Cabezon Stock Assessment Review (STAR) Panel Report

NOAA Fisheries, Northwest Fisheries Science Center
Barry Fisher Building, Conference Room 101
2032 SE OSU Drive
Newport, OR 97365

May 6-10, 2019

STAR Panel Members

Rishi Sharma, National Marine Fisheries Service Northwest Fisheries Science Center, Chair
Noel Cadigan, Center for Independent Experts
Robin Cook, Center for Independent Experts
Will White, Oregon State University

Stock Assessment Team (STAT) Members

Jason Cope, National Marine Fisheries Service Northwest Fisheries Science Center
Aaron Berger, National Marine Fisheries Service Northwest Fisheries Science Center
John Budrick, California Department of Fish and Wildlife
Ali Whitman, Oregon Department of Fish and Wildlife
Katelyn Bosely, National Marine Fisheries Service Northwest Fisheries Science Center
Theresa Tsou, Washington Department of Fish and Wildlife
Lisa Hillier, Washington Department of Fish and Wildlife
Kristen Hinton, Washington Department of Fish and Wildlife
Megan Wilson, Oregon State University

STAR Panel Advisors

Mr. Patrick Mirick, Oregon Department of Fish and Wildlife, Groundfish Management Team
Mr. Gerry Richter, B&G Seafoods, Inc., Groundfish Advisory Subpanel
Mr. John DeVore, Pacific Fishery Management Council

Overview

The STAR panel met the week of May 6th-10th in Newport, OR, to assess the status of Cabezon stocks on the west coast. This was the first time a complete west coast assessment was conducted on Cabezon from CA to WA. The resolution of the data differed extensively by state, and while some areas had a lot more data to be used, others were still limited by a paucity of data (CA) or data poor approaches (WA). The assessments addressed four sub-stock areas: Southern California Substock (SCS), Northern California Substock (NCS), Oregon Substock (ORS), and Washington Substock (WAS). The first three were full integrated stock assessments performed using Stock Synthesis 3 (SS3); the WAS assessment was a data-poor assessment using Simple Stock Synthesis (SSS).

The STAR panel recommends that all four substock assessments for cabezon constitute the best available scientific information on the current status of the stock(s) and that the assessments provide a suitable basis for management decisions. The panel further recommends that the SCS, NCS, and ORS assessments be considered category 1 (full integrated assessments), and that the data-poor WAS assessment be considered category 3.

The STAR panel further recommends that the next ORS assessment be a full assessment pending new information or explorations about within state stock structure, but would otherwise be suitable as an update. The panel also recommends that the next WAS assessment be a full assessment, based on the likely availability of new data that may permit a full integrated assessment. The panel recommends that the next SCS and NCS assessments be updates rather than full assessments.

Summary of Data and Assessment Models

The structure of the SCS, NCS, and ORS assessment models were nearly identical except that different data sets were available for each area, and hence different biological parameters were able to be estimated. These assessments use a recent version of Stock Synthesis 3 (Version V3.30.13). The WAS assessment took a data poor approach using Simple Stock Synthesis. The main sources of information and modeling details in each assessment are given below.

NCS Assessment

Data sources:

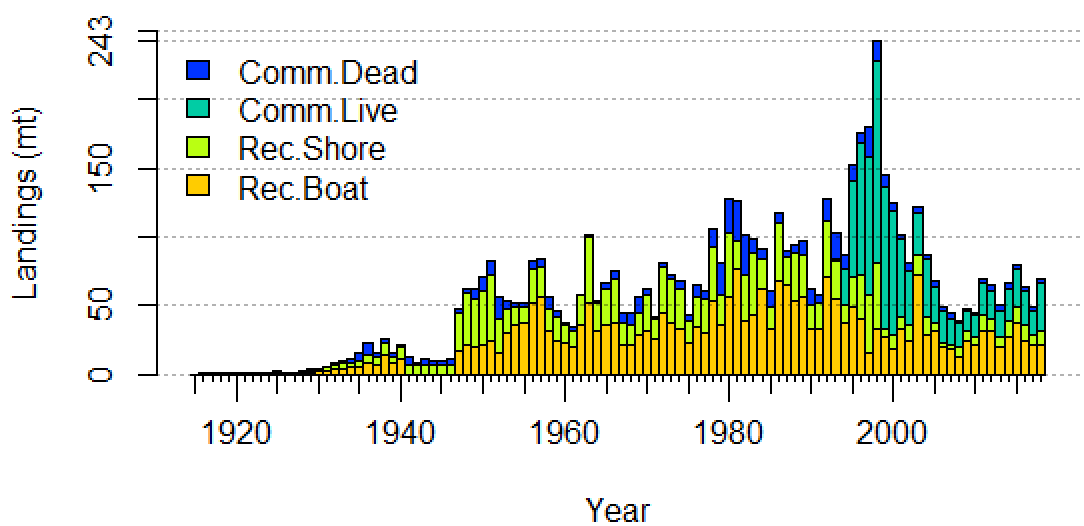
- Catch and length composition from four fisheries: commercial dead-fish, commercial live-fish, recreational boat (both charter and private boats; length from the charter boat fishery only), and recreational shore (though length compositions were not available from all fisheries in all years).
- Relative abundance (CPUE) indices based on the recreational charter boat fishery (Commercial Passenger Fishing Vessel, CPFV, a fishery-dependent index) and a scientific stratified-design hook-and-line survey (California Cooperative Fisheries Research Program, CCFRP).
- Conditional age-at-length data from a scientific research survey

Notable updates to data sources

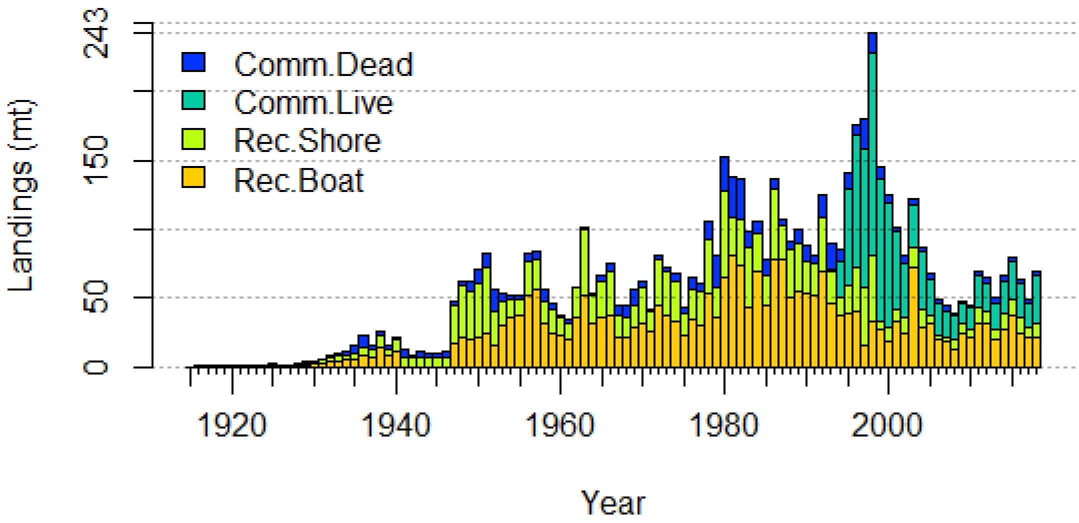
- Conditional-age-at-length samples from the previous assessment remained the only available samples for the current assessment. One change in treatment was to put all ages in the NCS model, as the very few samples that were in the past SCS model were not used in estimation. The inclusion of the full data set in the NCS model contributed to the estimation of the growth parameters that were then used in the SCS model.
- Historical catch times series remained largely the same from the last model, with new catches being added to complete the time series. One difference came in the new pull of recreational data from 2004-2008.
- A more significant change in the catch data was the reallocation of catches from the SCS region to the NCS region for years 1980-1995. This was in response to the recognition that past treatment of MRFSS catch estimates were based on stratification of California at 36° N Lat. as opposed to 34°27' N. Lat. at Point Conception, as the assessment structure assumes. This change was made after the pre-STAR draft assessment was prepared, but prior to the STAR panel, so this updated model was presented to the panel (see Figure below).

NCS model

Pre-STAR reference model catch

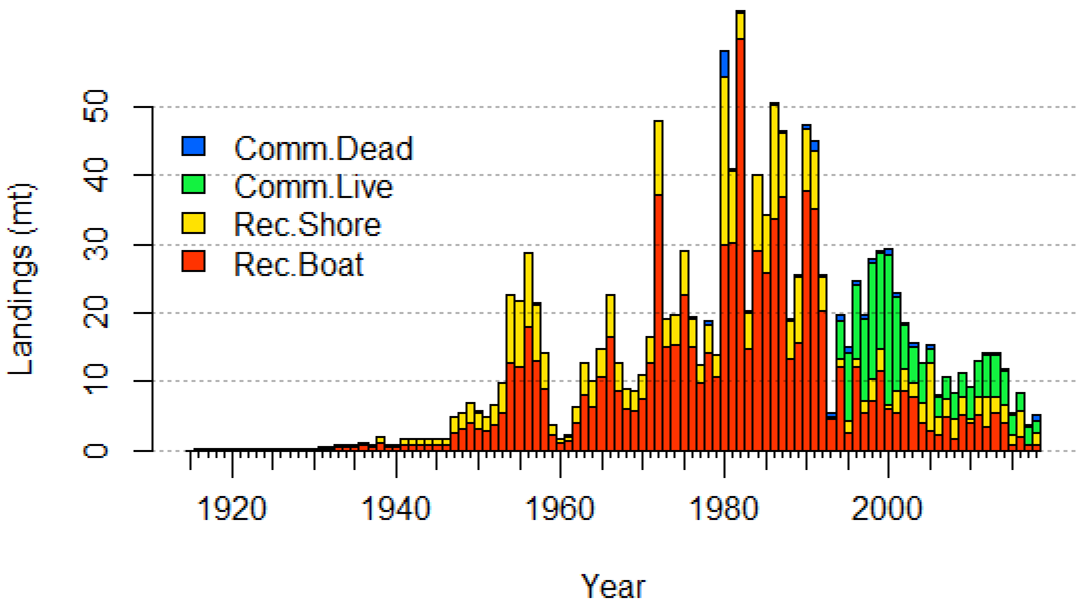


Alternative catch (now in Post-STAR model)

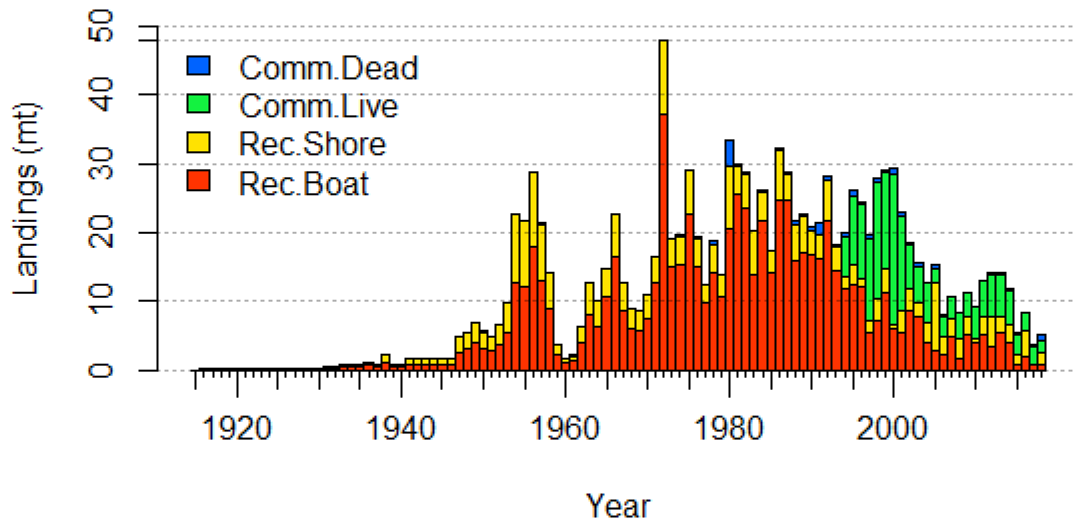


SCS model

Pre-STAR reference model catch



Alternative catch (now in Post-STAR model)



- A change was made in the assumed timing assigned to the catches. The previous assessment assigned catch to the month 1, whereas the new assessment assigned them to the mid-year 6th month.
- The CCFRP survey is a new addition since the 2009 assessment.
- Mean weights were excluded from the new assessment. The length of the time series of mean weights had been reduced from the last assessment with the recovery of some true length measures in the early MRFSS time period. Once that happened the final reference models showed low information content in the remaining mean weights, justifying removal from the current model.
- Length compositions again were similar to the previous assessment, with the addition of subsequent years. Month assignment was switched from 1 to 6, as done in the catches. Additional lengths for the CCFRP survey were also used in the new assessment. The same length bin structure was retained from the last assessment.

Model details (including details changed from the previous 2009 assessment model):

- Model time coverage starts from the same date used in the last stock assessment (1916) and continues through 2018.
- The two sexes are modeled separately, as growth is very different between females and males.
- The yearly time step is 12 months, though 6 sub-seasons were defined in order to allow for more flexibility in the treatment of fleets.
- The accumulator age was dropped from 35 to 25 to reduce model dimensions. Given the

likely natural mortality range, 35 years was a very high consideration.

- The number of fleets was reduced to 4 from the previous 6. This reduction came from combining the man-mad and beach/bank mode into a shore mode, and combining the private and charter boat modes into a boat mode. Model development showed similar length compositions between the modes and no appreciable difference in model outputs when using 6 vs. 4 fleets. The reduction in fleets, though, did reduce the number of estimated selectivity parameters as well as increase sample size for within-year length compositions.
- Selectivity block years were slightly adjusted in the new model to better match changes in the length compositions and known management changes. Numbers of blocks remained the same.
- The current NCS model estimated the length at age 0 to be close to 0 for both sexes, so that parameter was subsequently fixed, which improved model estimation. All other growth parameters were freely estimated with no prior. The von Bertalanffy k parameter was originally estimated using a prior based on the Grebel (2003) data, but this was changed in response to panel concerns.
- Additional biological parameters were fixed to the same values as in the previous model.
- Steepness and recruitment variability were fixed to the same values as in the previous model.
- Recruitment estimation differed as the current assessment used the method of Methot and Taylor (2011) to identify years of recruitment estimation and the treatment for bias adjustment to make it more consistent with the assumed recruitment variability. The previous assessments assumed all estimated recruitment years received a full bias adjustment (equal to 1), with years of estimated recruitment 1970-2006 in both models. The current NCS model estimated recruitments from 1962-2016, with the ramp from 0 in bias adjustment starting in 1964 and reaching its maximum value of 0.63 from years 1983-1998 (years of peak information), ramping again down to 0 in 2017.
- Change from Pope's approximation of F to the hybrid method.
- Both the previous and current model analytically calculated the catchability coefficient for each survey. Additional variance was also estimated for both CPFV surveys, but was set to 0 for the CCFRP survey as attempts to estimate this parameter always resulted in a value close to 0.
- Selectivity curves treatments remained mostly the same from the previous assessment: The commercial dead and recreational shore and boat fleets were estimated as asymptotic; the commercial live fishery was allowed to be dome-shaped. Selectivity for the new length composition data from the CCFRP survey was also free to go dome-shaped. Time-varying blocks were applied to the commercial live and the recreational boat fleets.
- Data-weighting was treated differently than the previous model. The 2009 model used the harmonic mean approach (McAllister and Ianelli 1997) whereas the Francis method (Francis 2011) was applied in the current models. A sensitivity to this choice of data weighting (as well as no data weighting) was explored as model sensitivities.

Overall Data Issues

The datasets included both catch, length composition and standardized indices of abundance.

Other issues related to the datasets and structural uncertainty examined are covered below:

1. Extensive sensitivity analyses were provided, both to data and structural model uncertainty.
2. The combination of rec shore (last assessment 2 fleets) and boat (2 fleets) compared to 2009 assessment seemed appropriate.

Parameter choices for reference models

An M prior was developed using life-history theory and other techniques (see report). This seemed appropriate, although estimated M's from the reference model were lower than the prior mean due to conflict in survey series. This is investigated further below.

The review panel (RP) concluded that data weights based on Francis reweighting methodology were appropriate. However, down-weighting data that seems to be more representative (Commercial live catch and recreational boat fleets) seemed counterintuitive as more weight is given to other data sources that are limited, i.e., Recreation Shore and Commercial Dead fisheries.

Other modelling issues covered:

a) The reference model used a prior on the von Bertalanffy (VonB) k parameter from Grebel and Calliet (2010) age study, but also used the Grebel age-length data directly in the model. That is double use of data. The RP recommended not using any priors on k to avoid this circularity. This issue was explored in the panel and shown to not change the reference model.
b) We discussed the shape of selectivity (domed vs. asymptotic). CIE reviewers argued that there should be at least one asymptotic fleet, which was the case in each reference model.
c) Jitters were iteratively redone until a global minimum was achieved and no model provided estimates lower than the values from the reference run.
d) The RP requested length comp diagnostics from a no-block run that was conducted, to get a better sense of why the model estimates different selectivities in blocks. AICs were provided along with these models, and supported that the selectivity blocking was warranted.
e) The RP was concerned if the reference model was sensitive to rare small and large lengths. We examined the total annual negative log-likelihood for fleet length compositions and we did not see large negative log-likelihoods associated with years with large Pearson residuals. Hence, it seems that sensitivities to such cases is not an issue.
f) It does not seem reasonable to use recruitment deviations historically or to generate non-equilibrium starting population size. However, it was not clear to the STAR panel what the problem is. Analysts explained that historic rec deviations looked unrealistic, but why? Why would SS3 estimate unrealistic recruitment deviations (that are penalized to be zero by sigma_R) to fit historic catch when it could be resolved in the annual F values?
g) The indices, particularly CPFV, favor a very low value of M. It was not clear to the RP why this occurred. The RP requested additional model fitting results to get a better understanding of what is going on. This was done (see below) but given the data weightings that index had low influence on the overall model estimate of M.

Requests by the STAR Panel for NCS and Responses by the STAT

STAR Request No. 1: Update reference models by re-weighting the alternative catch histories in the SCS and NCS models.

Rationale: These are more plausible catch histories by area.

STAT Response *Request No. 1:*

The NCS weightings are:

```
[1] "Francis Weights - len: FISHERY1: 1.0274 (0.6458-5.1035)"

[1] "Francis Weights - len: FISHERY2: 1.0118 (0.4674-3.7903)"

[1] "Francis Weights - len: FISHERY3: 1.0579 (0.6657-2.2036)"

[1] "Francis Weights - len: FISHERY4: 1.0237 (0.5528-2.8635)"

[1] "Francis Weights - len: SURVEY2: 1.0207 (0.6358-3.1814)"
```

SCS

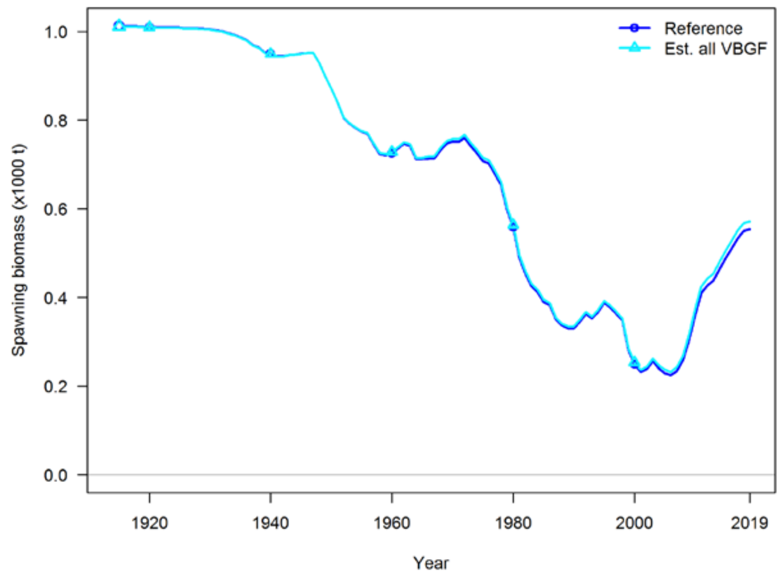
```
[1] "Francis Weights - len: FISHERY2: 1.0529 (0.5875-3.7356)"
[1] "Francis Weights - len: FISHERY3: 0.8999 (0.5561-2.0078)"
[1] "Francis Weights - len: FISHERY4: 1.1682 (0.8303-1.9336)"
```

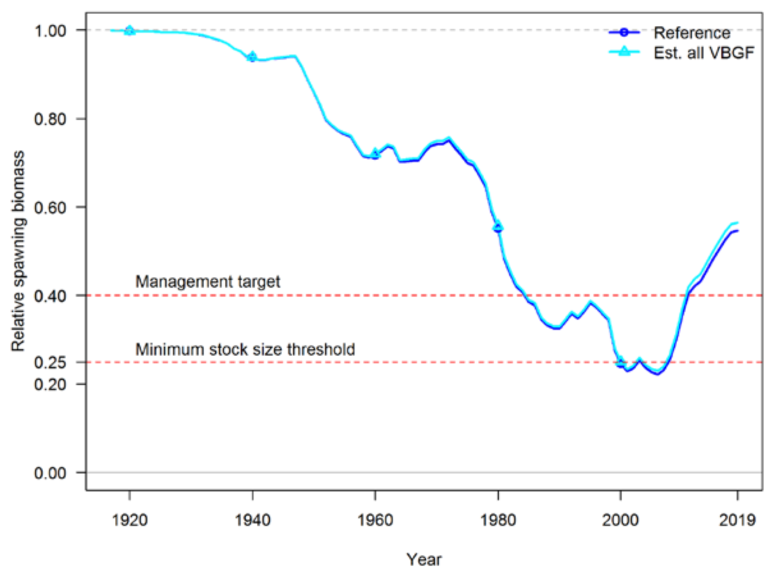
Conclusion: No need to revise weighting at this time.

