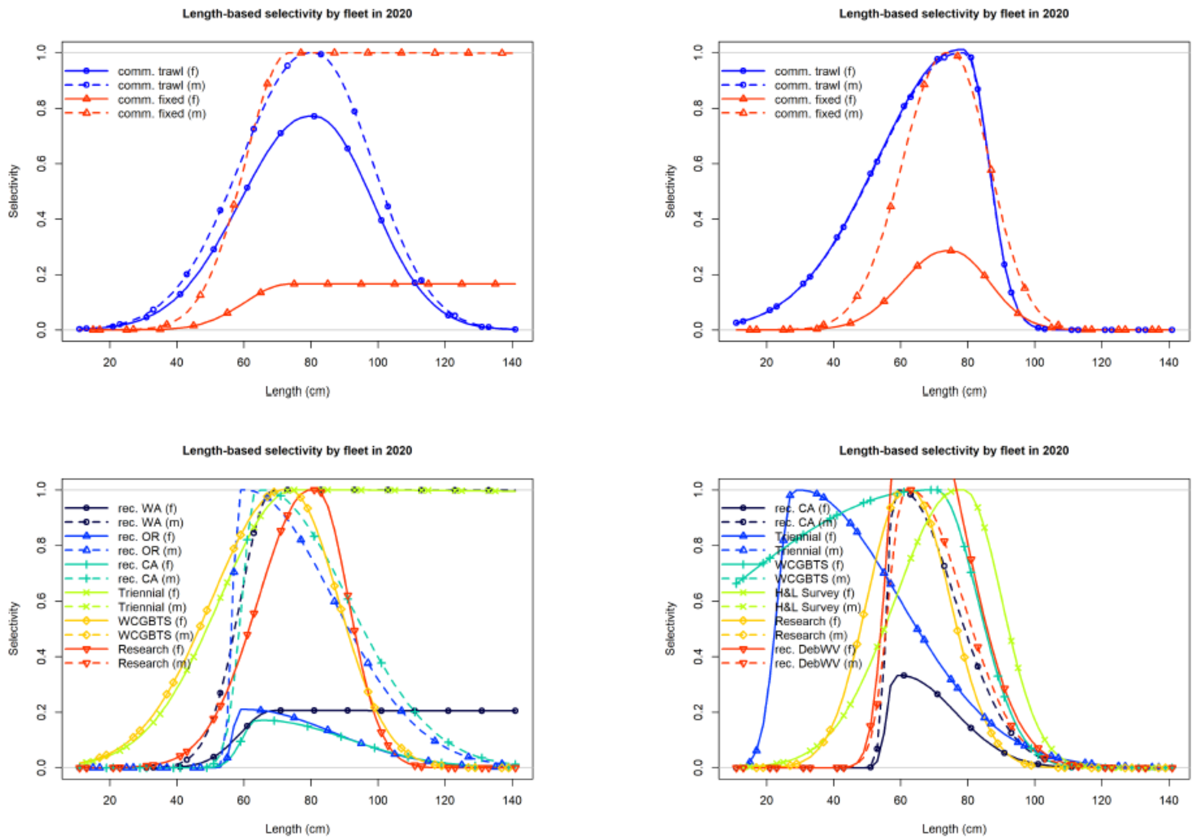


Figure 4. Selectivity of males and females using the sex selectivity offset. Figures on left are for the Northern stock, while figures on the right are for the Southern stock.

Request No 12. Explore freeing up the peak and the descending limbs of the offset selectivity curves for lingcod in both the northern and southern models, including the base model and the model from #11. Try this parameter by parameter, or in a combination deemed reasonable by the STAT. Provide model comparison slides.

Rationale: Adding sex-based selectivity offsets to the model improved some aspects of model fit but still produced implausible selectivity curves (among other issues). It is possible that allowing more latitude in fitting the selectivity curves would improve model behavior.

STAT Response: There are a large number of selectivity parameters given the number of fleets and time blocks, so a select number of options were presented in response to this request. These options include several scenarios in which certain parameters were shared (or not) across time blocks for particular fleets, and were (or were not) allowed to differ between sexes. The general outcome was that models allowing sex-specific differences in selectivity in one or more fleets led to much lower estimates of M than the base model, while non-sex-specific models produced M estimates more similar to the base model. Some of the estimated selectivities had a knife-edge right-hand side, which seems implausible. The most complex model considered had sex-dependent variation in both the peak and descending limb of the selectivity curves, and it did produce selectivity curves that could be consistent with a fixed-gear targeting of smaller inshore males. No alternative model produced large differences in estimated stock status relative to management targets.

Panel Conclusions: These exploratory analyses confirm that the data support some form of sex-dependent selectivity in the fisheries. Yet the attempts to model that feature lead to a trade-off between estimating selectivity and M, such that either both M and selectivity are estimated to be low (fish live long but are not caught) or high (fish are highly selected but die before they are caught). Without further research and likely a different approach to handling fleet structure and age data (see research recommendations) there is not an acceptable way to capture these dynamics with the current model structure.

Request No 13. Provide a likelihood profile for female M for the model that the STAT deems most plausible based on the result from request 12. M values should range from 0.1 to 0.5.

Rationale: As both base models have changed, and both continue to be highly sensitive to changes in M, this will provide the opportunity to re-evaluate sources of tension in the model.

STAT Response: The attempt to produce this profile was not entirely successful because the model had difficulty converging at low values on M. There is some evidence for a bimodal likelihood surface with negative-log-likelihood minima near both 0.2 and 0.5.

Panel Conclusions: This analysis appears to confirm the suspicion based on the results from Request 12 that the data as currently structured favor two alternative modes in parameter space; one with low M and one with high M. Unfortunately neither M falls within the range the panel would expect for lingcod life history.

Request No 14. Investigate the discrepancies in catches between the Groundfish Expanded Multiyear Mortality (GEMM) reports and the northern model estimates. Explore modeling methods to attempt to improve the fit to the dead discard rates in the model to produce dead discard estimates that are more in line with the GEMM values. This adjustment would consider these earlier years (1998-2001) as well. This investigation should focus on the trawl fleet.

Rationale: There are large discrepancies in catches between the GEMM reports and the model estimates. WCGOP discard rates in the model are from the same data that produce the GEMM estimates, hence we would expect the model estimates to have coherence with the GEMM estimates.

STAT Response: Several options to address this issue were considered: two different adjustments to reduce the standard error associated with the discard rate data (to encourage closer fits to that data stream) and an artificial deflation of the discard sample size to reduce its weighting in the likelihood calculation. Reducing the SEs produced better fits to the discard rates but poorer fits to other likelihood components. Reducing effective sample size improved fits to other likelihood components at the expense of fitting discards. No option produced substantial changes in estimated biological parameters or stock status.

Panel Conclusions: Because the changes to overall model fit are small and the considered alternative solutions are very ad-hoc, the base model should not be changed to include these options. However, this issue is a future research need.

Request #15: For the North model, provide a run in which the standard deviation on the prior on female natural mortality is divided by 2, and do this for both the base model and for model with sex-specific selectivity (model 417). Present model comparisons and associated informative tables/figures as in prior requests. Provide the whole suite of r4ss files for panel review.

Rationale: There continues to be considerable tension between model structure and the dynamics of the population, with model estimates of M seeming to vary widely in response to relatively modest structural changes in the model.

STAT Response: The STAT provided the results of four models, the base model, the model with moderate improvement in the discard estimation (from Request 14), the earlier model run with sex-specific selectivity (from Request 12, denoted model slx.417 in the figure), and the sex-specific selectivity run with a with the more informative prior on female M. The results clearly indicated that a more informative prior resulted in very little change to the model result (Figure 5), as the improvement in fit with a substantially higher M is far greater than the constraints on the prior. In other words, the data are pushing the model to a higher value of M despite the prior, and the result was only a trivial change in M. When the change in discard estimation was included, the result was actually an increase in M to a value just above that estimated in the base model. The model with the sex-based offset in selectivity resulted in M hitting the lower bounds for male lingcod (suggesting that male lingcod live very long lives), a result that did not seem plausible to either the STAT or the Panel. The STAT suggested that the sex-based selectivity exploration become a research priority, and the Panel agreed.