STAR Request No. 2: Provide a run with uninformed priors on growth curves.

Rationale: These data are in the NCS model and the prior should not be used.

STAT Response *Request No. 2:*





Conclusion: Marginal differences in assessments as priors on k have little influence. The reference model should proceed without priors on k.

STAR Request No. 3: Produce r4ss plots for the no time blocking run on selectivity and calculate AIC for both models.

Rationale: To better understand the residual patterns with the time blocking models.

STAT Response *Request No. 3:*

AIC = 2p+2(-lnL)

Reference: 2*112+2*594.638 = 1413.276

No selectivity block: 2*101+2*645.121 = 1492.242

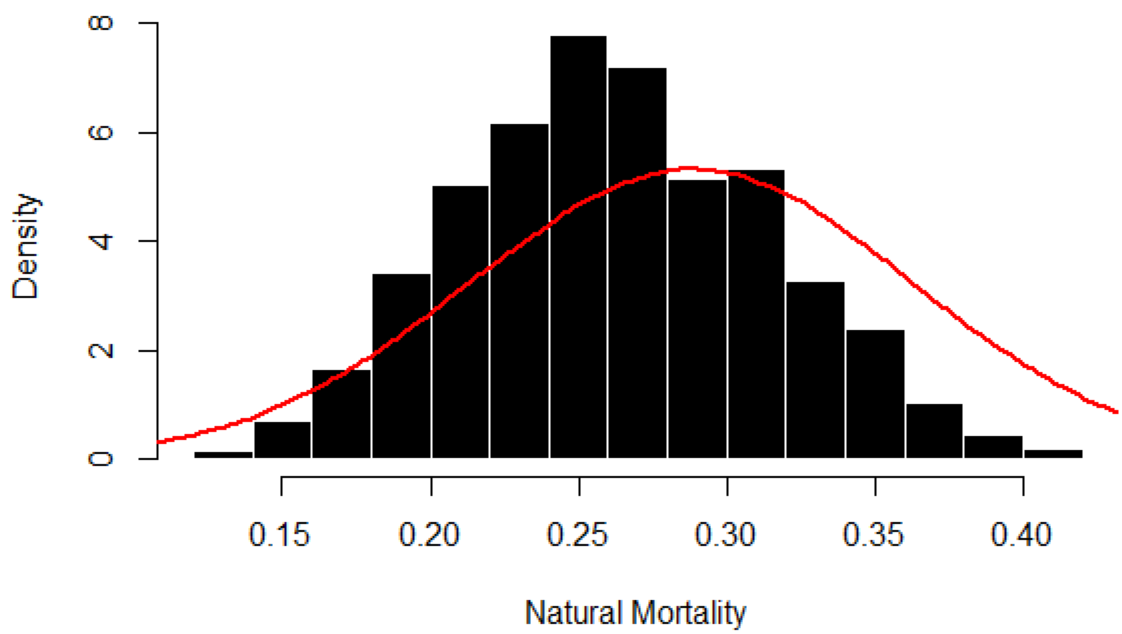
r4ss plots were provided for specific comparisons.

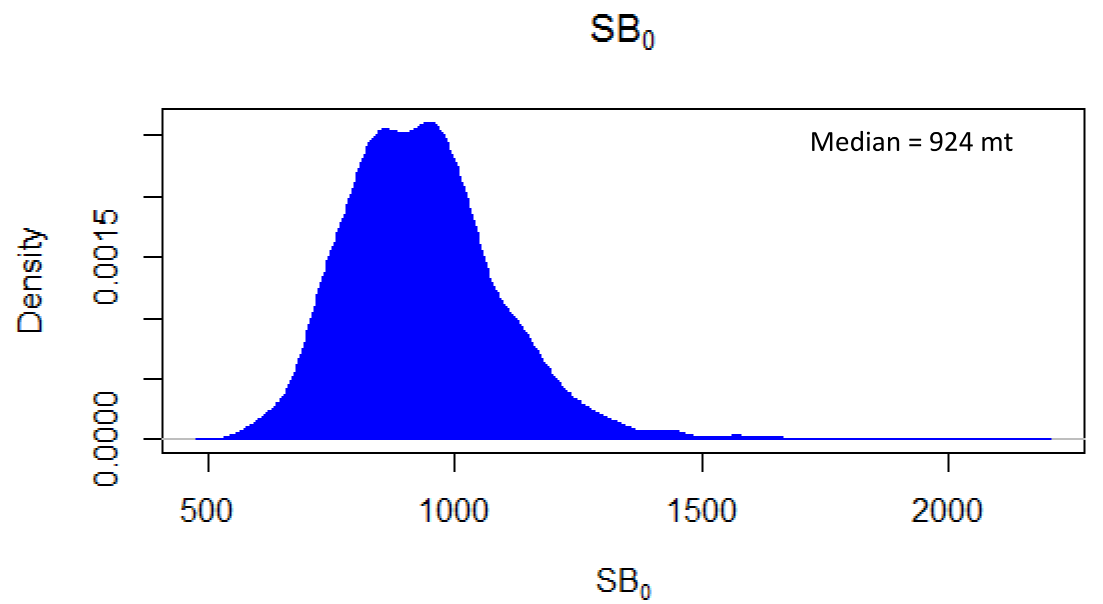
Conclusion: The blocking of selectivity is appropriate.

STAR Request No. 4: Provide comparative plots of prior and posterior values of M. Also provide posterior plots for five derived sensitivity quantities (B_0, B_{2019}, depletion, F $_{SPR}$, MSY at SPR).

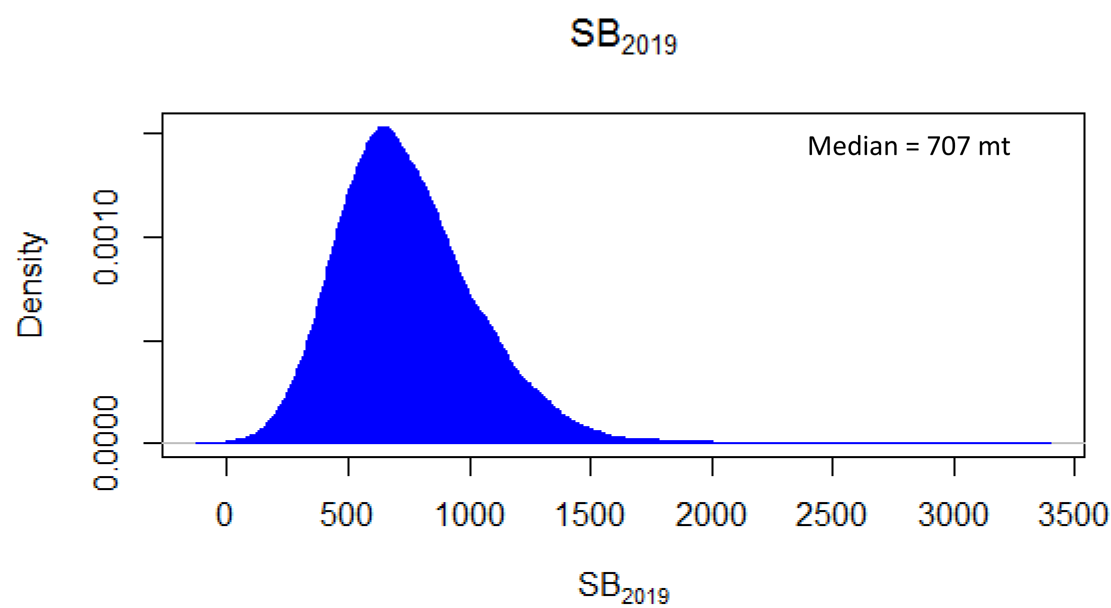
Rationale: To see the data influence on estimating M and to understand the Bayesian distributions on the derived quantities.

STAT Response *Request No. 4:*

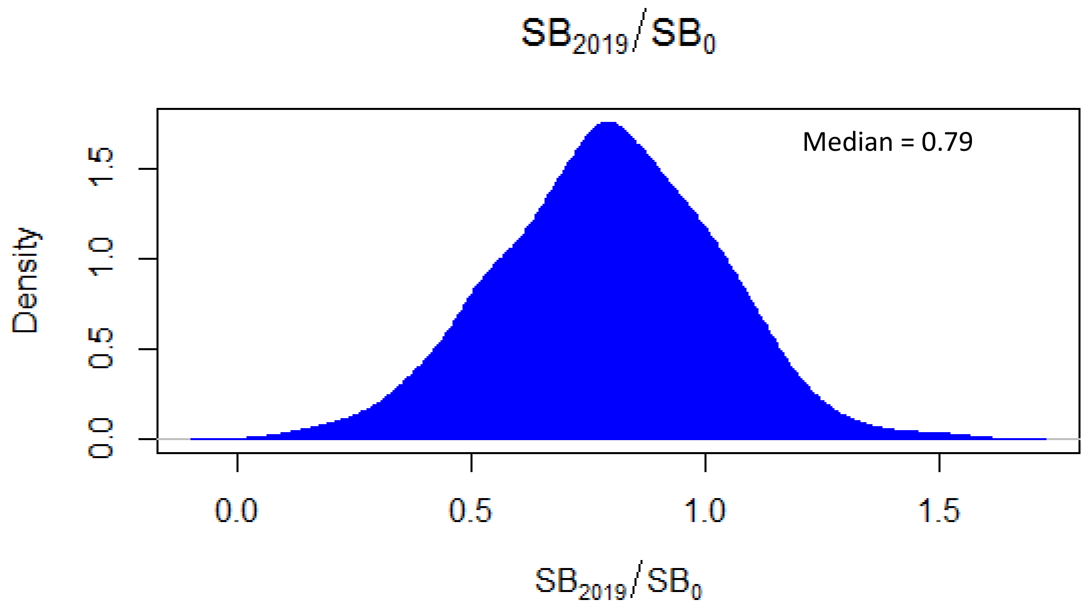




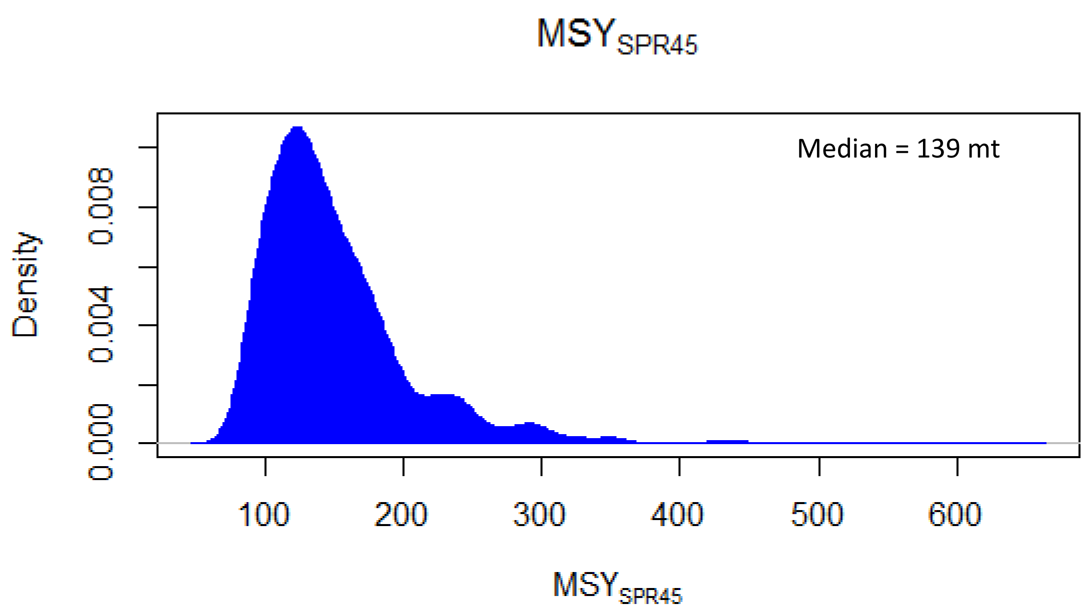
SB_0
Median = 924 mt
Density
0.0000
0.0015
500
1000
1500
2000
SB_0



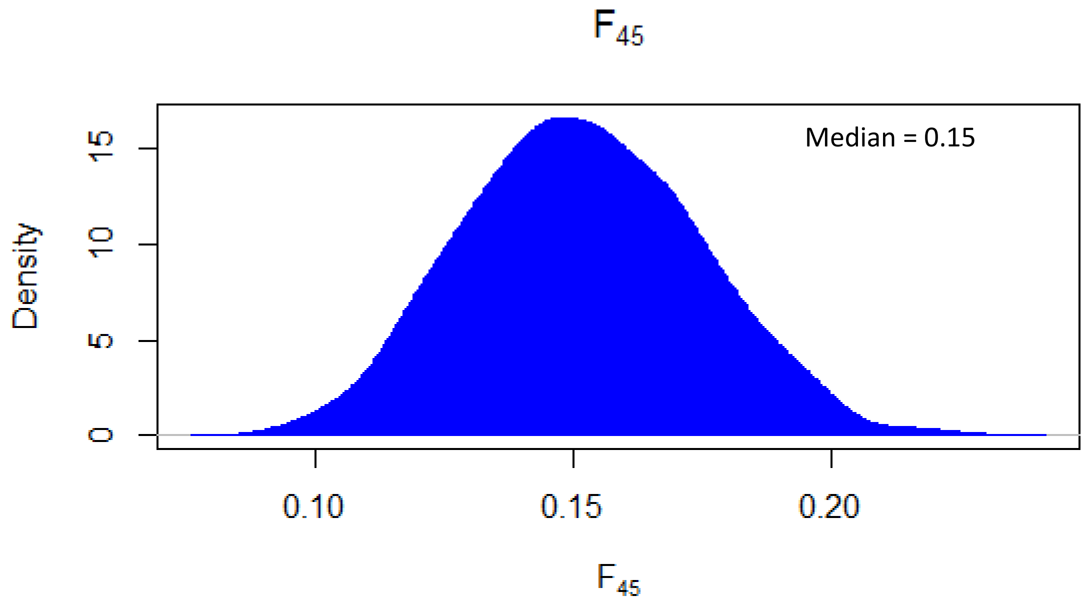
SB_{2019}
Median = 707 mt
Density
0.0000
0.0010
0
500
1000
1500
2000
2500
3000
3500
SB_{2019}



SB_{2019}/SB_0
Median = 0.79
Density
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
SB_{2019}/SB_0



MSY_{SPR45}
Median = 139 mt
Density
0.000
0.004
0.008
100
200
300
400
500
600
MSY_{SPR45}



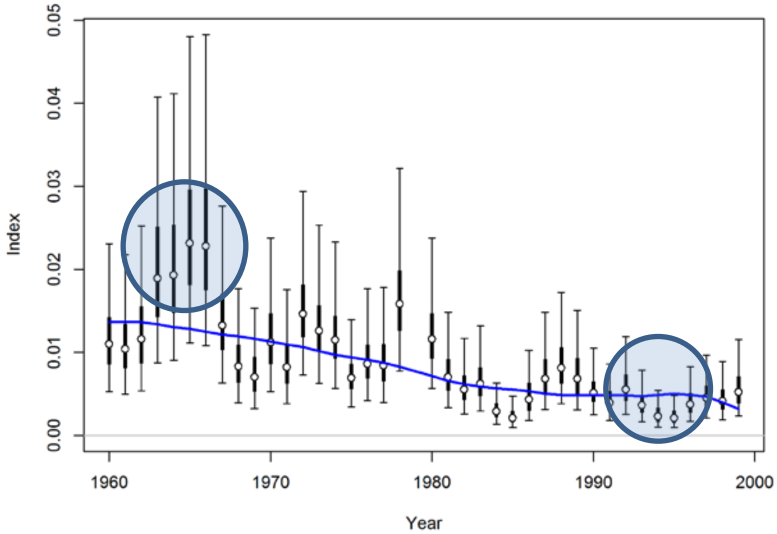
Conclusion: The posteriors are approximately symmetrically distributed around the median (except MSY SPR0.45). The data is shifting the posterior distribution of M lower than the prior, but posterior mean value appears to be biologically reasonable.

STAR Request No. 5: Provide a low M run (M = 0.1) and provide diagnostics.

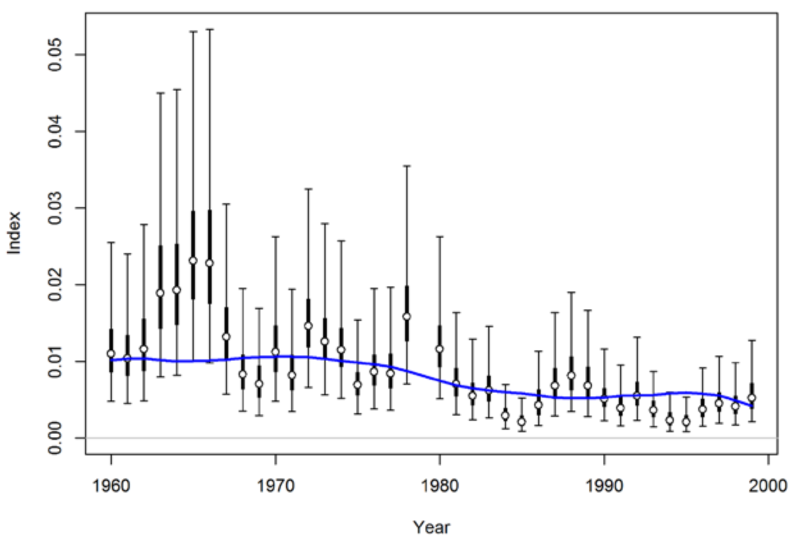
Rationale: To better understand why a low M model better fits the CPFV survey.

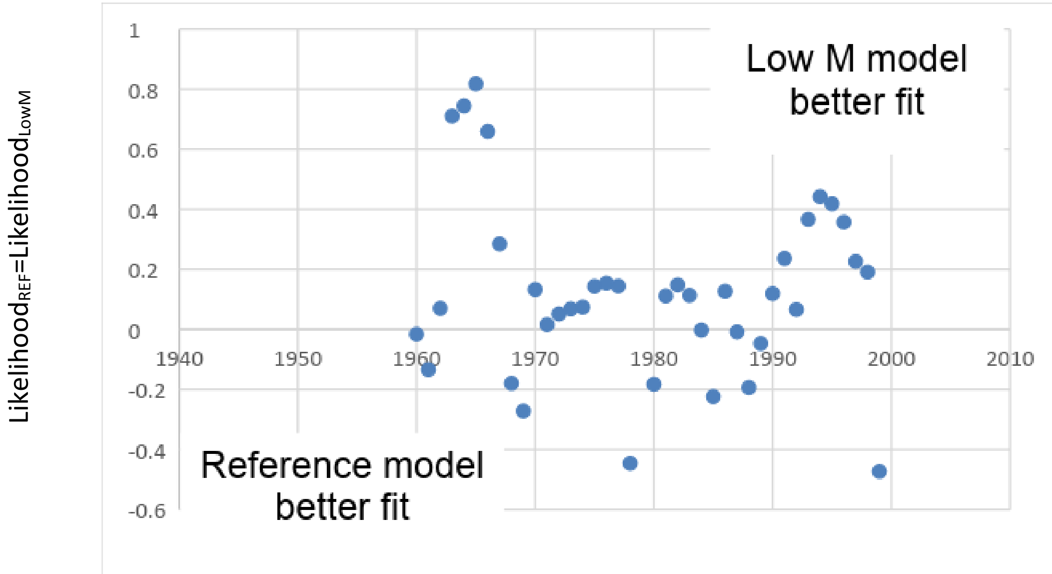
STAT Response *Request No. 5:*

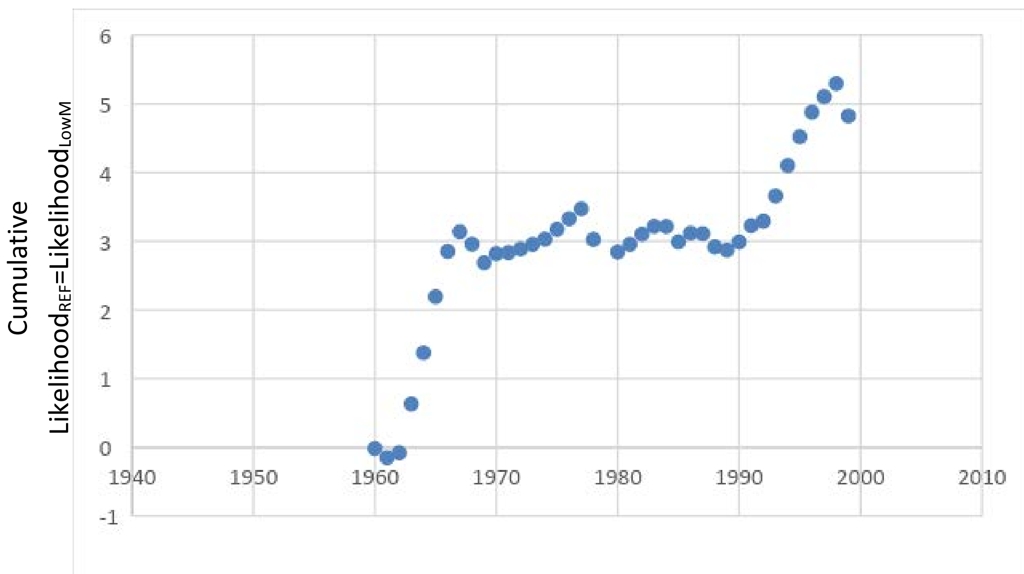
Low M model CPFV index Likelihood -17.09



Reference CPFV index Likelihood -12.27







Conclusion: The low M run resulted in a substantially lower negative log-likelihood for the CPFV index. The improvement in fit to that index was concentrated on two time-periods.