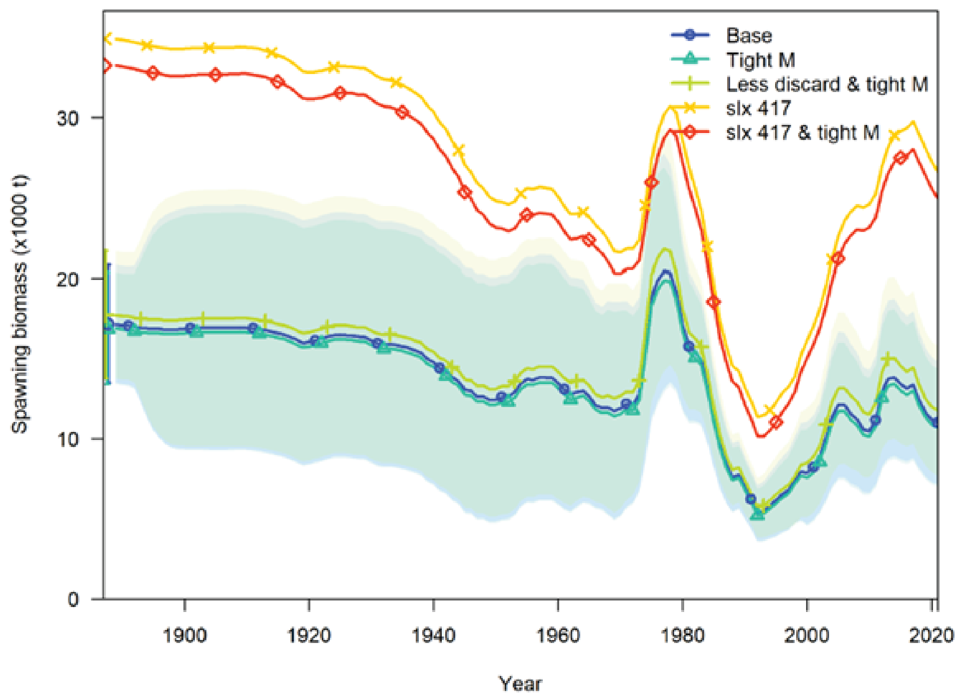


Figure 5: Model runs with and without a more informative prior on the female natural mortality rate.

Panel Conclusions: There was also general agreement that there seem to be two alternative local minima in the negative log likelihood surface suggested by the data, associated with low and high natural mortality rates. It was hypothesized that this could potentially be confounded with the boundary between the two stocks. It was noted that fixing M at the mean of the prior (0.3) could essentially be a form of model averaging, however the data clearly seem to want to avoid middling values of M in favor of values either much greater or much lower than this.

The results also indicate that the management quantities do not change that much among those two states. There was general agreement that although these explorations have allowed the Panel and STAT to better understand the model, we have mostly learned that there is a lot more to do in order to better fit the data and come up with a model that appropriately captures the complex sex-specific dynamics of this population.

Request #16: For the north model, provide a run in which female M is fixed at 0.3. Present model comparisons and associated informative tables/figures as in prior requests.

Rationale: The tighter uncertainty bounds on the prior did not influence the model result, but rather reinforced the notion that the model may have two stable states.

STAT Response: The model result was considerably more pessimistic regarding stock status than the base model or the models with alternative selectivity functions, as had been initially suggested by the likelihood profiles in the draft document, despite a relatively modest degradation in overall fit. Fits to age data improved, fits to both index and length data degraded. There was general discomfort by both the STAT and the Panel with this approach.

Panel Conclusions: There is general agreement that what is happening here is a bit more than just sensitivity to M, and that fixing M at the prior estimate is forcing the model to a solution more contrary to the data than the base model or other alternative models. The core challenge seems to be that there are conflicts between two types of errors- observation error in the data (leading to high tension between length and age data) as well as structural error in the model dynamics. Specifically, the inability to implement a reasonable sex-based selectivity function in the model when there is strong evidence for differences within and among fleets in selectivity by sex suggested in the data presents a key challenge in adequately capturing the dynamics of the population.

Request 17: Provide a retrospective analysis that goes back 5 years for the northern and southern models. Report the Mohn's rho values –(Woods Hole and Alaskan/Hurtado-Ferro).

Rationale: There are significant scaling issues associated with the retrospectives that merit a closer investigation

STAT Response: The STAT ran retrospective analyses for the final models. During this process, it was noticed that the approach of removing a single year's data is too simplistic given that these models have different time blocks and fixed parameters that could differentially interact with the retrospective analyses. As an example, the WA recreational fishery had a 22 inch minimum size limit lifted in 2017 resulting in a selectivity block in the model that would impact the N model retrospective analyses. Furthermore, typical of many assessments, the estimate of the terminal year recruitment is the most uncertain in these assessments. It would be more appropriate to address the time blocks and therefore slightly change the model in each peel. The results produced for the panel were a first step in this direction of better characterizing the effect of 1 year's less data in the model, but this approach is only a first step rather than a final solution to the problem. It is also important to note that the retrospectives are not re-tuned models. Not retuning the model may overemphasize model sensitivity to the peels.

Panel Conclusions: For the North model: Some of the peels extend beyond the uncertainty bounds of the final model, which is a source of concern because the new base case is not comprehensively characterizing uncertainty. There are also recruitment patterns in some of the retrospective runs that are not shown in the base model. These results support the sensitivity tests and highlight that there is tension in the model between M, selectivity and productivity. The Mohn's rho results are also an indication that the assessment should be classified as a category 2 assessment.

For the South model: the South has a much higher uncertainty interval in the final model, but the retrospectives show that SSB size still changes even above the final model's confidence intervals. In this case, there is also some uncertainty to the fraction unfinished (unlike in the northern model).

In general, the retrospectives of both assessments are showing that the recent years' data are providing information about scaling, but that there is great uncertainty in scaling the resource when these data are removed. This has already been shown through the different sensitivity tests.

However, there is also some degree of stability in the fraction unfished reflecting that there is more information in the relative change in biomass and recruitment over time.

With regard to high Mohn's rho values, it is important to make sure that starting values of the projections are appropriate when there are retrospective patterns in the assessment. For example, the stock-recruitment function is based around the historical SSB values and, if biased, then the steepness estimate used in the projection may not appropriately reflect the future.

Although there are very large sigma values in the final model, in the last year sigma is estimated at 35% for the south and 18% for the north model. The management uncertainty applied to a category 2 assessment would be much bigger than these values, which is appropriate given the both base models appear to significantly underestimate uncertainty.

The STAT was asked to include the recreation deviations in the retrospective report as these are very informative.

Request 18. Provide the r4ss files for the revised northern and southern Base models.

Rationale: This is a simple diagnostic check and the panel uses the information to review outputs.

STAT Response: These were provided to the panel.

Panel Conclusions: There was some discussion about the differences in the selectivity in the surveys between the North and South final models. This reflects that the size composition data are very different between the two regions. In the Triennial survey in the N there are not as many small animals than in the South. It was noted that the Triennial surveys does not cover the full coast of CA. However, there is a difference in the bathymetry between the N and S regions that could be contributing to the difference in the survey data, and could cause different interactions between the surveys, the fisheries and distribution of the resource. This reflects that there are spatial scale processes that are not fully captured in these assessments.

Request 19: For the southern model, develop the runs that would fill in a decision table based on the high and low quantiles of M as inferred by the likelihood profile. Provide diagnostic outputs as appropriate.

Rationale: This is likely to be a reasonable basis for the decision table for the southern model.

STAT Response: The team ran into convergence issues when diagnostics (such as the jitters and profile likelihood) were run in the base model as agreed on day 4 of the meeting. This meant that further changes had to be made by the STAT team to the southern model. This included changing some of the bounds and estimating parameters that were fixed for the triennial selectivity parameters. This resulted in minor changes to the model but avoided these convergence issues. This was proposed as the new base case.

Panel Conclusions: The Panel supported this model as the final base case

Request 20. For the northern model, develop model runs that might encompass the different types of both observational and structural uncertainty by 1) excluding fishery dependent age data from a model and, 2) running a model with sex-specific selectivity (as in model 420) to capture the "process" uncertainty. Include in the comparison plots and tables the model run in which M is fixed at 0.3 for females.