STAR Request No.6: Investigate alternative model formulations that address overfitting of the CPFV survey with low M.

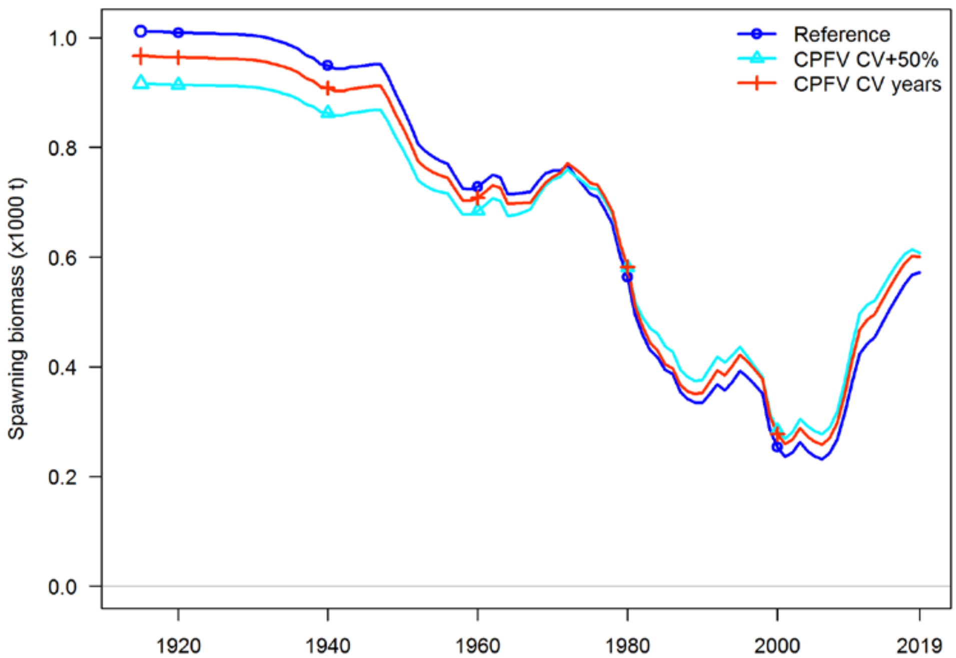
Rationale: To better understand why a low M model better fits the CPFV survey.

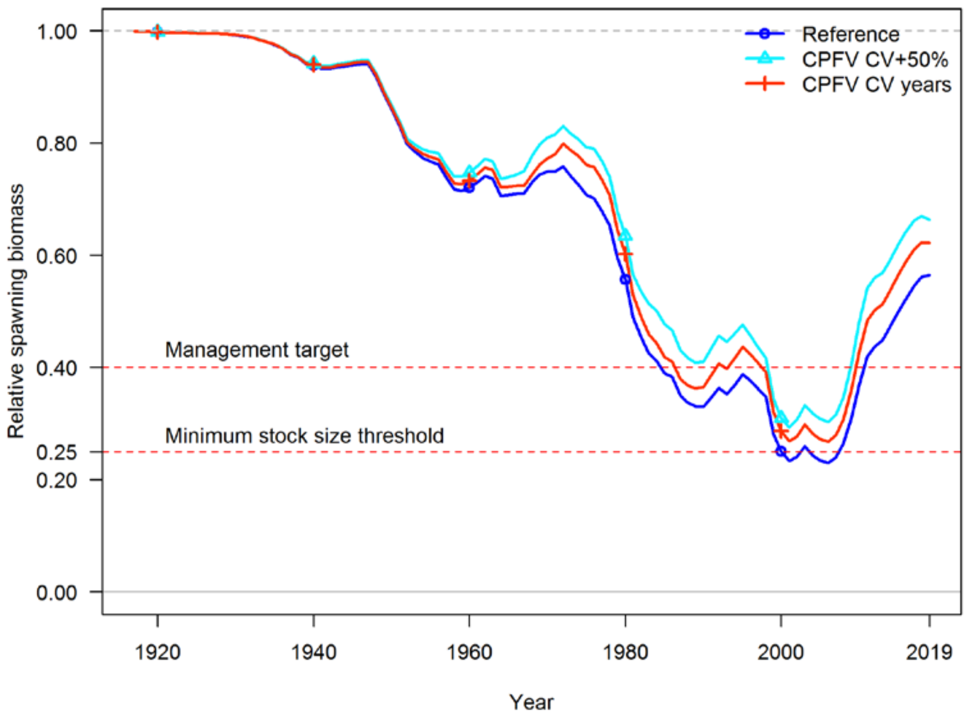
STAT Response *Request No. 6:*

Two models were explored:

1. Add a fixed 50% to CPFV CV (as opposed to the estimated 30%).
2. Add 20% to the years 1963-1967 and 1993-1998 (years with the biggest differences in likelihoods).

	Estimated M	
Model	Female	Male
Reference (+ 30% CV)	0.214	0.256
1)Add 50% CV	0.256	0.287
2) Add CV to years	0.234	0.275





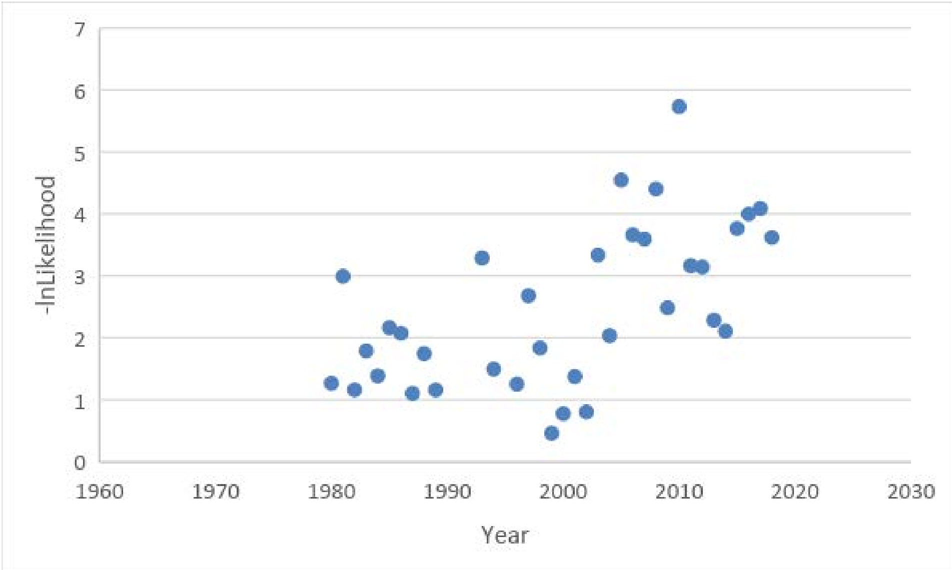
Conclusion: All reductions in weights for the CPFV index resulted in higher M estimates. However, these models did not result in much difference in SSB or status estimates. The down-weighting seemed too subjective and the RP recommended using the reference model weights; however, we should document that the apparent assessment information about M in the CPFV index does not seem real.

STAR Request No.7: Provide a plot of yearly length composition likelihood values for the recreational boat fleet in the NCS model.

Rationale: To better understand the residual pattern for these length compositions.

STAT Response *Request No. 7:*

NCS Recreational Boat –log likelihood values by year. Red dots are huge residuals



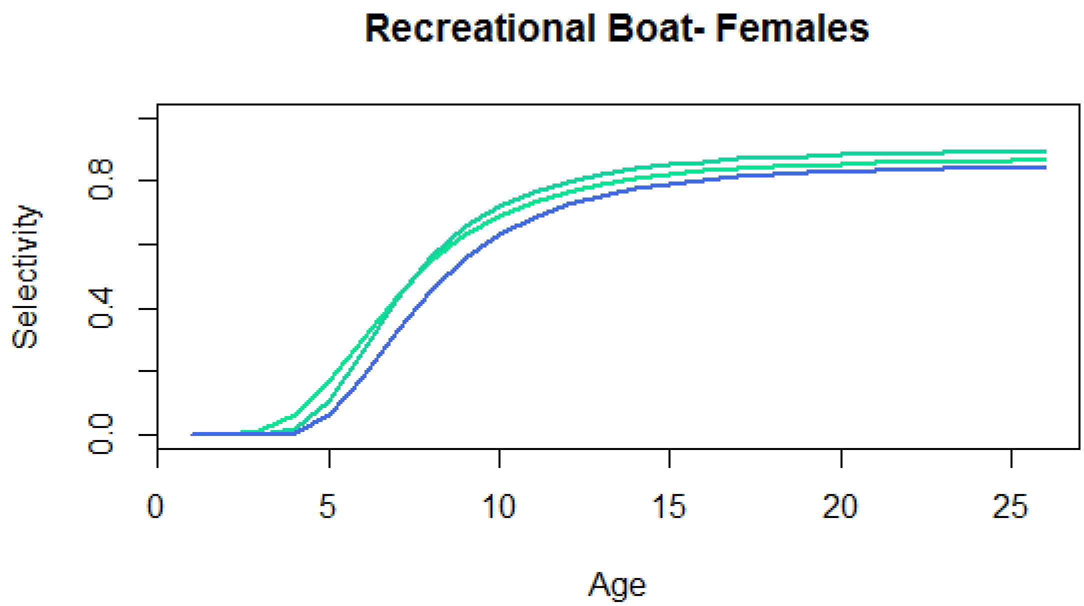
Conclusion: Length composition likelihoods get worse over time.

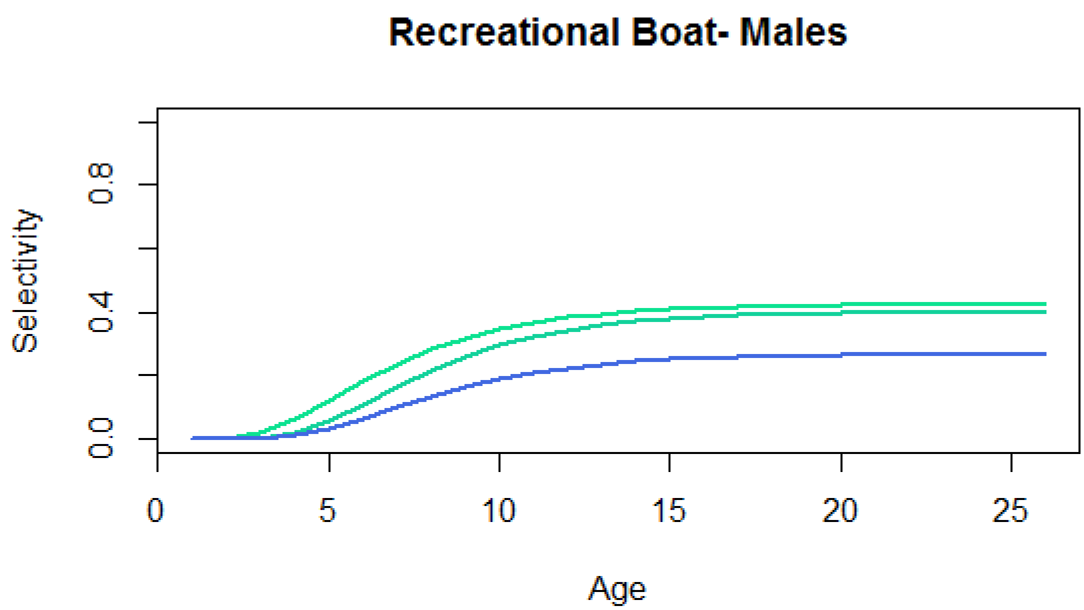
STAR Request No.8: Provide a plot by year of the age selectivity 2 vector.

Rationale: To confirm that the Asel2 vector changes over time.

STAT Response *Request No. 8:*

Figures below show recreational boat derived-age selectivity plotted for all years for females and males.





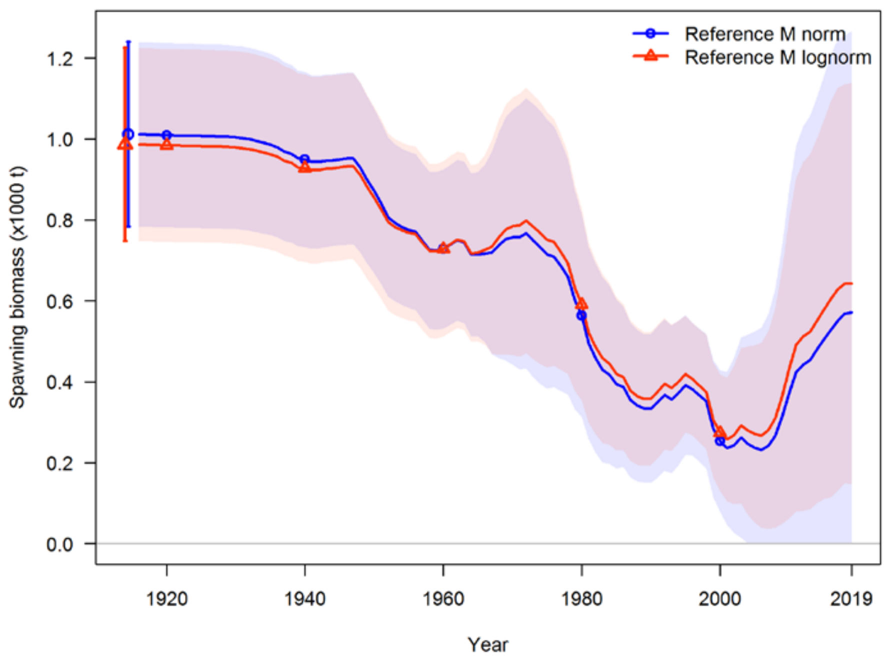
Conclusion: Selectivity at age is a function of selectivity at length multiplied by length at age for selectivity at age. This changes as selectivity is time blocked for 3 periods.

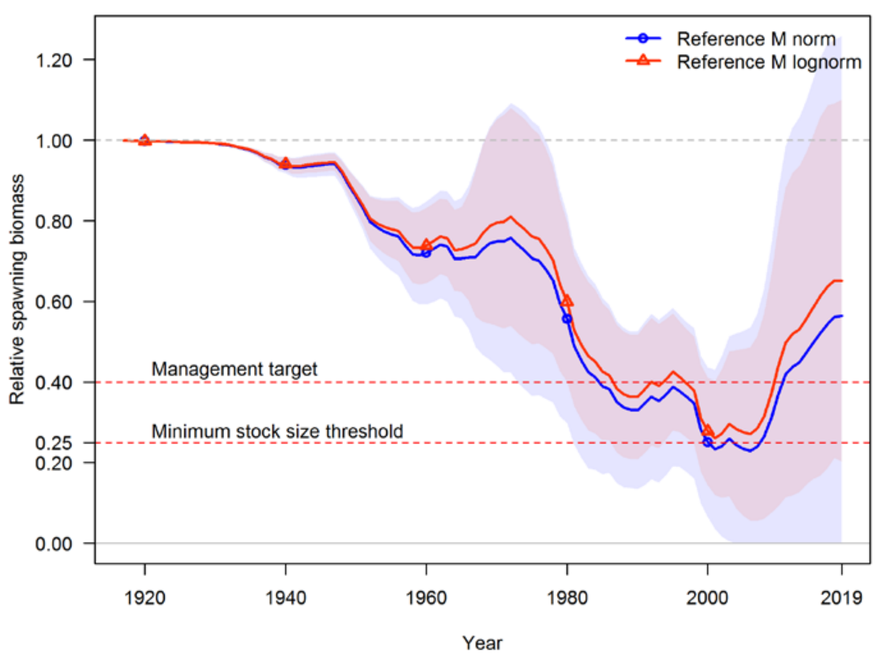
2nd Round of Requests for the CA Cabezon STAT Regarding the NCS Model

3. Recalculate the M prior as a lognormal and present the model fits and diagnostics.

Rationale: To explore the possibility of a traditional decision table that does not require an MCMC approach.

NCS





4. Run two models: fix female Ms at 0.2 and 0.33 (these values come from the Bayesian posterior distribution shown to the panel). Alternatively, use the male M offset from the reference model and estimate the male M offset. Provide model fits and diagnostics.

Rationale: To explore possible low and high states of nature for the decision table.

Response: Unnecessary given results of request #1.

SCS Assessment

Data sources

- Catch and length composition from three fisheries: commercial live-fish, recreational boat (both charter and private boats; length from the charter boat fishery only), and recreational shore (though length compositions were not available from all fisheries in all years). There was also a small amount of catch (but not length data) from the commercial dead-fish fishery.
- A relative abundance (CPUE) index based on the recreational charter boat fishery (California Passenger Fishing Vessel, CPFV, a fishery-dependent index).

Notable updates to data sources

- As mentioned above for the NCS model, all age data were now included in the NCS model only, as the very few samples that were in the past SCS model were not used in estimation. Consequently the growth parameters estimated in the NCS model were then fixed for use in the SCS model.
- Historical catch times series remained largely the same from the last model, with new catches being added to complete the time series. One difference came in the new pull of recreational data from 2004-2008.
- A more significant change in the catch data was the reallocation of catches from the SCS region to the NCS region for years 1980-1995. This was in response to the recognition that past treatment of MRFSS catch estimates were based on stratification of California at 36° N lat. as opposed to 34°27' N lat. at Point Conception, as the assessment structure assumes. This change was made after the pre-STAR draft assessment was prepared, but prior to the STAR panel, so this updated model was presented to the panel. The effect of this change was proportionally greater for the SCS, where there is a smaller catch overall than in the NCS.
- A change was made in the assumed timing assigned to the catches. The previous assessment assigned catch to the month 1, whereas the new assessment assigned them to the mid-year 6th month.
- Mean weights were excluded from the new assessment. The length of the time series of mean weights had been reduced from the last assessment with the recovery of some true length measures in the early MRFSS time period. Once that happened the final reference models showed low information content in the remaining mean weights, justifying removal from the current model.
- Length compositions again were similar to the previous assessment, with the addition of

subsequent years. Month assignment was switched from 1 to 6, as done in the catches.

Model details (including details changed from the previous 2009 assessment model):

- All model details are as in the NCS description above, with the exception of the use of CCFRP data and the details noted below.
- As in the previous assessment, growth was estimated in the NCS model and the NCS estimates were fixed for use in the SCS model.
- Recruitment estimation differed as the current assessment used the method of Methot and Taylor (2011) to identify years of recruitment estimation and the treatment of bias adjustment to make it more consistent with the assumed recruitment variability. The previous assessments assumed all estimated recruitment years received a full bias adjustment (equal to 1), with years of estimated recruitment 1970-2006 in both models. The current SCS model estimated recruitments from 1970-2016, with the ramp from 0 in bias adjustment starting in 1970 and reaching its maximum value of 0.45 from years 1977-2011, ramping again down to 0 in 2017.

Data Issues

The datasets covered examined both catch, length composition and standardized indices of abundance. While M was kept free, other issues related to model mis-specification in terms of selectivity for the recreational fleet were discussed. In general there appeared to be very low numbers of sampled length composition data, and too much weight was being put on very small samples in the recreational shore based length composition data. Other issues related to the datasets and structural uncertainty examined are covered below:

1. Recruitment deviations – several big peaks appear to be associated with La Niña conditions – 1984, 1987, 2001. The deviations from that pattern are in 2000 and in 2010. From 2010 the recruit deviations get very autocorrelated. The analysts argued that information is absent in that part of the time series and it is part of the ramp down in the recruitment deviation bias correction that contributes to lack of certainty on current stock status. The CV's in the recruitment deviates are large and they are incorporated in the uncertainties of stock summary statistics. However, some thought needs to be given as to how CV's on recruitment deviates are estimated as time periods with little information are considered to be more precise by SS3 than time periods with data. Hence, the status evaluation is expected to be highly uncertain because of a lack of data, especially recently.
2. Covariation in the growth parameters k and L_∞ needs to be explicitly accounted for in the priors
3. In the SCS model, the 1984 year class is strong and there is evidence for this in the recreational boat length compositions and in the CPFV index. Assuming the CPFV index reflects recruitment variation, then there is evidence of an early large year class (evidenced by increase in the CPFV index in the early 1960s) that the model cannot account for because the model does not include recruitment deviations early in the time series. The STAT suggests that the CPFV cannot by itself detect a strong year class, but can support other data sets, so the early 1960s pattern may not be a real recruitment signal.

4. The combination of recreational shore (last assessment 2 fleets) and boat (2 fleets) compared to 2009 assessment seemed to make sense.

. Error bounds span the limit reference point and should be assessed as to how to capture the uncertainty correctly. MCMCs were more optimistic in their behavior with respect to relative biomass trends.