Rationale: Model sensitivities without the age data had the best fit to length data and indices, while the model with sex-specific selectivity implemented as freeing up descending limbs of the selectivity curve had the best overall fit to the data. The results of each also bracket a wide range of possible model outcomes.

STAT Response: Four options were provided for the northern model decision table to describe uncertainty – the final model, the final model without ages (representing the high stock state), the model with sex specific selectivity (which represents the low productivity alternative state and structural uncertainty) and the fixed M=0.3 model (the low stock state). There was some concern expressed regarding the M=0.3 run in that the likelihood of this model compared to the base model is almost 17 units above the base model and does not fit the data well.

Panel Conclusions: Although this is an unusual sensitivity table, it clearly reflects the uncertainty within the north model. These are two axes of uncertainty reflected therein. The first emphasizes data uncertainty by reflecting the tension within the composition data that produces uncertainty with regard to scale and productivity of the resource. When the age data are excluded, the virgin SSB is increased, but the natural mortality is considerably lower than the base model (0.33 for females, 0.29 for males). As a result, the results associated with excluding the age data from the final model are a good reflection of this axis of uncertainty.

The second source of uncertainty is structural uncertainty, which is demonstrated by the inclusion of sex-specific selectivity. The possibility of sex-specific selectivity is likely related to processes associated with different movement patterns of male and females, or of the different habitat types that are more or less available to the trawl fleet, for example. In this configuration, the SSB is higher than the final model, but natural mortality is very low – similar to the south model.

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

The structure of the final base models for both northern and southern lingcod changed only modestly from those provided to the review panel, and details are described in the assessment documents and summarized in the "Summary of Data and Assessment Models" in this report. Minor corrections to catch streams were made, the early length composition data in the southern model were removed, and considerable changes to model structure were explored, but ultimately not adopted. Those changes led to a greater appreciation of the magnitude of what more there exists to do in order to better fit the data and more appropriately capture the apparently complex sex-specific dynamics of this population.

Both models estimated most key parameters, including sex-specific natural mortality and steepness, however the resulting natural mortality estimates were highly divergent between the two models, and the models were very sensitive to the M estimates, which in turn were very sensitive to other aspects of the model structure. Interestingly, neither model estimated natural mortality rate values consistent with either the prior (0.3 for females) or the previous assessment (0.25), with the northern model estimating a significantly greater natural mortality rate (0.41) and the southern model estimating a significantly lower rate (0.17). The focus of the model exploration during the review revolved around better understanding of the drivers of these unusual patterns and discrepancies. Similarly, steepness estimates were also considerably different between the northern and southern models, with the northern model steepness estimated at 0.80 and the southern model estimated at 0.51.

An overarching conclusion by both the STAT and the STAR Panel is that neither model is capturing key aspects of the population dynamics, there are many clues as to why this is, but no clear short-term fixes from a modeling perspective.

With respect to the results of the base models, both north and south assessments estimated that the stocks were above the $B_{40\%}$ reference point in 2021. The northern stock was estimated as having never been overfished (below $B_{25\%}$), and to be at approximately 64% of the unfished level in 2021. The Southern stock was estimated as having been overfished during the period 1990-20 but at almost exactly target levels (41% of unfished) in 2021, following a series of strong year classes between 2008 and 2013. Despite a range of issues in both models related to the overall scaling of abundance, the results of both models (perhaps more so the northern model) tended to be more robust with respect to relative stock status, which likely reflects that each model has several, informative indices to provide information on relative abundance.

Uncertainty

For the southern model, uncertainty over the model estimate of natural mortality was determined to be an appropriate axis of uncertainty, with the high and low states of nature based on the high (0.22) and low (0.11) quantiles of female natural mortality, as inferred by the likelihood profile (in which the base estimate was 0.17). Catch stream advice was provided by the GMT, and was based on recent average catch, the estimated ACL with a category 2 designation and a P* of 0.45, and the estimated ACL with a category 2 designation and a P* of 0.40. As the base model indicates relatively strong recruitment (positive recruitment deviations) from approximately 2008-2013, and relatively weaker recruitment (negative recruitment deviations) from 2014 through recent years, most model trajectories in the decision table indicate stable or declining trends for all three catch scenarios. All trajectories in the decision table leading to depletion estimates within the precautionary zone (e.g., between 25% and 40% of the unfished level).

For the northern model, the model estimated M was considerably greater than both the southern model estimate and the mean of the prior, and while natural mortality was clearly a key uncertainty, the Panel considered that a “typical” symmetrical evaluation of the uncertainty around the base model estimate not to be the best approach for this model. Consequently, consistent with the terms of reference guidance that “expert judgement” can be used to develop decision tables when warranted, an alternative approach that addressed the key uncertainties in the model was developed. Catch streams were developed using the same basis as those for the southern model.

In this approach, the "high" state of nature is reflected by excluding the fishery-dependent age data, which are a source of considerable tension in the model (and as such were also excluded in the 2017 model), and without which the model estimates of productivity are greater. Alternatively, the "low" state of nature reflects the model with sex specific selectivity (reflecting model structure uncertainty), in this model the resulting SSB is higher than the base model, but natural mortality and productivity are very low (similar to the south model). Although the ending relative spawning output is generally comparable across these three states of nature, the equilibrium MSY levels have a broader range, such that they are approximately 34% of the base model equilibrium MSY for the "low" state of nature, and 125% of the base model equilibrium MSY for the "high" state of nature.

Consequently, population trajectories under the alternative catch streams diverge substantially across these three alternative states of nature. Specifically, under the "low productivity" state of nature, the stock remains stable with catches at recent average levels, but declines sharply with catches close to ACL levels (with either P*). Under the base and high productivity states of nature, none of the catch scenarios result in the stock declining below management targets. The Panel and STAT both note that because the relative depletion estimates appear to be robust among the three models, presumably due to informative survey indices, there will be an ability to qualitatively evaluate relative abundance trends (based on those indices) for this stock to help ensure that the stock is not at significant risk of depletion. This is helpful if indeed the "low" productivity state of nature best represents the actual dynamics of the population.

Sensitivity Analyses

The draft assessments included a broad range of sensitivity analyses, including retrospective analyses, likelihood profiles over key model parameters, and a suite of structural assumptions (many more of which were explored during the STAR Panel). The retrospective analyses suggested substantial changes in scaling of total biomass, well outside the confidence bounds of the base model for the north, within the confidence bounds but still highly divergent from the base model for the south, which was fairly concerning for both models. However, the depletion estimates were within the uncertainty bounds of the base model for both areas as well. Both models were intuitively sensitive to alternative assumptions regarding steepness, and highly sensitive to assumptions regarding both natural mortality and alternative model structures, as discussed throughout this report.

Technical Merits of the Assessment

Both the northern and southern assessment are relatively data rich, each are informed by both fishery independent and fishery dependent index data, each have robust time series of length composition data for the major fisheries, and the northern model has a considerable amount of age data (the southern model has limited age data). Although there is clear tension among data sources in the model, for most plausible parameterizations of the model the management quantities and estimates of relative abundance are within the confidence bounds of the base model.

Notable improvements since the 2017 benchmark assessment included several efforts that were recommended in the research and data needs for that assessment, including

- Incorporation of ages into the Northern model as CAAL rather than marginal
- Work done to change the boundary between northern and southern stock based on genetic analysis
- Improvement to historical catch estimates (both commercial and recreational)
- Estimation of key parameters, particularly steepness and natural mortality rates

Moreover, both the northern and southern assessments, the analyses were very thorough with respect to documentation, diagnostics, sensitivities and complementary information on all of the above. Significantly, surveys appear sufficiently informative to estimate the relative abundance of the stocks over recent time periods.

Technical Deficiencies of the Assessment

Tension between age and length/index data in the northern model remains a key challenge for the northern assessment, while the lack of fishery-dependent age data is a key uncertainty in the southern model. The southern model also includes a split in many data sources north and south of Point Conception (particularly relative abundance indices, most of which are primarily north or south of this region). The substantial differences between model estimates of natural mortality among the two areas are disconcerting, although not unheard of or unprecedented, and may be related to the described differences in stock structure and genetics, or potentially habitat-driven differences in dominant life history patterns (investigations along these lines were discussed by the STAT).