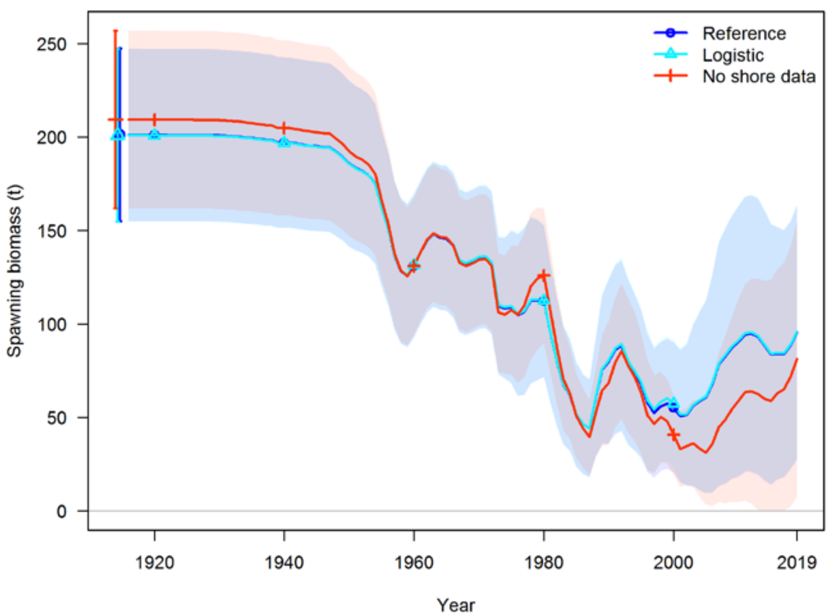
Requests by the STAR Panel for SCS and Responses by the STAT

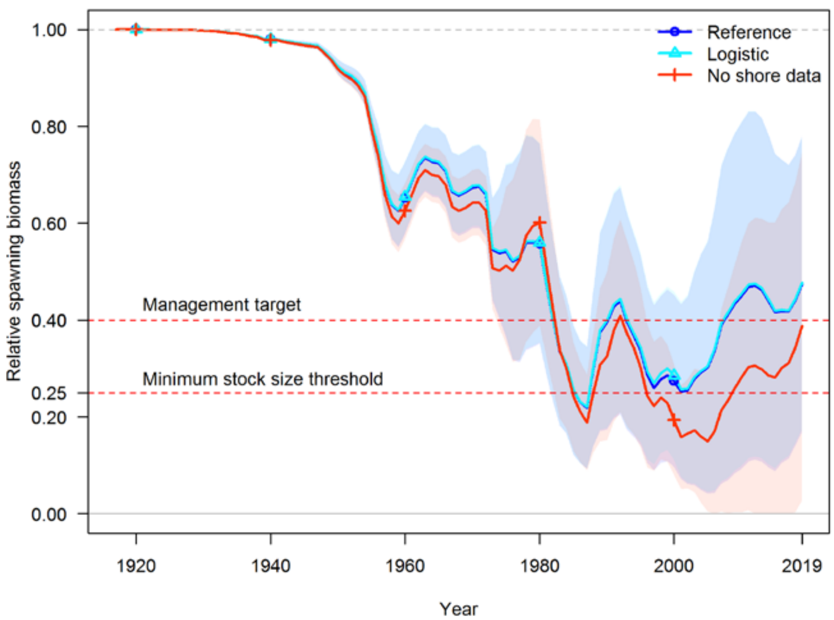
1st Round of Requests for the CA Cabezon STAT Regarding the SCS Model

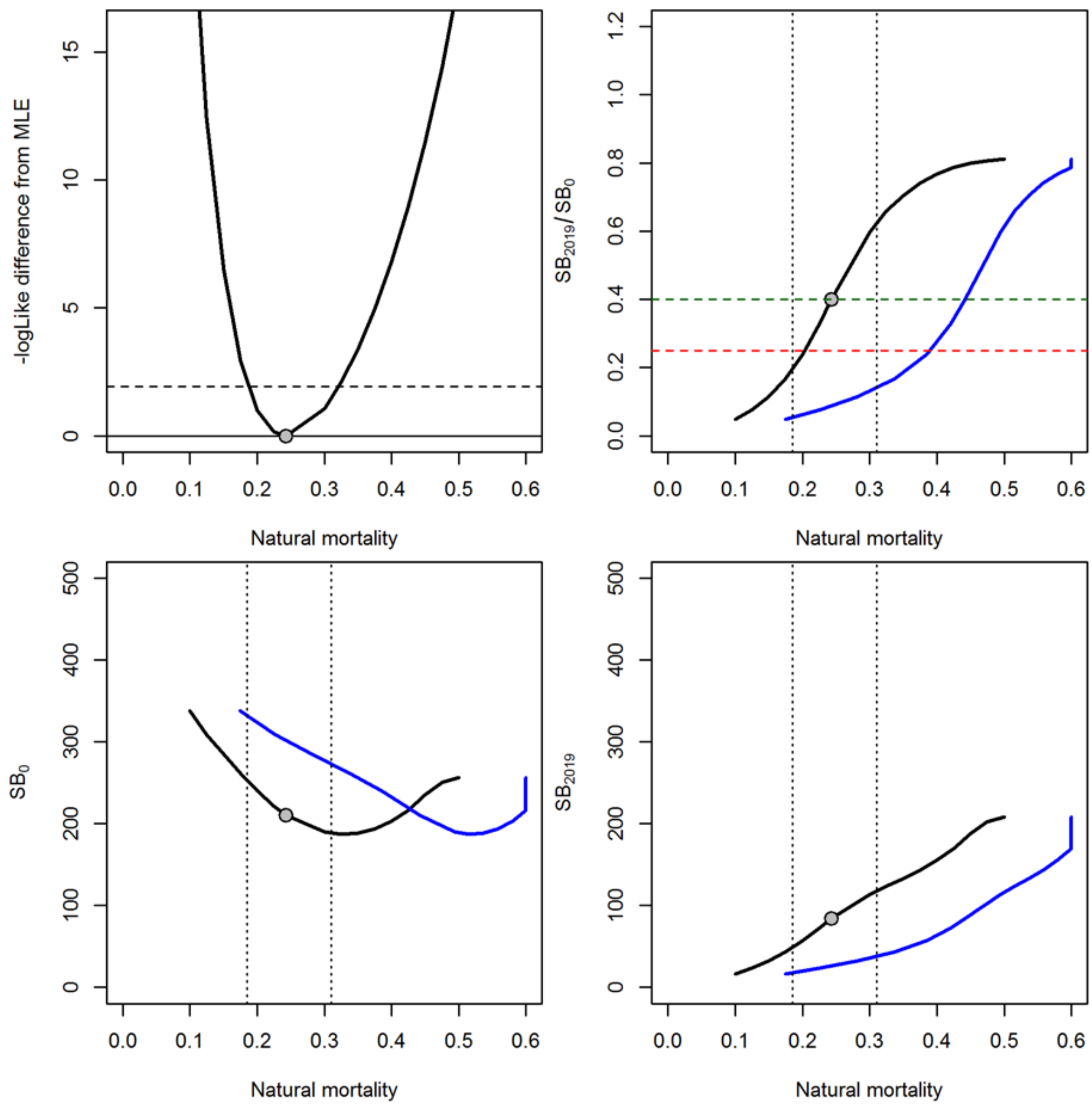
1. Assume the alternative catch histories for the SCS (and NCS) models and then:
   a) force the rec. shore fleet to have logistic selectivity;
   b) fix the selectivity to the outcome in 1a; and
   c) remove the rec. shore length compositions. and redo the M profile.

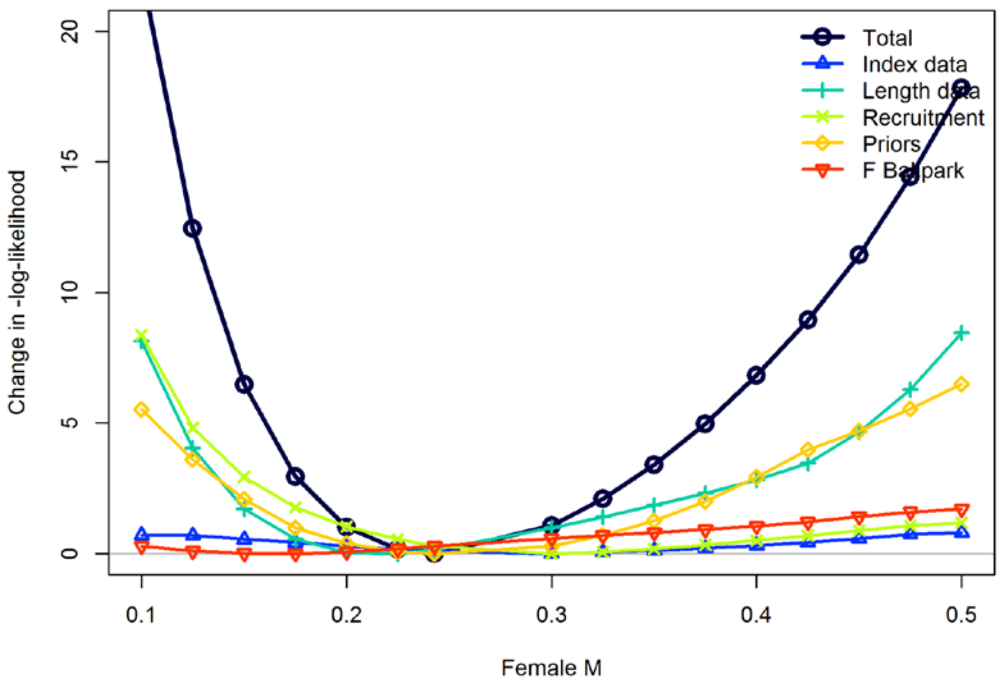
Rationale: To explore the effect of rec. shore length compositions. on the M estimate.

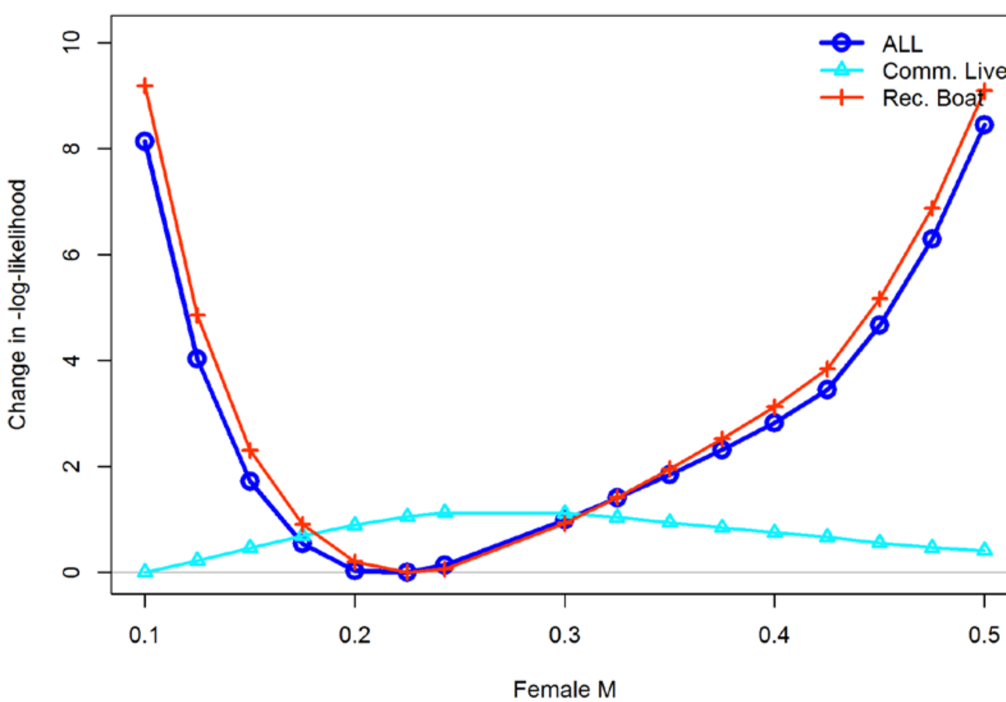
Responses by STAT:




Reference
Logistic
No shore data
Relative spawning biomass
1.00
0.80
0.60
0.40
0.25
0.20
0.00
Management target
Minimum stock size threshold
1920
1940
1960
1980
2000
2019
Year


-logLike difference from MLE
Natural mortality
SB_{2019}/SB_0
Natural mortality
SB_0
Natural mortality
SB_{2019}
Natural mortality


Total
Index data
Length data
Recruitment
Priors
F Ballpark
Change in -log-likelihood
Female M

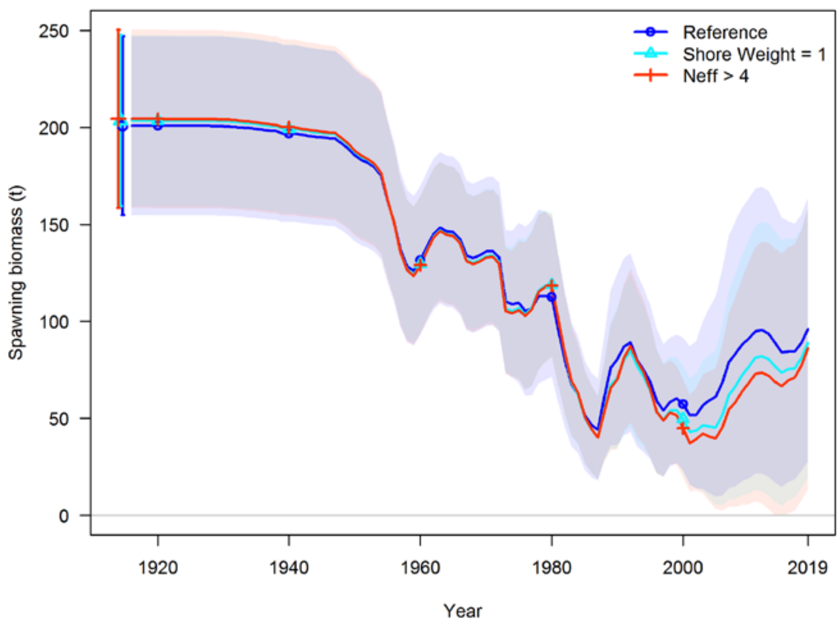

ALL
Comm. Live
Rec. Boat
Change in -log-likelihood
Female M

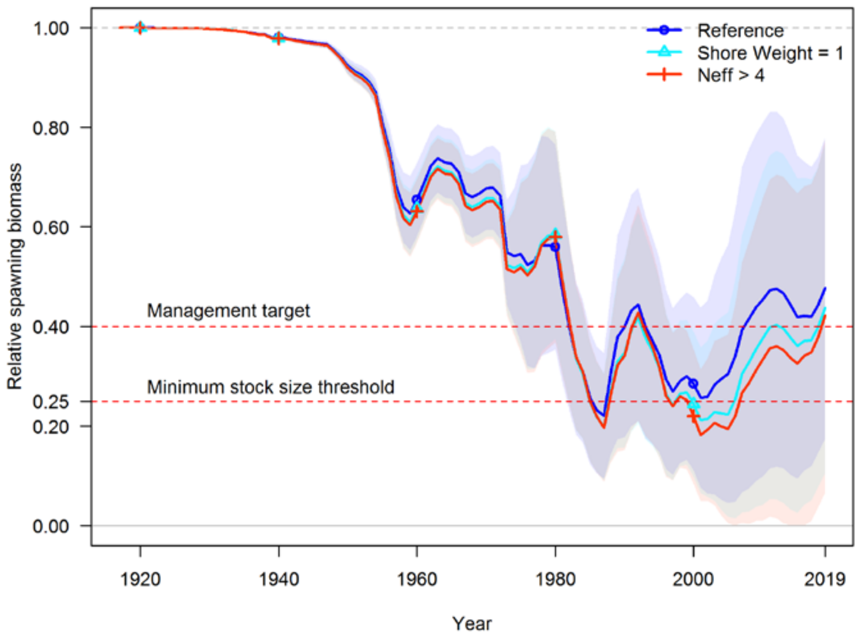
Conclusion: While rec shore length compositions are influential, they do not affect the overall M profile.

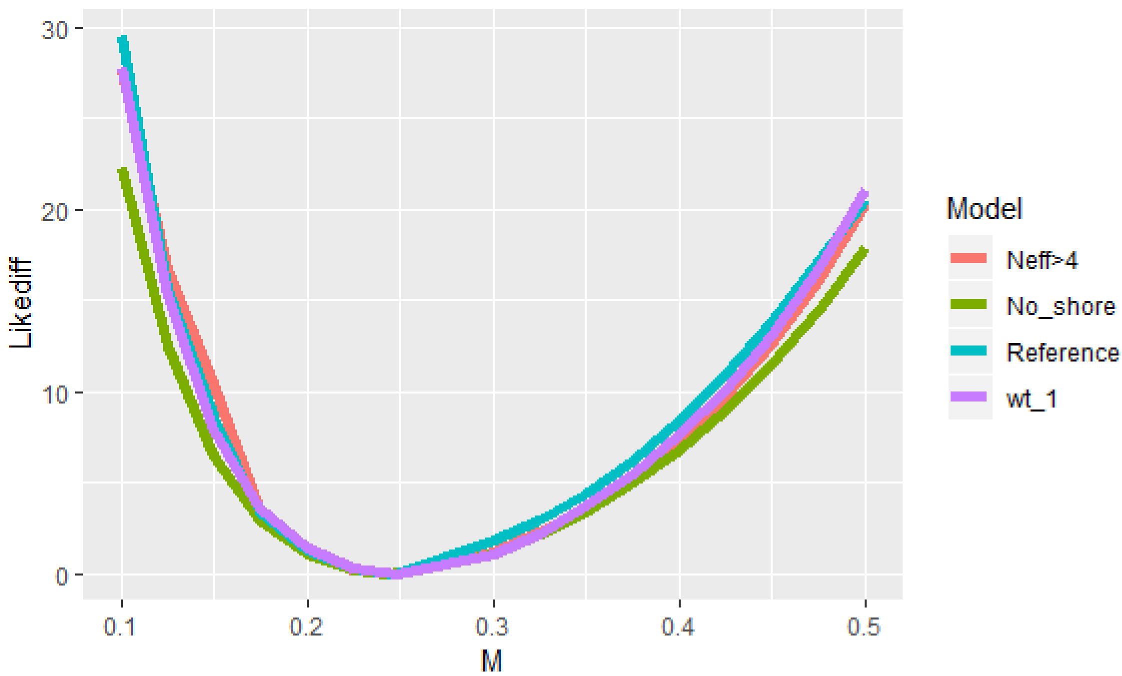
2. a) Cap the data weighting for the rec. shore fleet to 1; b) modify effective N in the rec. shore fleet by removing samples less than 5 or only exclude samples with very tight intervals; c) redo the M profile; and d) provide the comp. plots and diagnostics.

Rationale: Explore how much influence the rec. shore fleet has on model results by down-weighting these data.

Response:




Reference
Shore Weight = 1
Neff > 4
Management target
Minimum stock size threshold
Relative spawning biomass
1.00
0.80
0.60
0.40
0.25
0.20
0.00
1920
1940
1960
1980
2000
2019
Year

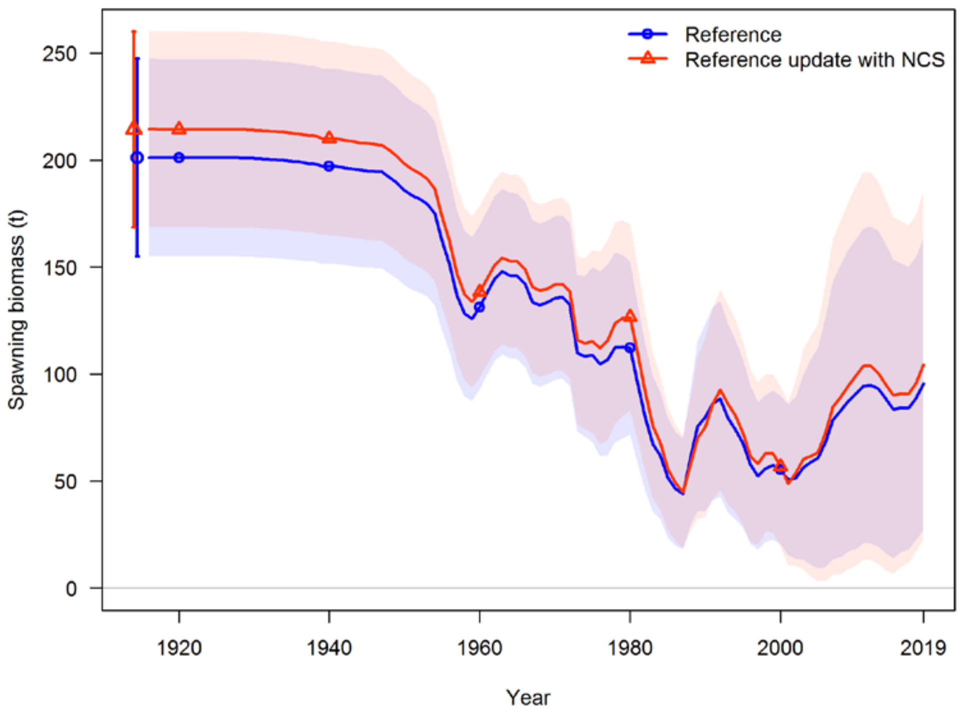


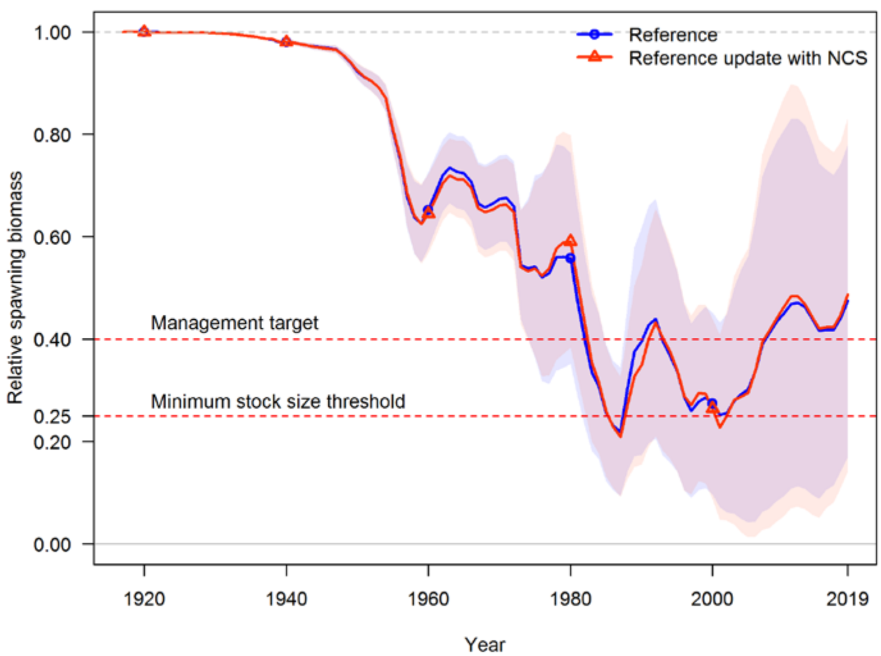
Model	Estimated M Femal e	Male
Reference	0.256	0.395
No shore weight	0.24	0.45
Shore weight = 1	0.248	0.426
Shore Neff > 4	0.246	0.44

Conclusion: While rec shore fleet has an effect it is quite marginal on estimates of M

3. Recalculate the M prior as a lognormal and present the model fits and diagnostics.

Rationale: To explore the possibility of a traditional decision table that does not require an MCMC approach.