The northern border of the northern stock may not be quite correct from a stock perspective given the recent genetic analyses. The absence of a recent assessment for Canadian waters is also concerning, as is the lack of information along the southern border of the south model. The Assessment team did contact fisheries managers in both of the areas adjacent to the boundaries of these assessments in an attempt to better understand the data that might exist for those regions.

Other differences between north and south also appear non-intuitive. These include survey and fishery selectivities between northern and southern models being strikingly different, which is troubling. The lack of consistency between estimated biological and fishery parameters between the two models is an indication of the overall model uncertainty. Some of the substantial differences in ecology among the sexes are not fully captured by the model (e.g., selectivity), and this is also a considerable source of uncertainty.

Areas of Disagreement Regarding STAR Panel Recommendations

Among STAR Panel members (including GAP, GMT, and PFMC representatives): None
Between the STAR Panel and the STAT Team: None

Management, Data, or Fishery Issues raised by the GMT or GAP

Representatives During the STAR Panel Meeting

The GMT and GAP representatives raised several management, data, and fishery issues during the STAR panel meeting. Those issues are summarized in the bulleted list below:

- The poor fit of the northern model modeled dead discards to the dead discards in the trawl fishery from 2002-2010 found in the GEMM product. The STAT team attempted to address the issue, but additional examinations should be conducted in future assessments
- Corrections to the Washington recreational discards, which were completed during the week
- Need for a method to disentangle retained (dead) vs. discard dead in California recreational catch estimates
- The use of fishery independent data sources (e.g., ODFW Marine Reserves hook and line data, lander surveys) and how we can work in between cycles to be able to incorporate that data.
- The disconnect between some age data appearing to be biased and not meeting the diagnostic criteria for random sampling.
    - What does that mean for sampling programs/ sample design, and their implementation
    - How can States ensure that data collected will be used in stock assessments.
    - Do we need to change the way we collect things?
- Need for better understanding of the live fish fishery in terms of fisher behavior and magnitude of the fishery, as well as the biases and nuances of data collected from that specific sector
    - Additionally, differences in the nearshore and offshore FG strategies
- The GAP representative mentioned the need for additional port sampling in CA which was a big concern coming out of the 2017 assessment and continues to be a concern

Unresolved Problems and Major Uncertainties

A key challenge with the northern model in particular is that the northern model is data-rich, but information-poor. It is important to note that these challenges were present in the 2017 model, but avoided by excluding most age data and fixing the natural mortality rate at the mean of the prior. As the current northern model includes additional data and other improvements relative to the 2017 assessments and provides relatively consistent management advice to the earlier models, the Panel agrees that the 2021 assessment models are an improvement and should be used to inform management.

However, the Panel recognizes that the results of both models, particularly the northern model, indicate so much tension in the models that they almost indicate two divergent life histories, one high and one low productivity, which leads to a high level of discomfort with the model results. This is true both between the northern and southern models, as well as within the northern model. The apparent lack of comparable divergence and tension in the southern model is most likely due to the more data-poor condition of that model, particularly relative to age data. Both models also had retrospective patterns that diverged and rescaled model estimates relative to the base runs, contributing to additional concerns about model stability. Consequently, the Panel would recommend that these models are most appropriately considered category 2 models for the purposes of providing management advice. Similarly, the Panel does not recommend a routine update for this model, but rather recommends that future assessments be full assessments.

Recommendations for Future Research and Data Collection

High priority

- Consider alternative means of modeling sex-varying selectivity curves, as the explorations initiated by the STAT and explored in this review suggest that there are likely to be differences in selectivity by sex that are driving a considerable fraction of the tension in the model.
- Better information and exploration of the differences in fishing strategies throughout the fixed gear fisheries, particularly the live fish fishery in the northern model, and how that fishery is modeled. Specifically, there is likely merit in exploring a model in which the fixed gear live fish sector is treated as a separate fleet, given that they likely fish more shallow waters than the fixed gear (dead) fishery, and due to sex-specific nest guarding behavior, the selectivity offsets by sex that should be explored in future models may differ by fishing strategy. Moreover, the lack of age-structure collection from the live fish sector may bias age composition data for the fixed gear fishery as a whole, this could be less of a problem if this fishery is treated as a separate fleet.
- Further exploration of natural mortality for both Northern and Southern stocks which may include; tag based investigations, simulation modeling, incorporation of Lorenzen/Charnov/Then et al. information, as well as explorations by region and/or year. Consider such investigations in the context of dome-shaped selectivity.
- The STAT indicated that discard length composition data are better sampled than retained length composition data, yet when the data are treated within the same fleet the tuning and weighting approaches treat the two types of data equally. The ability to weight these two components of the length composition data separately would be very helpful in disentangling sources of tension in this (and other) models.
- Consider including the CCFRP dataset in future assessments as a fishery-independent data source. Also, consider the other recreational indices of abundance that were available in the southern model but not used, such as the 1999-2019 California onboard observer CPFV index.
- Explore consequences of modeling recreational catch and discards as numbers rather than converted biomass
- Explore possible approaches to improve estimation of trawl discard mortality rates, given mismatch between model estimates and GEMM estimates
- Additional and expanded age structure collection for the southern model continues to be a high data need for this stock, from both fishery dependent (e.g., hook and line survey) and fishery independent sources.

Medium Priority

- Explore time-varying growth, as that could potentially be one mechanism for tension between length and age data, particularly among different time periods. Given the high population growth and turnover rates, environmental conditions could lead to significant differences in growth among different environmental regimes.
- Investigate the Northern stock border and connectedness with Canadian/Alaskan stocks, including the Salish Sea. Also consider issues related to the southern border with Mexico in the southern model, albeit while recognizing robust data from Mexico do not currently exist.
- Support ongoing efforts to compare fin ray and otolith aging comparisons as well as further comparisons in aging between NWFSC, DFO, and ADFG.
- The VAST method used to standardize some of the indices include a component of spatio-temporal auto-correlation. Depth was also often a factor in the final Bayesian delta-GLM analyses. These results reflect the biological information that there are depth and time differences in the movement between the sexes, and these would be affecting the representativeness of the composition data. It is usual that age and length data is not standardized, but in this fishery some form of in-depth analysis of the biology of the species relative to alternative approaches to standardizing the composition data is warranted, in the north model particularly.
- Although earlier models estimated separate sigma R (standard deviation of recruitments off of the spawner-recruit curve) for northern and southern models (0.75 in south, 0.5 in north), this assessment used the same value (0.6) for both models. However, sensitivities suggested that recruitment variability may be greater in the south rather than the north, which would be consistent with some general patterns of recruitment variability in other stocks. This should be re-evaluated in future assessments.
- More consideration needs to be made on how best to implement retrospective analyses in models that include time blocks and address parameter inefficiencies, for example. Only peeling back 1 year of data is not really reflecting how the assessment would be done at that time and may not reflect the model uncertainty to the loss of a year’s data. This means that some form of re-tuning and changes to settings (e.g., time blocks) may need to be undertaken in retrospective analyses. What would also be useful is a quick method to compare the likelihoods of the retrospectives for the years that data are shared. This information is presently provided in SS, for example, but not automatically provided in packages such as r4ss. It should be highlighted here that this is a general issue to both data rich or data limited assessments and are not particular to this assessment. Presently it is difficult to understand the cause of retrospective patterns in a model, or what an appropriate management response to those patterns should be.
- Analyze the trawl fishery relative abundance index data by individual stock rather than both combined.
- Continue to investigate the potential sources of bias to the nature of fishery dependent age-data sampling, particularly with respect to the implications of non-random sampling of aged data even when using a conditional age-at-length approach. Consider improvements to both sampling approaches and the preparation and treatment of the age data in assessments.

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

Longo, G.C., Lam, L., Basnett, B., Samhouri, J., Hamilton, S., Andrews, K., Williams, G., Goetz, G., McClure, M. and Nichols, K.M., 2020. Strong population differentiation in lingcod (*Ophiodon elongatus*) is driven by a small portion of the genome. Evolutionary applications, 13(10), pp.2536-2554.

Thorson, J.T. and Barnett, L.A., 2017. Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Science, 74(5), pp.1311-1321.