Conclusion: The use of the log normal M fixes the issues with unrealistic stock status trajectories with unrealistically low M's creating low stock status conditions.

OCS assessment

Data sources

- Catch and length composition from four fisheries: commercial dead-fish, commercial live-fish (catch only, no length), recreational boat (both charter and private boats; length from the charter boat fishery only), and recreational shore.
- Fishery-dependent relative abundance (CPUE) indices based on commercial logbooks, an onboard observer program on recreational charter boats, and two dockside recreational surveys, the Marine Recreational Fisheries Statistics Survey (MRFSS) and the Oregon Recreational Boat Survey (ORBS).
- Conditional age-at-length data from a scientific research survey.

Notable updates to data sources

- An accounting error was discovered in the weeks leading up to the STAR panel. The STAT resolved this issue prior to the STAR and all presented results during the STAR week were using the updated catch data. The error was due to double counting discards for the recreational ocean boat fleet during the period 2001-2018. This amounted to a very small amount of the total catch (~1%) and very little change to overall model results.

Model details (including details changed from the previous 2009 assessment model):

- Model time coverage was moved to 1970 (previously 1973) and a linear ramp of recreational ocean boat catch from 1970 to 1973 and shore catch from 1970 to 1979 was used rather than estimating an initial fishing mortality parameter.
- Two sexes are retained, as growth is very different between females and males.
- Population length bins spanned 4 cm to 70 cm (previously 6 cm to 92 cm) and the accumulator age was set to age-20 (previously age-35).
- The yearly time step with 12 months is retained, though 6 subseasons were defined in order to allow for more flexibility in the treatment of fleets.
- Updated female and male weight/length relationship using additional data.
- Two commercial (live and dead) and two recreational (ocean boat and shore) fleets were specified. The shore fleet combined the man-made jetties and beach/bank modes, and the ocean boat fleet combined the private and charter boat modes. Model development showed similar length compositions between the modes. Model development showed similar length compositions between the modes.
- Updates to the historical recreational catch time series were made following methods outlined in the assessment document (section 2.4.2). Catch and length and age composition data was updated to 2018.
- Mean weights were excluded from the new assessment. Mean weights had been reduced last assessment with the recovery of some true length measures in the early MRFSS time period. Once that happened the final reference models showed low information content in the remaining mean weights, justifying removal from the current model.
- Selectivity curves were treated similarly to the previous assessment: the commercial dead and boat fleets were estimated as asymptotic; the commercial live fishery was allowed to be dome-shaped (though went asymptotic during one time block). Time-varying blocks were applied to the commercial live and the recreational boat fleets.

- Block years for selectivity were slightly adjusted in the new model to better match changes in the length compositions and known management changes. Numbers of blocks changed from 3 to 2, because of indistinguishable differences between two of the previously specified time blocks (i.e., drop one of the previous time blocks; 2000-2003)
- An additional three indices were developed and used in the reference model, compared to one index in the previous model.
- Natural mortality was fixed for both females and males at the value indicated by the longevity-based Hamel prior, given the difficulty in estimating natural mortality internal to the ORS model. This was changed during the STAR panel (see Description of the Base Model, below)
- Growth was again estimated in the ORS model. One difference is that the current ORS model estimated the length at age 0 to be close to 0 for both sexes, so that parameter was subsequently fixed, which improved model estimation.
- Additional biological parameters were fixed to the same values as in the previous model.
- Steepness and recruitment variability were fixed to the same values as in the previous model.
- Recruitment estimation differed as the current assessment used the method of Methot and Taylor (2011) to identify years of recruitment estimation and the treatment of bias adjustment to make it more consistent with the assumed recruitment variability. The previous assessments assumed all estimated recruitment years received a full bias adjustment (=1). The current ORS model estimated recruitments from 1980-2016, with the ramp from 0 in bias adjustment starting in 1964 and reaching its maximum value of 0.63 from years 1983-1998 (years of peak information), ramping again down to 0 in 2017. The current SCS model estimated recruitments from 1970-2015, with the ramp from 0 in bias adjustment starting in 1984 and reaching its maximum value of 0.7 from years 1984-2015, ramping again down to 0 in 2016.
- Change from Pope's approximation of F to the hybrid method.
- Both the previous and current model analytically calculated the catchability coefficient for each survey. Additional variance was also estimated for the ORBS dockside. The model did not estimate additional variance (even when allowed to) for the other three indices.
- Data weighting was changed from the previous assessment that used the harmonic mean (McAllister and Ianelli 1997) approach for all composition data to using the Francis (Francis 2011) approach for length composition data and the harmonic mean approach for conditional-age-at-length data in this assessment. The Francis method was initially attempted for the conditional-age-at-length data, but resulted in extremely over-weighting the commercial dead data (>7), thus the harmonic mean approach was used.

Other issues identified are noted below:

1. von Bertalanffy fits to external growth estimates (Rasmuson study are underestimating length at high age – L_inf seems to be too low?). This could reflect a selective fishery that targets faster-growing young fish. The data could be fit externally and compared with what

the base model estimates to make sure some other piece of the objective function is not biasing this fit.

2. 1978 recreational ocean landings seemed high, and thus there is a need to check this.
3. Length composition data is not fit well for all fleets; i) the commercial live length compositions seem to have residual patterns, 2001/2011 seems to have strange data, ii) Commercial dead length compositions may have similar data problems, iii) the recreational Ocean fleet has blocks of length composition residual patterns, and iv) the recreational shore fleet may have the data problem.
4. The bridge model revealed some issues with the change to the new version of SS3. A run with parameters fixed to the earlier assessment values produced identical dynamics. However, there seemed to be a problem when SS3 was used in parameter estimation because a different stock trajectory was produced.

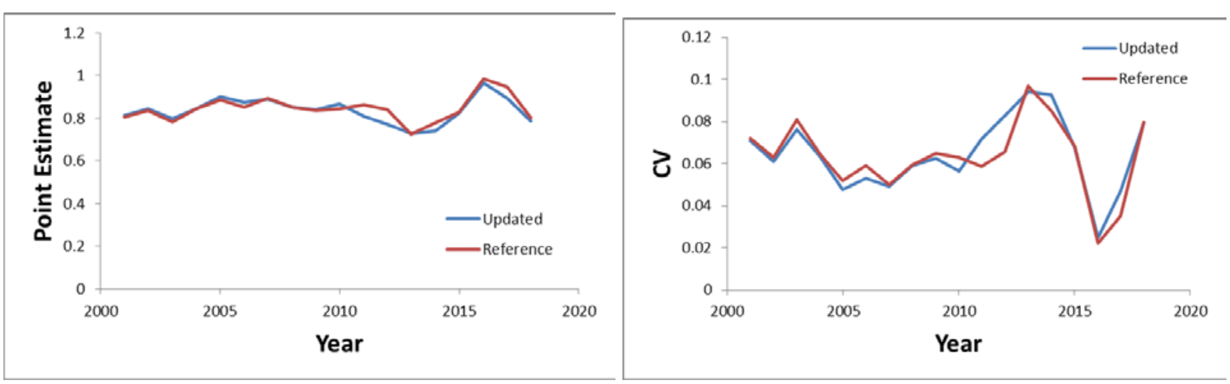
Requests by the STAR Panel for ORS and Responses by the STAT

1st Round of Requests for the OR Cabezon STAT regarding the ORS model

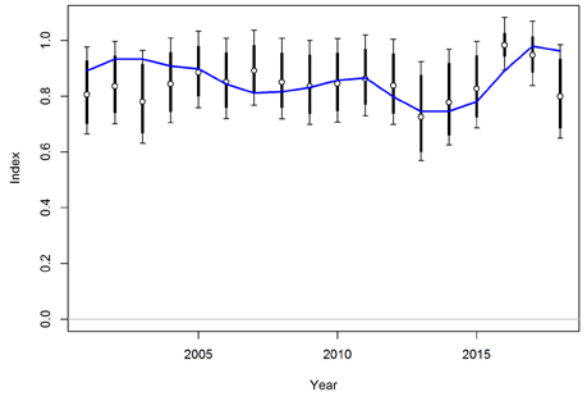
1. Remove the year:boat type interaction term in the ORBS index and re-compute the index. In the short term, assume the CVs from the current configuration (recalculate CVs for the post-STAR version of the base model). Compare the resulting indices and synthesis results.

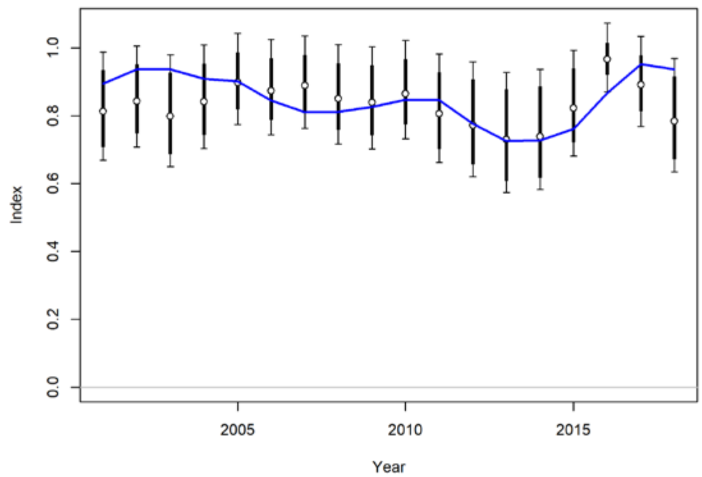
Rationale: The year:boat type interaction was not properly accounted for in the current index.

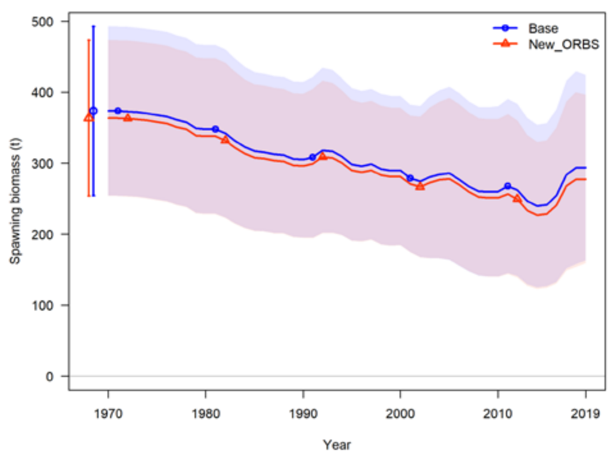
Response: The index was recalculated (both point and uncertainty, CV, estimates) and the model re-run and is a straightforward improvement to the index itself. The area-weighted index changed only slightly and model results are insensitive to the adjusted index.
Difference in point estimates and CVs for the two index versions.

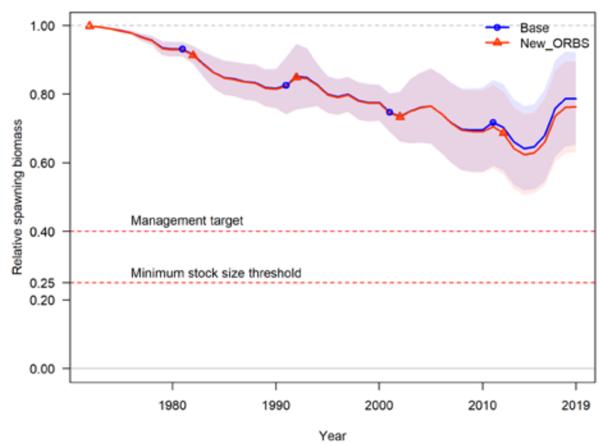


Top figure is the reference model version of the ORBS index. The bottom panel is the updated ORBS index (year*boat type interaction removed).









The STAT can also flip through the plots folders for these runs to better show more subtle differences.

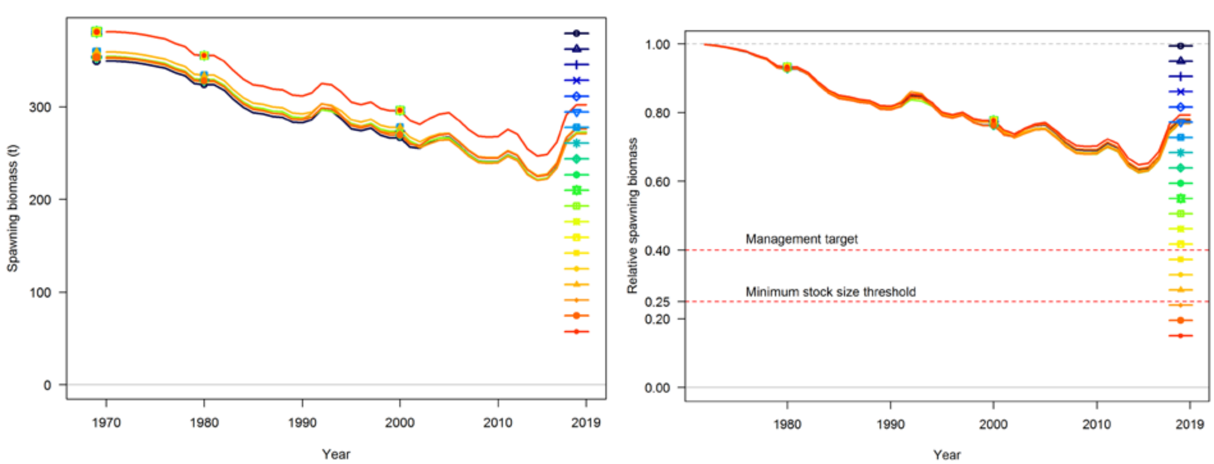
Conclusion: The effect of the index is marginal at best and changes in standardization fo not effect the overall trajectory much.

2. Provide comparative biomass and depletion trends for jittered models within 2 negative log likelihoods.

Rationale: To evaluate whether small changes in the current model are creating the big changes seen in the bridging analysis.

Response: Twelve jitter runs were identified that had a negative log-likelihood that was less than a 2 unit difference from the reference model. These small changes in the model likelihood had relatively no effect on model results.

A set of twenty jitter runs (<10 NEGATIVE LOG-LIKELIHOOD units difference from the reference model) are shown below. Increasing from within 2 NEGATIVE LOG-LIKELIHOOD units to within 10 NEGATIVE LOG-LIKELIHOOD units resulted in increased changes relative to the reference model.



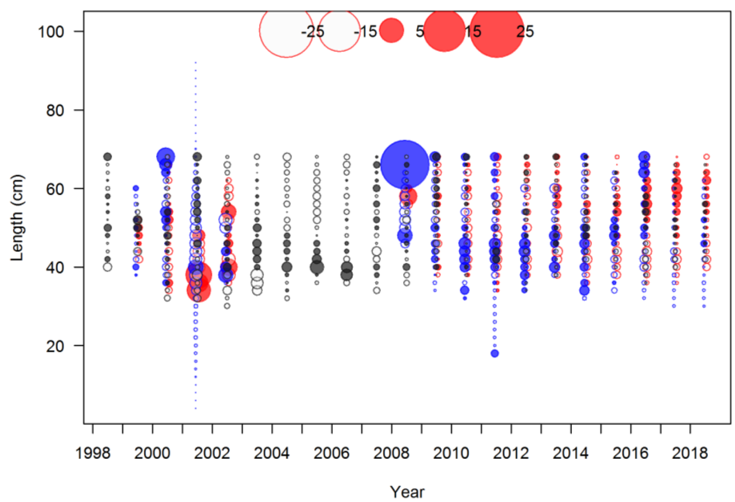
Conclusion: The new versions of the model are not creating any divergent dynamics.

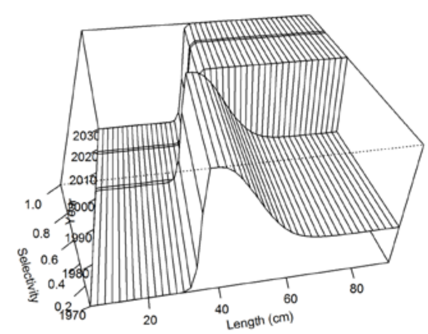
3. Provide a model on the commercial live-fish fleet with another time block in 2015.

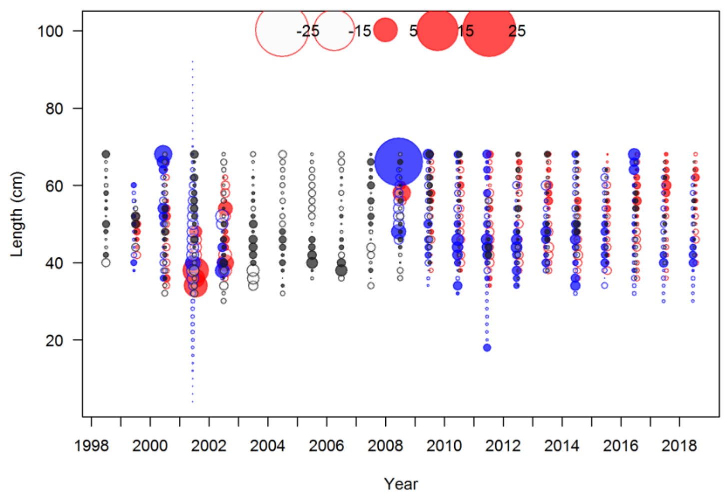
Rationale: To evaluate whether this fixes the length composition residual pattern seen for this fleet.

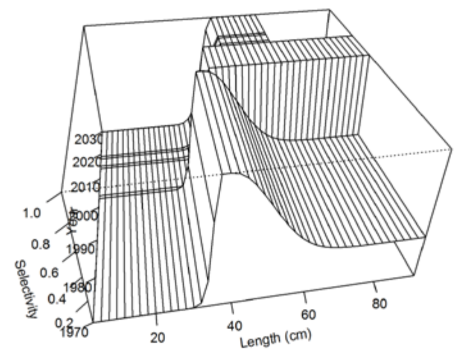
Response: An additional time block from 2015-2018 reduced the residual patterns (top: reference model; middle: add 2015 time block (pre-2015 the same); bottom: add 2015 time block (pre-2015 freed up). Total likelihoods (# parameters) were:

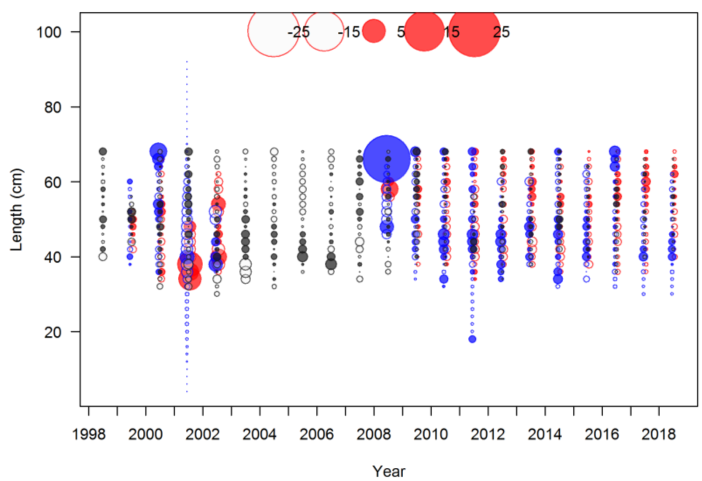
Reference model:	**704.947 (62); AIC= 1533.9**
Add 2015 time block_v1:	**699.674 (67); AIC= 1533.3**
Add 2015 time block_v2:	**699.31 (70); AIC= 1538.6**

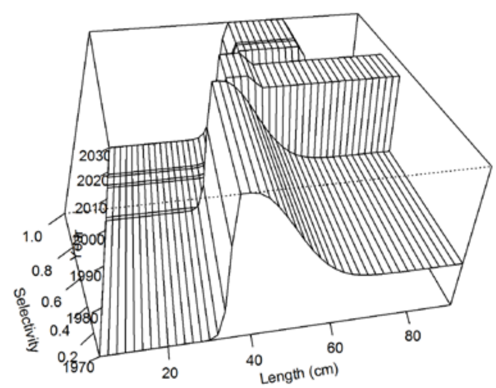












The STAT can show the associated r4ss plots during the meeting.

Conclusion: While the apparent misfit for the later periods is fixed with time varying selectivity the AIC values do not suggest adding more parameters.

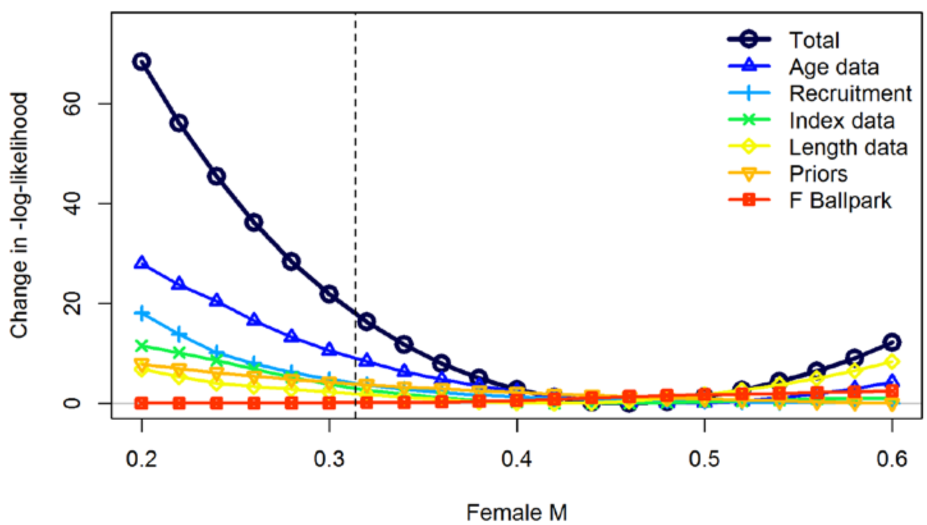
4. a) Provide the fleet-specific age and length likelihood profiles over M.

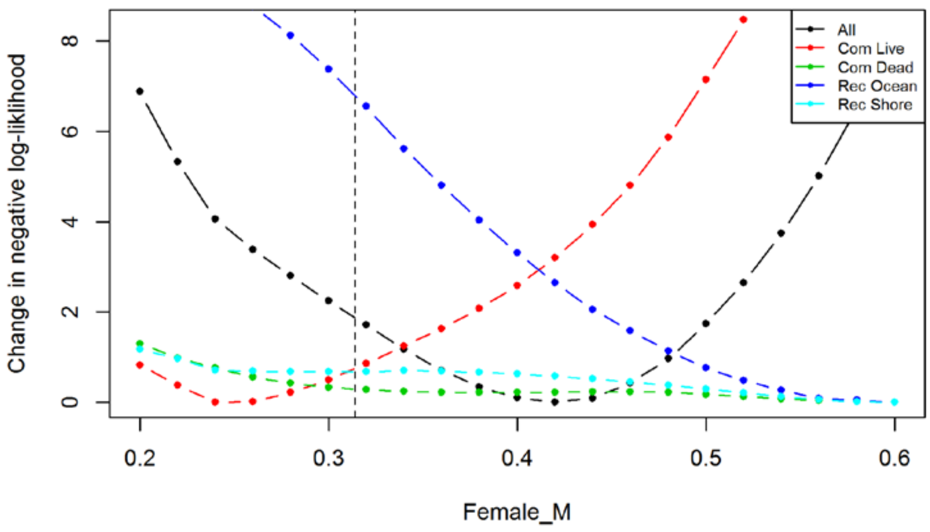
b) Down-weight age components leading to high M; re-run the model while freely estimating M (a model with both sexes freely estimated and a model with only the male offset freely estimated) and growth parameters.

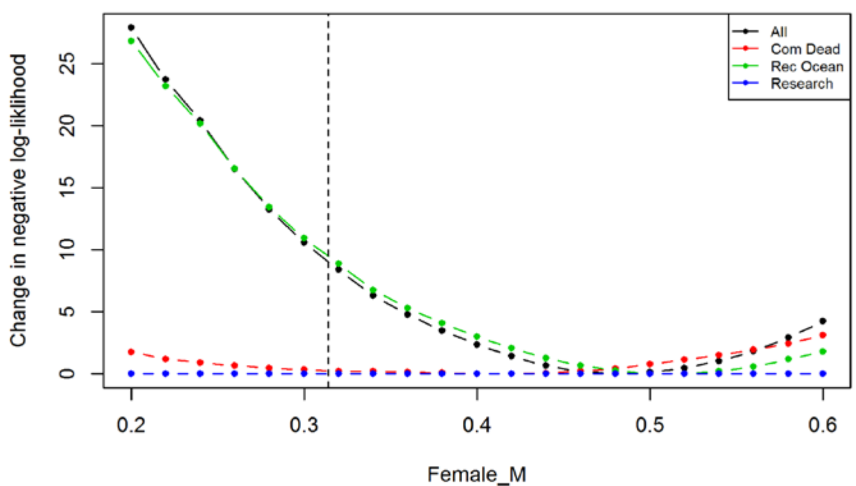
c) If there are age components identified leading to high M that does not fix the aberrant growth patterns seen, fix the CVs at older ages (using the NCS model), to see if that makes the growth patterns more reasonable.

Rationale: To better understand what is leading to the low CVs estimated in older ages in growth curves and the high M estimates in the current base model.

Response (a): Fleet-specific length (top) and age (bottom) likelihood components have been shown below. The information suggesting a high M is coming from the Rec Ocean boat age data; the data source that contains a large part of the total available age data.







(b) Down-weight age components leading to high M; re-run the model while freely estimating M (a model with both sexes freely estimated and a model with only the male offset freely estimated) and growth parameters.

Response (b): Alternative down-weighting values (lambdas) were evaluated when estimating both female and male natural mortality. Italics are fixed values.

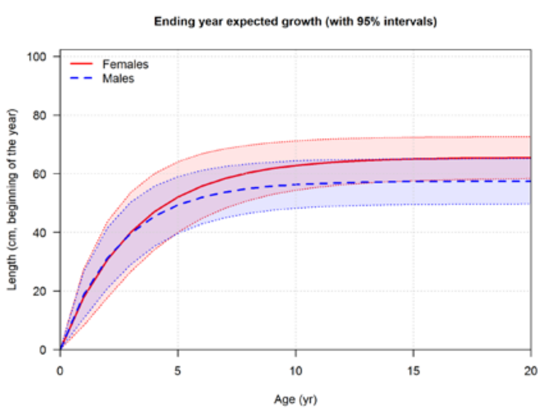
Data Source	Lambda	Female M	Male M

	1.00	0.458	0.430
ean Age only	0.50	0.443	0.412
ean Age only	0.25	0.413	0.381
ean Age only	0.05	0.361	0.342
ean Age only	0.50	*0.314*	0.304
;e Data only (fix growth)	0.00	0.426	0.405

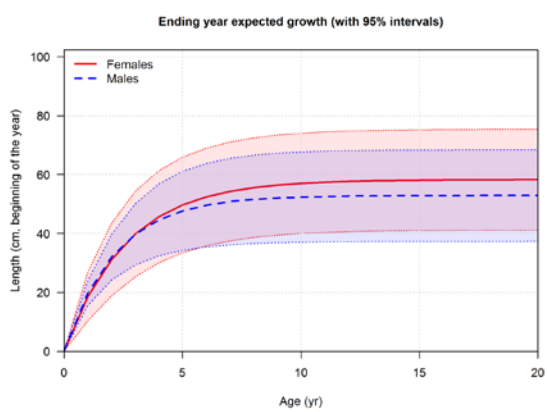
(c) If there are age components identified leading to high M that does not fix the aberrant growth patterns seen, fix the CVs at older ages (using the NCS model), to see if that makes the growth patterns more reasonable.

Response (c): Fixing growth related CVs for older (Lmax age = Linf) for male and female at <u>0.15</u> (similar to that estimated for CA).

Data Source	Lambda	Female M	Male M
	1.00	0.482	0.442
ean Age only	0.25	0.394	0.360

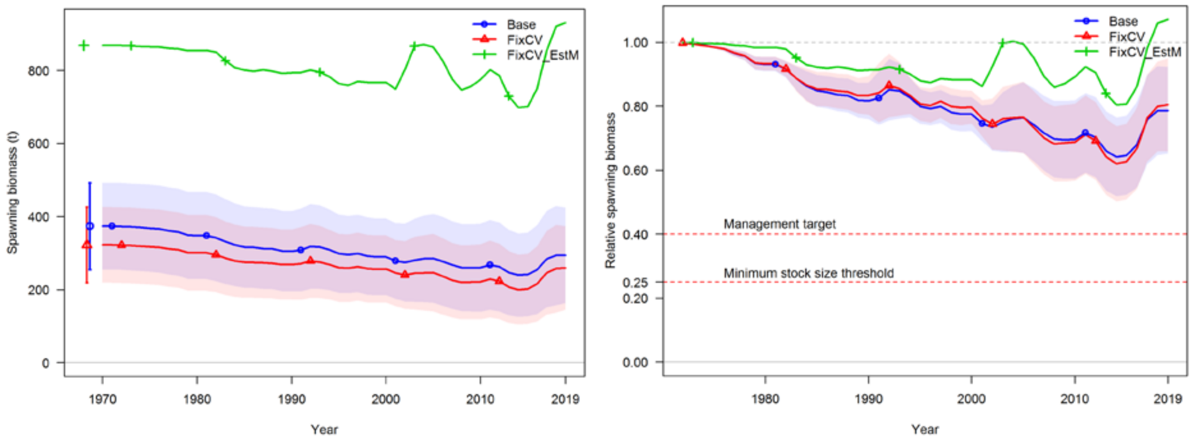


Reference model

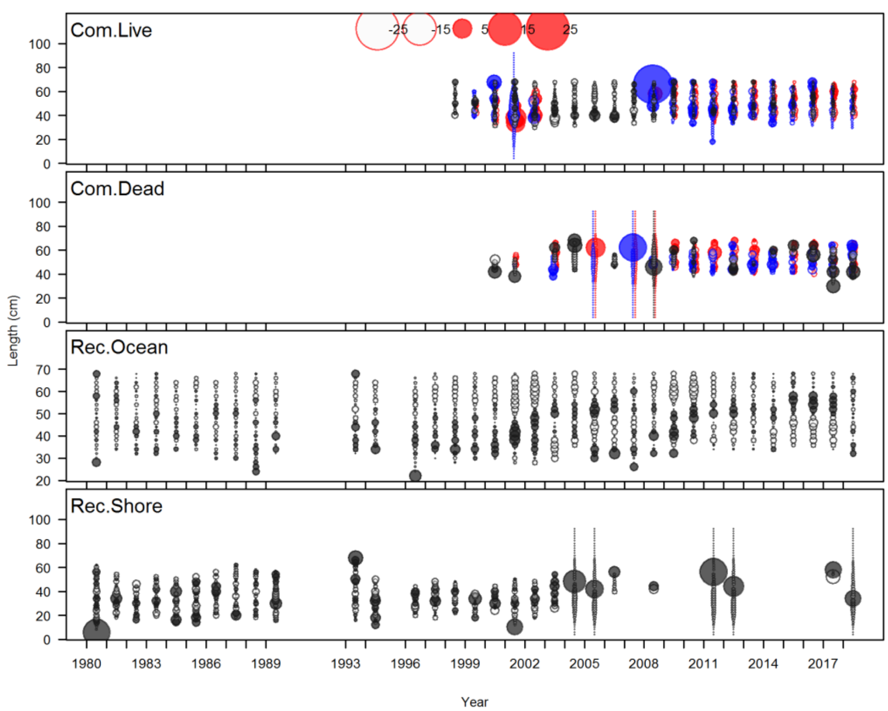


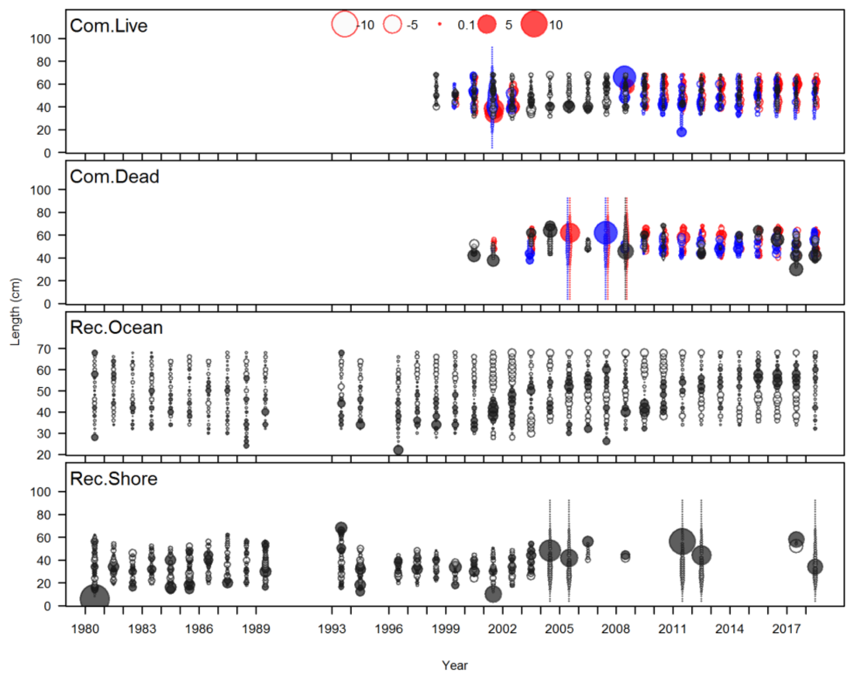
FixedCV_Lmax

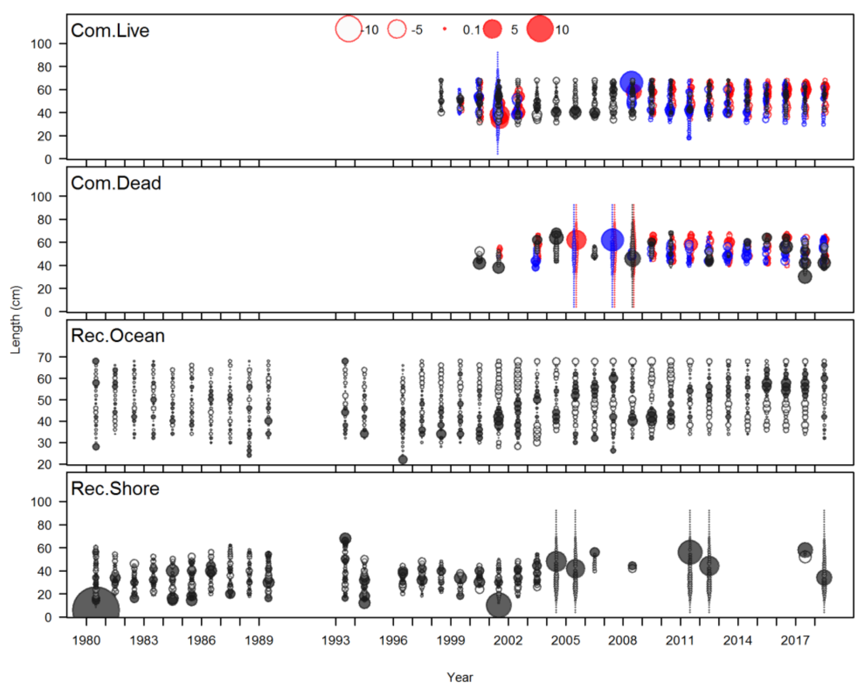
Length Comp residuals for alternative runs when fixing CV at Lmax.



Length Comp residuals for the reference model (top), Fix CV at Lmax (middle), and Fix CV at Lmax and estimate M (bottom).







Conclusion: None of these issues fixed either the estimate of M or the CV estimated at older ages

Further thoughts:

Estimate male offset (fixed female M):

Reference: **704.947 (62; AIC = 1533.894)**
Est male offset: **712.279 (63; AIC = 1550.6)**

Little additional uncertainty in 2019 biomass/status added.

2nd Round of Requests for the OR Cabezon STAT regarding the ORS model

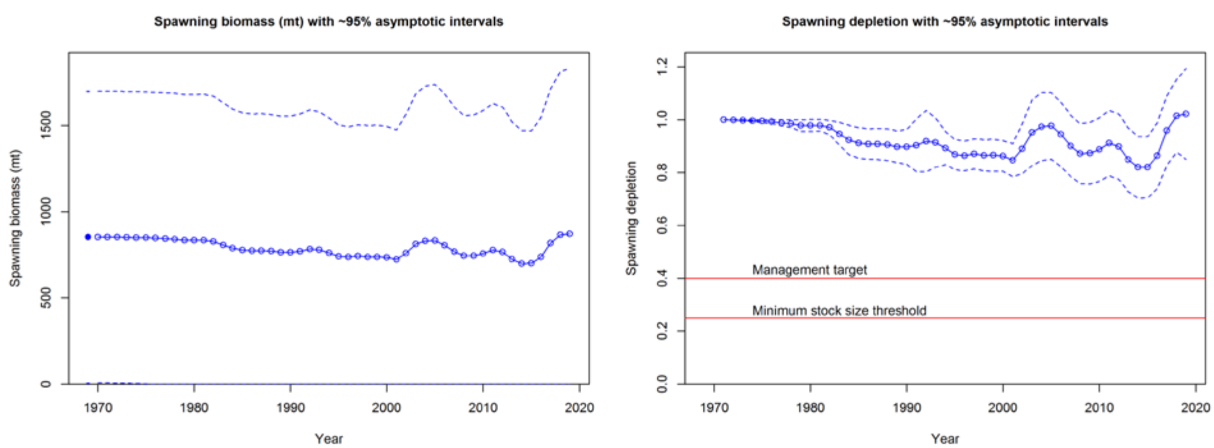
5. a) Estimate female and male M with a more informed Hamel prior (female M = 0.318; CV = 0.219). Provide diagnostics and profiles over M.

b) If the results for request 5a prove unsatisfactory, make M more consistent with the NCS model (0.25 females , 0.318 males at Hamel prior)

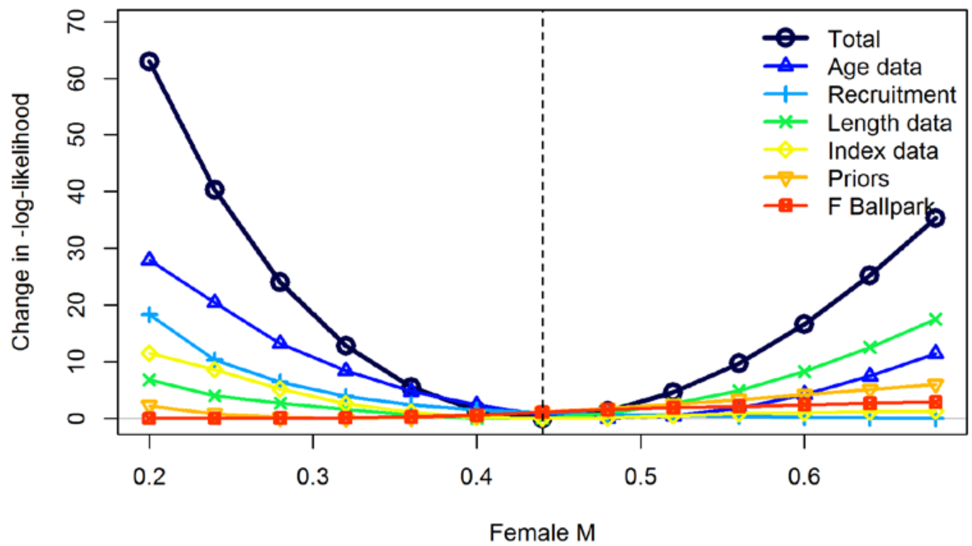
Rationale: a) To evaluate whether the model can estimate plausible values of M.

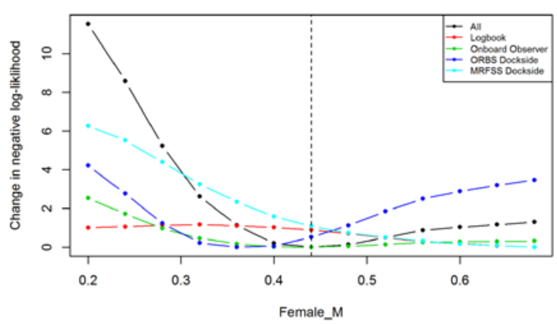
b) To understand how to better bound model results with more plausible values of M to evaluate states of nature in a decision table.

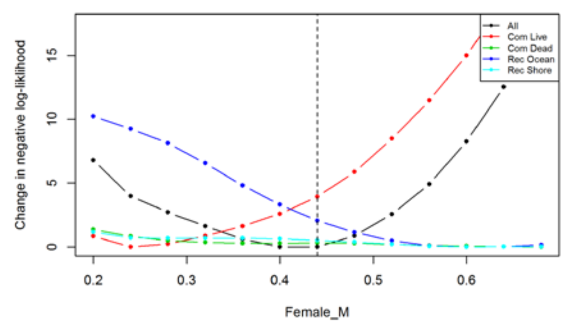
Response (a): The informed Hamel prior on M resulted in higher than expected estimates of natural mortality (female M = 0.440; male M = 0.415). Despite a more constrained prior, the reference model continues to be unable to provide reasonable (given life history and estimates from NCS model) estimates of natural mortality.

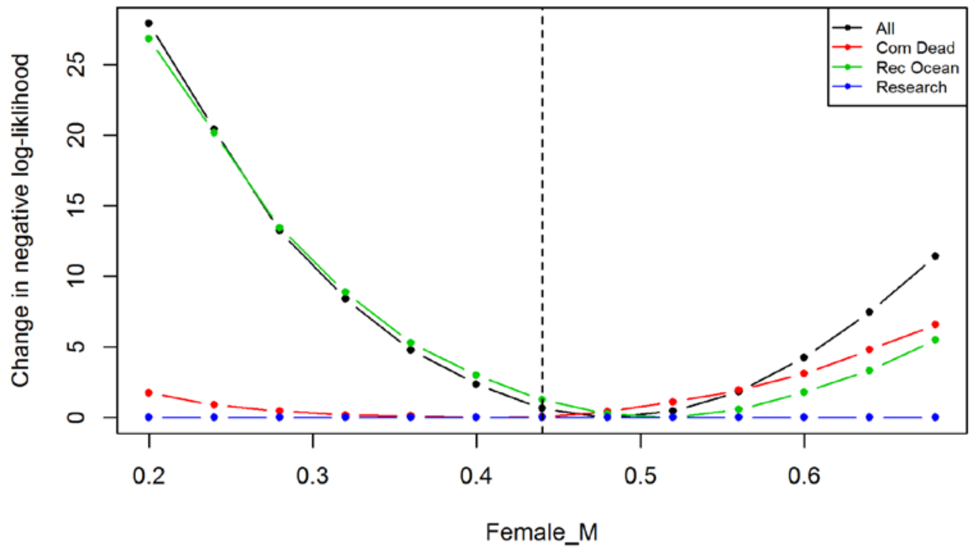


Profiles for informed Hamel prior on M are:


Change in -log-likelihood
0
10
20
30
40
50
60
70
0.2
0.3
0.4
0.5
0.6
Female M
Total
Age data
Recruitment
Length data
Index data
Priors
F Ballpark

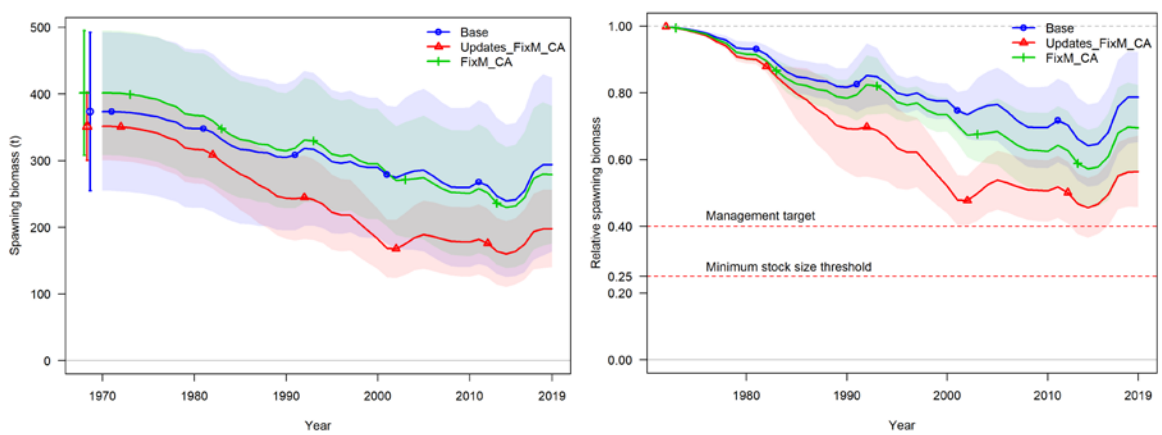

Change in negative log-liklihood
0
2
4
6
8
10
0.2
0.3
0.4
0.5
0.6
Female_M
All
Logbook
Onboard Observer
ORBS Dockside
MRFSS Dockside


Change in negative log-liklihood
0
5
10
15
0.2
0.3
0.4
0.5
0.6
Female_M
All
Com Live
Com Dead
Rec Ocean
Rec Shore



Response (b): These runs are shown below. The “FixM_CA” refers to the natural mortality specifications under this item above.

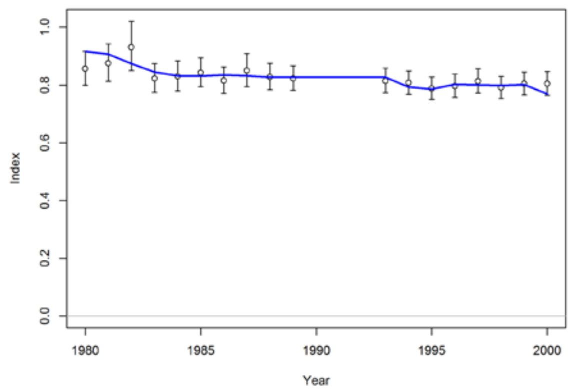
Conclusion: None of these fixed the estimation of M

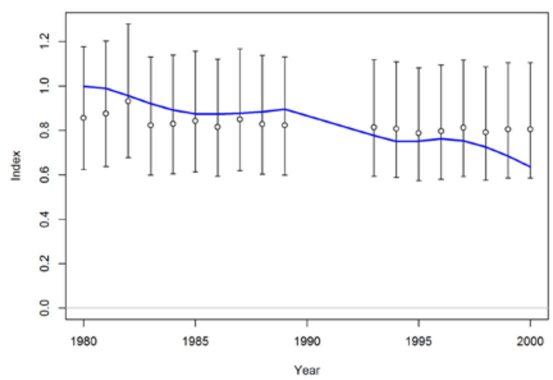


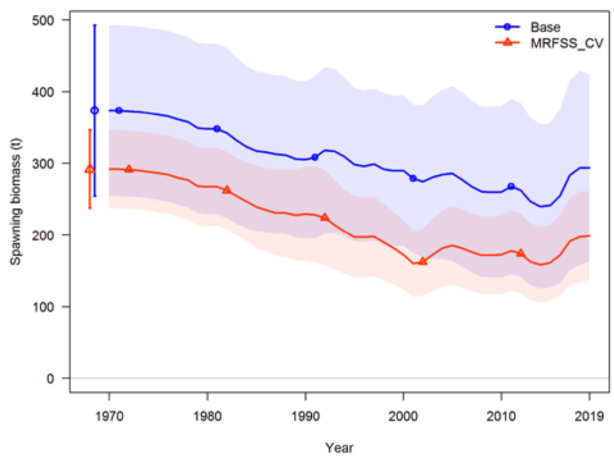
6. Increase the CV on the MRFSS index to a level consistent with the other rec. indices.

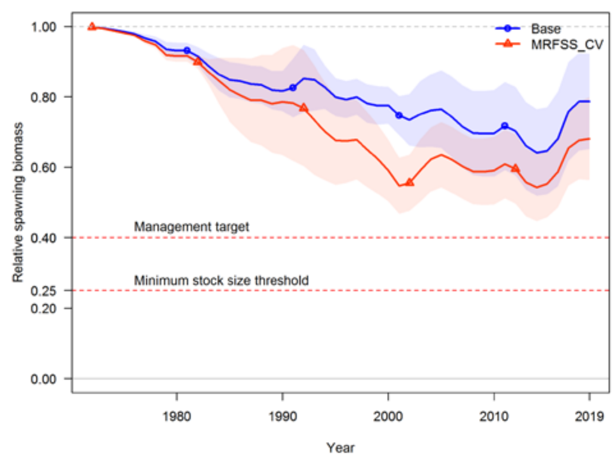
Rationale: To allow more flexibility in the model dynamics.

Response: Increasing the CV to the average of the CV for the onboard index (0.162) resulted in changes to the dynamics of the stock through the index period (more steep decline in population size) and resulted in a lower estimate of stock status. When allowed to estimate an extra SD parameter on the MRFSS index, the reference model did not prefer to add extra SD (parameter at ~ 0). However, the input CVs are overly precise (CV’s ~ 3-5%) and thus extra added variance (down-weighting) the MRFSS index seems reasonable.





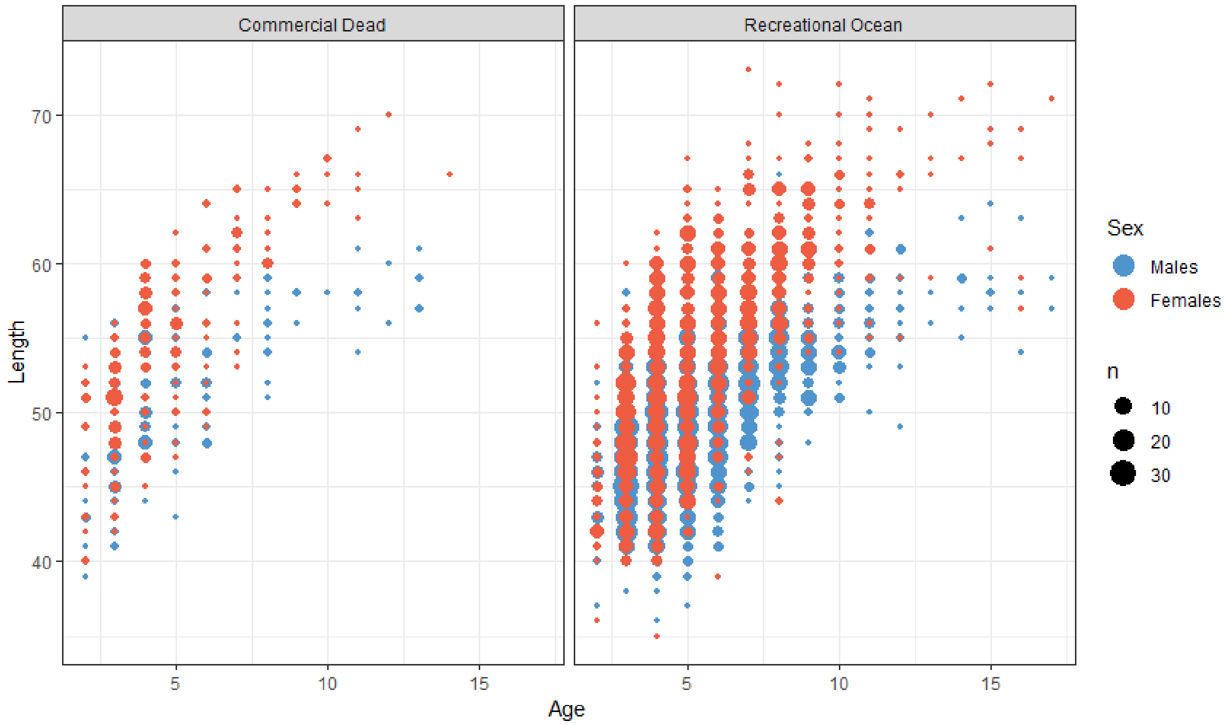




Conclusion: Freeing up the CV's gave the model more flexibility to capture dynamics that are more in line with what field data suggests.

7. Provide plots of lengths and associated ages by fleet and sex.

Rationale: To understand why the estimated CVs are so tight for older ages.



Response: Sample sizes at older ages are limited in number and are only present in the Rec Ocean fleet. In general, the spread of length at age appears to be greater in the Rec Ocean fleet; however, there are large discrepancies in sample size between these fleets.

Conclusion: Uncertain whether the age length data represents the live capture fleet. Also not sure whether these growth curves differ by fleet.

3rd Round of Requests for the OR Cabezon STAT regarding the ORS model

8. Present an updated reference model that includes the fixed ORBS index, down-weighting the MRFSS index (added CV), and fix female M at 0.240 and male M at 0.280 (consistent with the NCS model). Provide plots and diagnostics.

Rationale: To verify the proposed reference model for ORS. Fixing M is necessary since the ORS does not have adequate data to estimate M and the M values are easy to explain.

Conclusion: Represents the realistic dynamics as suggested by STAT. However the next analysis makes it more evenly distributed.

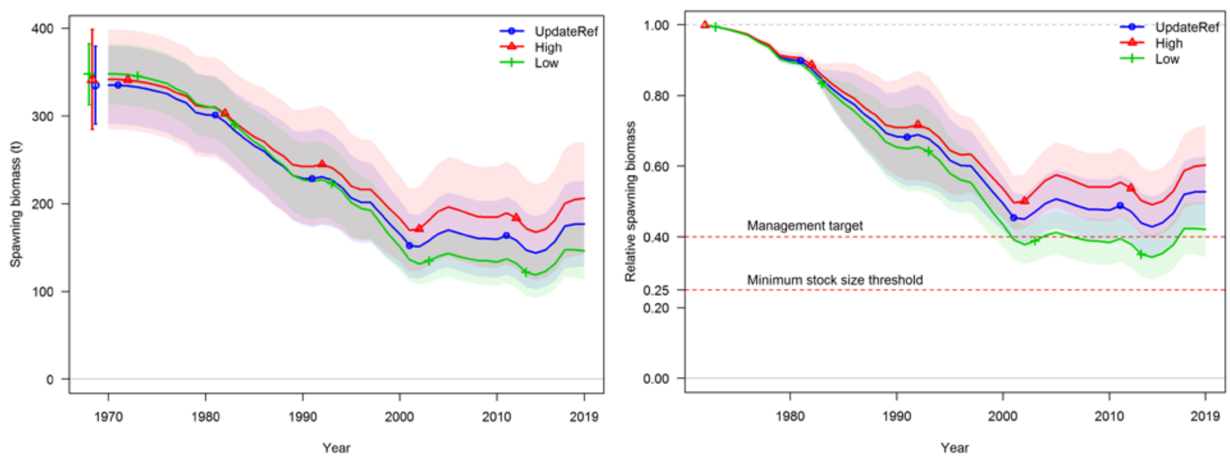
Response: This run was done and the STAT will show the suite of r4ss output.

9. Explore runs for the low and high states of nature by deriving the 12.5% and 87.5% quantiles of 2019 biomass to estimate ranging M values in the ORS model. If that does not provide adequate contrast, use the low and high Ms from the NCS model.

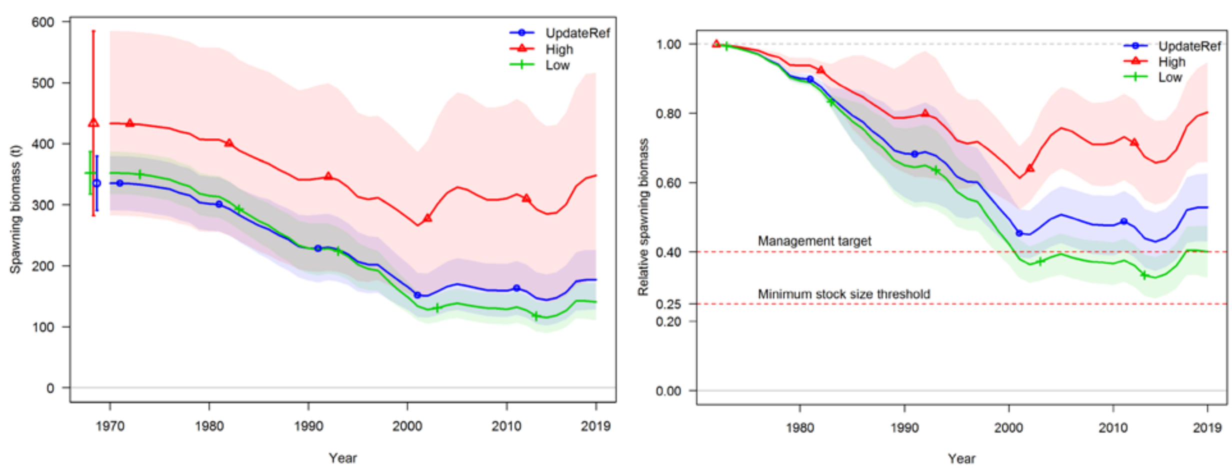
Rationale: To explore potential states of nature in the ORS decision table.

Response: The STAT created two versions of the possible decision table.

Preferred version: alternative female M's were iteratively chosen (male offset remained the same) to match the 12.5% and 87.5% quantiles of 2019 biomass (147 and 205 mt, respectively) and then those models were run as high and low states of nature. This resulted in female M values of 0.19 and 0.27.



Alternative version: alternative female M's were set at 0.18 and 0.34 (matching that used for the NCS states of nature).



Conclusion: The is set of graphs are more appropriate given the CI's overlap.

Final Thoughts:

Proposed Changes to the reference model:

- Update ORBS index
- Widen CVs on MRFSS index

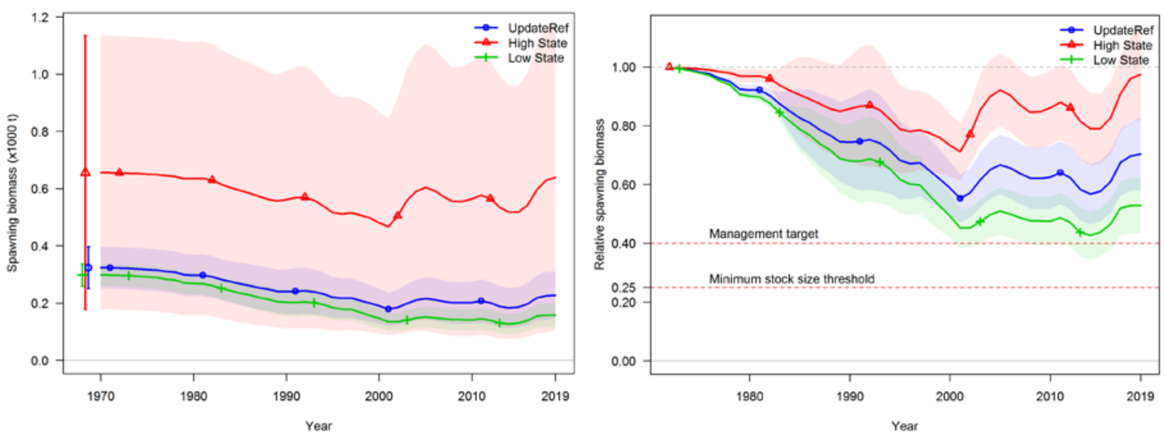
Further evaluations related to estimating natural mortality.

Model	Likelihood	Female	Male
Updated: Fix Fem/Mal to Hamel prior	732.968	0.318	0.318
Updated: Fix male offset (NCS), Est Female	750.351	0.314	0.377
Updated: Fix female to model above and Est male offset	732.73	0.314	0.300

Alternative states of nature:

1. High M: value close to that estimated in ORS - ~0.41
2. Low M: value close to that used in NCS - ~ 0.25

Re 2: see run in response in 5b above. That run fixes M for males at median of Hamel Prior (0.318), fix females at a value consistent with NCS model (0.25).



WAS Assessment

Data sources

- Catch data from two fisheries: commercial (dead-fish; there is no live-fish fishery in WA) and recreational. Neither source includes length compositions; the recreational landings data are in terms of numbers only, so three different catch scenarios were formulated to represent likely average weights of the reported catch.
- Length-weight data from some recreational samples
- Age-length data from some recreational samples

Notable updates to data sources

- The age-length data are a new addition since the previous assessment, making the SSS method possible in the current assessment.

Model details (including details changed from the previous 2017 assessment model):

- SSS is used instead of deletion-based stock reduction analysis (DBSRA). The change allowed for a) exploration of alternative selectivities, rather than assuming it was equal to maturity (and not knife-edged), b) applying sex-specific growth values and c) the use of the F_{MSY} proxy to calculate the overfishing limit (OFL).
- This is a two sex model with the same length and age population structure as in the NCS and SCS models, which are very similar to the ORS model.
- There is one recreational fleet represented in the model (commercial harvest has been restricted to tribal landings since 1999 and are negligible).
- This method uses no measured indices of abundance (it does use a "stock status survey" as described below) or biological data.
- Relative stock status in 2019 is estimated using length-based spawning potential ratio (LB-SPR), calculated based on the biological parameters and the available recreational length data.
- The relative stock status input is implemented as a "survey" with high precision that forces the model to match a specific stock status in a given year and drawn from a distribution specified by the user. A beta distribution is used to express the uncertainty in the relative stock status, with the LB-SPR estimates used to establish a range of relative stock status values. A beta distribution was used as it was in the previous OFL estimation, but the source of stock status year and prior are different. The previous method borrowed stocks status from Oregon in 1997 (before the live fish fishery started in Oregon), whereas the current application uses limited length compositions from Washington to establish a value in 2019.
- Natural mortality a normal distribution and prior was established using the Natural Mortality Tool. The last application used the 2009 female value with a default value of 0.4.
- Growth parameters are fixed to the values estimated for Washington state.
- Maturities are assumed equal to values reported in the Cabezon sub-stock in Oregon waters (Cope and Key 2009; Table 2).
- Steepness is used instead of F_{MSY}/Mand SB_{MSY}/SB_0, which are the productivity parameters as expressed in DBSRA. The steepness value is the one assumed for the other stock assessments (SCS,NCS, ORS). Steepness values used on the west coast are often more productive than the default F_{MSY}/M and

SB_{MSY}/SB_0, values used last time (see Section 3.3.3 for more information).

- Selectivity is set asymptotic at the 18-inch (45.7 cm) minimum size limit and the length of 50% maturity is set to 43.7 cm in Washington.

STAR Panel comments

The approach based was a data poor technique used to estimate OFL's for the WA area. Since this is a tier 3 assessment that only relies on catches and some assumptions of M, h and depletion, we can estimate the overall stock trajectory and the allowable exploitation levels on this stock using DBSRA like SSS runs. The approach had some issues such as the prior of F_{MSY}/M seemed skewed to values less than one, and this created values of steepness in a Ricker and Beverton-Holt type model to be extremely low (h=0.33, or h=0.45) and appear unrealitsic. Depletion was estimated from a stable equilibrium model with available length composition data that would provide, an F and depletion estimate based on the sampled length compositions. Other points noted on this approach were identified:

1. Longer projections of up to 10 years might be useful
2. The M prior was taken directly from the natural mortality tool and needed more documentation

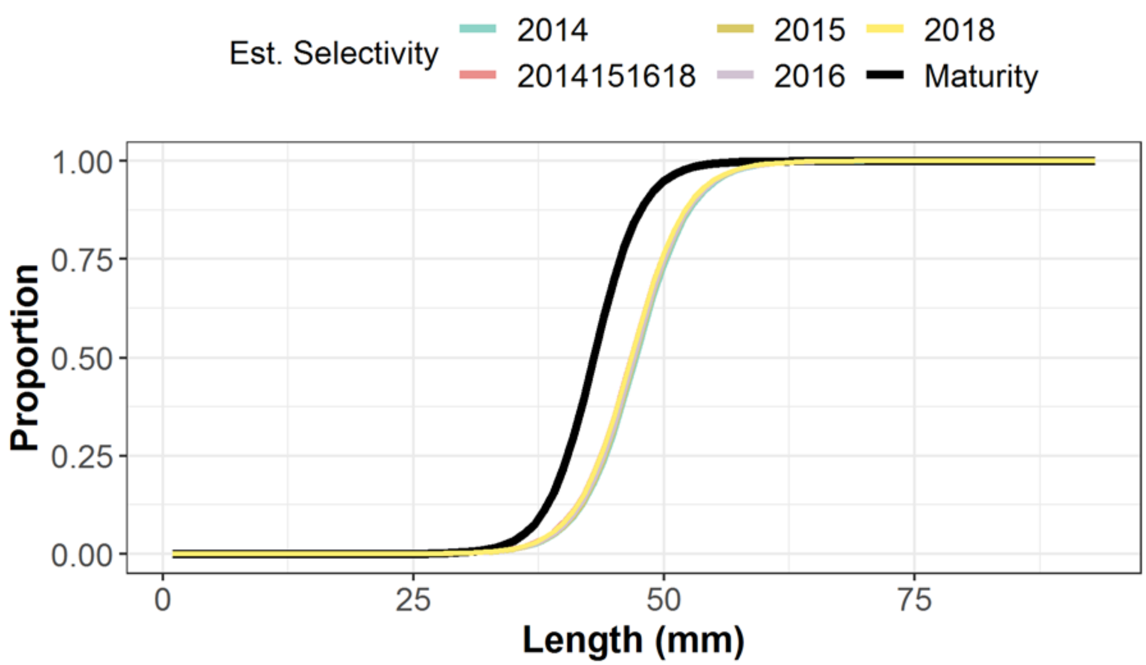
Requests by the STAR Panel for WAS and Responses by the STAT

1st Round of Requests for the WA Cabezon STAT Regarding the WA Model

1. Aggregate length in years 2014-2018 and use that one length composition in LB-SPR to estimate depletion and the associated uncertainty.

Rationale: See sensitivity of LB-SPR to alternative treatment of length compositions.

Response: Years 2014-2016 and 2018 were used in the presented WA results, so these were the years aggregated into one length composition. The aggregate composition data (labelled "2014151618") SPR was then compared to the SPR value used in the WA analysis (65%). Selectivity curves were similar among all years and the aggregate composition. Both approaches estimated SPR to be 65% with very similar uncertainty estimates.



Conclusion: This does not change the state of depletion for the stock. And in addition selectivity is constant for the time period examined.

Years	SPR	SL50	SL95
2014	0.61 (0.48 - 0.74)	49.25 (46.4 - 52.1)	57.93 (53.38 - 62.48)
2015	0.66 (0.47 - 0.84)	46.48 (44.02 - 48.94)	53.46 (49.17 - 57.75)
2016	0.62 (0.51 - 0.74)	48.9 (46.48 - 51.32)	57.25 (53.41 - 61.09)
2018	0.67 (0.56 - 0.77)	44.17 (42.76 - 45.58)	52.21 (49.82 - 54.6)
Aggregate	0.65 (0.59 - 0.72)	46.44 (45.32 - 47.56)	54.85 52.99 - 56.71)

2. Use male growth values to compute LB-SPR to estimate depletion and the associated uncertainty.

Rationale: See sensitivity of LB-SPR to using the males given the length compositions are mixed sex.

Response: The male life history parameters were unable to detect any decrease from an unfished state. Given the relatively small L_∞ value (65 cm) compared to the observed length frequencies, this result is unsurprising. Overall, this approach is not recommend because the LB-SPR approach assumes female spawning biomass is being measured. The data did not have sufficient samples of sexed individuals to isolate female length compositions.

Years	SPR	SL50	SL95
2014	1 (1 - 1)	48.59 (46.66 - 50.52)	56.75 (53.35 - 60.15)
2015	1 (1 - 1)	46.7 (44.54 - 48.86)	53.77 (49.72 - 57.82)
2016	0.99 (0.77 - 1)	48.42 (45.59 - 51.25)	56.33 (52.02 - 60.64)
2018	1 (1 - 1)	44.71 (43.12 - 46.3)	53.02 (49.54 - 56.5)
Aggregate	1 (1 - 1)	46.74 (45.83 - 47.65)	55.25 (53.61 - 56.89)

Conclusion: This does not work as selectivity and F is related to females and not males.

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

NCS Assessment

Several changes were made to the pre-STAR base model during review by the STAR panel:

- Natural mortality was estimated using a lognormal prior (based on the Natural Mortality Tool), rather than a normal prior. This produced more biologically plausible mortality estimates.
- The selectivity of the recreational shore fishery was constrained to be asymptotic.
- The von Bertalanffy *k* parameter was estimated without priors, because the draft assessment had used a prior based on a dataset that was also included in the posterior estimate.

Alternative models for bracketing uncertainty

The STAR panel and STAT agreed that the major axis of uncertainty was in natural mortality (M). Uncertainty in spawning biomass was used to determine natural mortality values to bracket the uncertainty around the reference model. Quantile values of 12.5% (low state) and 87.5% (high state) for 2019 spawning biomass were calculated from the asymptotic estimates of error from the reference model. Fixed female natural mortality values, while estimating male natural mortality, were then explored to find values that approximated those calculated low and high states of nature 2019 spawning biomass values. For the NCS model, the low and high states of nature based on female natural mortality were 0.18 and 0.3457, respectively.

SCS Assessment

Several changes were made to the pre-STAR base model during review by the STAR panel:

- Natural mortality was estimated using a lognormal prior (based on the Natural Mortality Tool), rather than a normal prior. This produced more biologically plausible mortality estimates.
- The selectivity of the recreational shore fishery was constrained to be asymptotic.
- Many years of the SCS model recreational shore fleet length compositions suffered from extremely limited effective samples per year, which caused unduly influence of these low samples in the model. A decision rule to use only use years with effective sample sizes of <5 excluded these low sample years.

Alternative models for bracketing uncertainty

The STAR panel and STAT agreed that the major axis of uncertainty was in natural mortality (M). Uncertainty in spawning biomass was used to determine natural mortality values to bracket the uncertainty around the reference model. Quantile values of 12.5% (low state) and 87.5% (high state) for 2019 spawning biomass were calculated from the asymptotic estimates of error from the reference model. Fixed female natural mortality values, while estimating male natural mortality, were then explored to find values that approximated those calculated low and high states of nature 2019 spawning biomass values. For the SCS model, the low and high states of nature based on female natural mortality were 0.18 and 0.3426, respectively.

ORS Assessment

Several changes were made to the pre-STAR base model during review by the STAR panel:

- The ORBS index was recomputed to remove the year:boat interaction; this interaction effect complicated the across-region aggregation of the index (due to the year:region interaction effect) and did not appear to strongly affect the overall index trajectory.
- The CV of the MRFSS index was judged to be overly precise (~3.5%) and were increased by adding variance (down-weighting) by manually increasing the CV to the average of the CVs from the onboard observer index. The original CVs were implausibly small given likely sampling accuracy, and appeared to greatly constrain early model dynamics. The revised CVs allowed greater variability in early stock dynamics in the early part of the model run.
- Natural mortality (M) values were fixed to those estimated for NCS (0.24 and 0.28, for female and male, respectively). The resulting values are comparable to the mean of the Hamel prior (for males), if one includes the offset of female mortality observed in NCS and SCS. Given

the apparent biological implausibility of M estimates obtained from the model, borrowing from the nearby stock region is the next best alternative.

Alternative models for bracketing uncertainty

The STAR panel and STAT agreed that the major axis of uncertainty was in natural mortality (M). Uncertainty in spawning biomass was used to determine natural mortality values to bracket the uncertainty around the reference model. Quantile values of 12.5% (low state) and 87.5% (high state) for 2019 spawning biomass were calculated from the asymptotic estimates of error from the reference model. Fixed female natural mortality values (with a constant relative value, or offset, of male natural mortality values) were then explored to find values that approximated those calculated low and high states of nature 2019 spawning biomass values. The low and high states of nature based on female natural mortality were 0.19 and 0.27, respectively.

WAS Assessment

No changes were made to the pre-STAR base model (as described above) during the STAR panel.

Alternative models for bracketing uncertainty

Three catch scenarios (based on the average weight of fish used to expand numbers to biomass, and the same scenarios as the 2017 DBSRA application) and five relative stock status values (40%, 55%, 65%, 75% and 90%) were explored for OFL calculation using SSS (for a total of fifteen scenarios). The middle relative stock status value is the mean SPR value from the LB-SPR analysis, with the other values presenting a balanced look at more or less probable relative stock status values, including one at the target biomass (40%). Each SSS scenario was run 1,000 times to produce OFL forecasts.

In addition to presenting each scenario individually, the 15 scenarios are also presented as two different ensembles. One ensemble weights each scenario equally, thus simply combining all scenarios into one distribution. The other weighting schemes assumes the middle catch scenario is twice as likely as the other two and the relative stock status scenarios weights are based on the standardized density values determined by the SPR estimate (mean 0.65 with standard deviation = 0.075).

Technical Merits of the Assessment

The California and Oregon assessments use SS3 as the modelling framework for the assessments. This allows a variety of disparate data to be included in a single analysis. Parameters are estimated via maximum likelihood to appropriately weight the data components. Priors can be applied to incorporate external information on parameters. Uncertainty in the estimates is characterized by the asymptotic variances of the parameter estimates. SS3 is a well-established and tested approach and appropriate for the assessments.

The species is assessed as four separate populations to capture spatial structure. These populations are the same as those adopted at the last assessment in 2009 and are supported by genetic and fishery evidence.

NCS

The assessment uses a range of data including catch (treated as exact) from four fleets, associated length compositions, two fishery-dependent CPUE series that are non-overlapping and a few years of conditional age compositions. These provide the assessment with sufficient information to estimate stock metrics (SSB0, SSB2019, Bratio2019, MSY_SPR and F_SPR) and their associated variance. The natural mortality tool was used to derive a prior on M which facilitated a more realistic quantification of model uncertainty. In addition to computed asymptotic variances, the reference model configuration was used to run an MCMC simulation to provide more realistic posterior distributions of the quantities of interest. This analysis indicated that the lower bound of the biomass trend computed from the asymptotic variances were likely to be too low. It better characterized the asymmetric variances of the parameters and derived quantities.

A systematic sensitivity analysis which considers the principal sources of uncertainty is presented. The analysis considers the influence of data components (indices, length compositions and conditional age) and model specification (M, growth, data weighting and recruitment assumptions) in the principal stock metrics. The results of these sensitivities are plotted to show where the estimates lie in the range of uncertainty as derived from the reference model. This provides a very clear indication of where the main issues lie.

Retrospective runs did not reveal any major problems as data are sequentially removed from the assessment. However, the analysis illustrates the dependence of the assessment on the catch data that is assumed to be known without error. Jitter analyses suggest that the model converges on the lowest negative log-likelihood.

SCS

While the SCS assessment follows a similar approach to the NCS approach there is much less data. An improved catch stream was used that accounts for some mis-classification between the northern and southern areas. In view of the more limited amount of data, growth parameters were taken from the NCS, but in most other respects the model configuration reflects the NCS assessment. Sensitivity and MCMC runs are provided in the same form as NCS which provide an informative overview of uncertainty. Retrospective runs did not reveal any major problems as data are sequentially removed from the assessment. However, the analysis illustrates the dependence of the assessment on the catch data that is assumed to be known without error. Jitter analyses suggest that the model converges on the lowest negative log-likelihood.

ORS

The Oregon assessment benefits from a larger amount of data that include four fishery dependent abundance indices. There is also age data for two fleets from around 2005 and some research data for a few recent years that will contribute to better estimates of growth and recruitment. Sensitivity and MCMC runs are provided in the same form as NCS which provide an informative overview of uncertainty. Jitter analyses suggest that the model converges on the lowest negative log-likelihood.

WAS

For the Washington assessment a novel approach had been developed for data poor stocks. This builds on the DBSRA approach used earlier but can be implemented within the SS framework. In particular, the productivity parameters in DBSRA can be replaced by conventional steepness used in SS making the assessment more consistent with the other cabezon assessments. It allowed the same steepness assumption (0.7) to be applied. The use of the length composition tool was used to estimate selectivity parameters from recent observed length compositions and current biomass depletion. The panel felt this was an innovative and appropriate modelling approach to compute OFLs and investigate uncertainty.

Technical Deficiencies of the Assessment

This section contains a description of the technical deficiencies that are common to all stocks, plus stock specific deficiencies, which are presented in terms of data inputs and model.

All Stocks

Data Inputs

Fishery catch rate indices for the NCS and OR stocks are affected by changes in trip limits and other management regulations. It was not clear that the standardization procedures fully accounted for the impacts of changes in management regulations. The STAT notes that index data are not included after a major management change in 1999.

Models

There is uncertainty in landings estimates that is not accounted for in all assessments. This is a structural feature of SS3.

There were poor fits to some survey indices in SCS, NCS, and ORS (e.g., MRFSS in ORS and CPFV in NCS), with systematic temporal discrepancies between those indices and model predictions. Those indices are based on fishery catch and effort sampling and may not reliably reflect trends in abundance. Therefore, it was difficult to evaluate the lack of fit to survey indices.

The SCS, NSC, and OR models did not include recruitment deviations in historic periods that only have catch estimates. This results in false precision about estimates of historic stock size. In fact, the model estimates of historic stock size may only indicate average stock size over a substantial number of years. This may arise from a structural feature of SS3.

Index and length composition residual diagnostics for the SCS, NSC, and OR models did not seem to conform to the underlying statistical assumptions used in model estimation (e.g., autocorrelated patterns of positive or negative residuals). This only occurred in some years for some fleets.

NCS

Data Inputs

The coverage of the commercial fleet seemed adequate, the recreational fleet data suffered from both methodological survey deficiencies (prior to 2004), and poor coverage of effort.

Models

Results indicated that the reference model was sensitive to some data (Live COM and REC Boat data).

There was a systematic discrepancy with respect to M sensitivities. All assigned M's in sensitivity runs were higher than the relatively low M estimated in the reference model formulation. It is possible that model mis-specification is confounding the estimation of M, leading to lower-than-expected estimates. It is also possible that given the data, the model cannot reliably estimate M. This implies that M is a major source of uncertainty in the model.

SCS

Data Inputs

Conditional catch at age data was unavailable so growth parameters could not be estimated. Length composition sample sizes are low, particularly in the recent period. There are no abundance indices since 2000.

Models

The SCS model has less data and more structural parameters are fixed, so the model estimates less uncertainty relative to the NCS model but more sensitivity. In absolute terms, the uncertainty is still quite high.

SCS growth parameters were fixed at the values for the NCS stock.

ORS

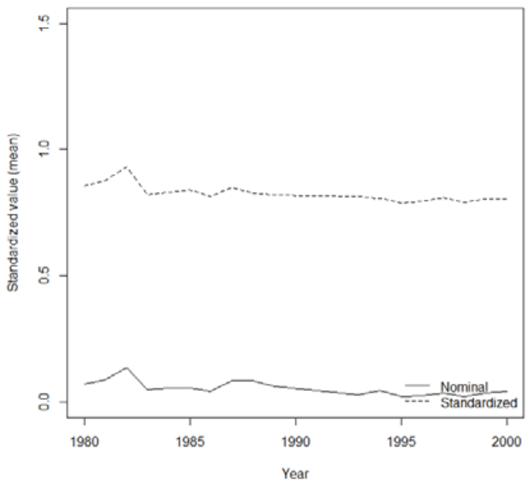
Data Inputs

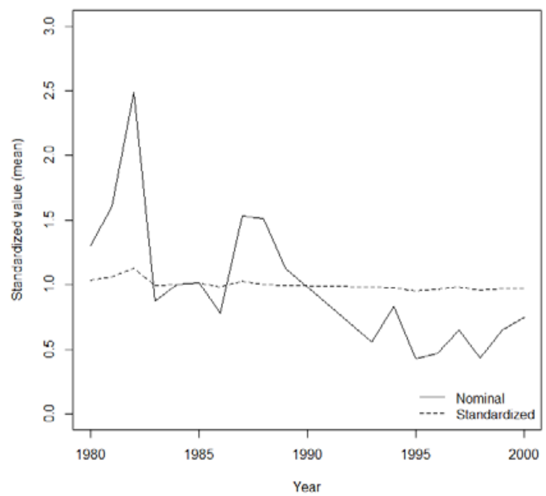
The Oregon Assessment is built on a number of additional datasets; however their utility as indices of abundance or representativeness of the growth at age is uncertain, as some fleets and areas are poorly represented and issues related to spatial and temporal scales are not fully understood.

This assessment is handled as a single area and the fleets-as-area approach has been incorporated. However, the spatial extent of the fleets spans over 360 miles and the data is mostly collected from the south for the commercial live and dead fleet and the central/north for the recreational boat and shore-based fleets (though there is some recreational and commercial catch in both regions).

Almost all of the logbook index data is out of Port Orford, and may not represent the stock (or even the southern coast) as a whole.

The MRFSS standardization resulted in an index that is different from the raw CPUE. The standardization removed much of the raw CPUE signal. Alternative structure in CPUE may indicate a seasonality that is not accounted for in the assessment structure nor in the CPUE standardization. The standard errors for the index seemed much too small and not reliable for stock assessment.





Non-normalized (left) and normalized (right) comparisons of the nominal (raw) CPUE and standardized (delta-glm) CPUE trends for the MRFSS index.

Models

The model estimates a high value of female M and results in unreasonably high uncertainty in stock size. M profile demonstrated that surveys, age data, and length compositions all favor a higher value of M. The RP concluded that the estimated M was not realistic and that this is probably related to some other model mis-specification (however, it was not clear what that mis-specification was) or simply an inability to reliably estimate M given the data.

WAS

Data Inputs

There were no indices of abundance for this stock, and only a short time-series of length composition was available.

Models

A data-limited method based on many assumptions about stock productivity. The length compositions were for unknown (male+female) sexes. However, the LBSPR calculations were based on female parameters. This would tend to underestimate current stock status because males are smaller than females at a given age.

Areas of Disagreement Regarding STAR Panel Recommendations

Among STAR Panel members (including GAP, GMT, and PFMC representatives):

There were no disagreements among the GAP, GMT, and PFMC representatives with respect to STAR panel recommendations.

Between the STAR Panel and the STAT Team:

There were no disagreements among the stock assessment teams with respect to STAR panel recommendations.

Management, Data, or Fishery Issues raised by the GMT or GAP Representatives During the STAR Panel Meeting (John Devore to writeup)

Unresolved Problems and Major Uncertainties

Critical uncertainties

The basic reproductive life history processes of Cabezon is not well understood, including the potential for batch spawning. There is uncertainty about larval connectivity at coastal scales, because genetic results suggest more spatial structure than may be expected given their pelagic larval duration. There is also uncertainty about how density-dependent processes may affect recruitment.

Cabezon appears to be fairly sedentary species so localized depletion could be a problem (see below). It is also important to note that in California there is an extensive network of no-take marine protected areas that should alleviate this worry to some extent. There is also a less-extensive set of no-take MPAs in Oregon. It is unclear how to address this spatial protection in a stock assessment context.

In many fisheries Cabezon is only a valued incidental catch which creates uncertainty in interpreting fisheries catch rate and size composition information.

No stock has a dedicated age-sampling program.

Stock structure research suggests there is more sub-stock structure than is reflected by the current assessment region boundaries. For example, genetics studies in California suggested 6 populations.

Stock boundaries were based on pragmatic considerations of spatial differences in biogeographic shifts in life history traits and data availability. Indices of abundance and length/age composition data were aggregated at the SCS, NCS, OR and WA spatial units, balancing the need for spatial resolution versus data availability and quality.

Sampling coverage and representativeness were discussed for all the data types, and perhaps the model developed for OR should be increased in complexity and split into a northern and southern stock due to the LH characteristics of Cabezon. However, this is beyond the scope of the current assessment and using the single area model maybe the only option but perhaps using fewer indices to represent dynamics and also examining model structure for mis-specification should be checked. Data weighting schemes were giving undue weight to length at age samples and creating conflict in the model to go to higher estimates of natural mortality.

Recommendations for Future Research and Data Collection

All areas:

-All assessment need to deal with catch uncertainty

-Evaluate the way in which the likelihood is calculated for recruitment deviations to properly capture the nested random effect structure that should be present.

-Stock structure issues need to be addressed as current spatial scales seem too large given the LH of the species. Issues on localized depletion cannot be addressed as currently formulated.

Small scale tagging studies may inform range of their movement and appropriate scale for sub-stock structure.

Recruitment models describing nesting/guarding/territorial behavior should be attempted.

1. Examine catch and effort data at finer spatial scales as time/data permits in the future to examine localized depletion issues.

2. Develop fishery independent surveys to better understand stock dynamics to use as an independent data source in model fitting.

3. Develop methods to include recruitment variability in uncertainty intervals for historic stock size and SBo.

4. Consider developing a tagging program to understand the spatial extent of the localized populations at different stages of their life cycle.

5. Future assessments should re-visit standardization of the California Collaborative Fisheries Research Program Nearshore Survey and provide more diagnostics.

California

- California Collaborative Fisheries Research Program Nearshore Survey: Future assessments should re-visit standardization and provide more diagnostics. But for this assessment re-analyzing this does not seem to be high priority.

- SCS needs to collect some data on indices of abundance and life history to update their series. At the very least start collecting some better length composition data as well.

Oregon

Finer spatial scales could be explored with the indices currently being developed and used.

If one area is used, an effort to get a single signal analyzing different set/trip level data across all fleets should be attempted.

A better understanding of the MRFSS index and the appropriate level of error to represent uncertainty should be examined.

A 2-3 area spatial assessment should be explored to address the differences in indices and LC data being collected at different locations. Perhaps a finer resolution time step could also be developed.

- OR: some evidence of different year-classes evident in commercial live length compositions recently compared to Recreational boat compositions. This is something to keep in mind but at present there is no way to address this.

Washington

Length at age data should be examined for growth and a more integrated model developing an index of abundance could be developed for WA in the future. Tagging studies to inform movement could also be conducted and genetic data examining stock structure could be addressed.

Acknowledgements

The STAR Panel thanks the council staff and NOAA Newport staff for a successful meeting. The STAT staff were well prepared and made this meeting easy to conduct and review.

References

Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68, 1124–1138. https://doi.org/10.1139/f2011-025.

Grebel, J.M., Cailliet, G.M., 2010. Age, growth and maturity of cabezon (*Scorpaenichthys marmoratus*) in California. California Fish and Game 96, 36–52.

McAllister, M.K., Ianelli, J.N., 1997. Bayesian stock assessment using catch-age data and the sampling - importance of resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences 54, 284–300.

Methot, R.D., 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, 1744–1760. https://doi.org/10.1139/f2011-092.