IPHC Management Strategy Evaluation to Investigate Fishing Intensity

PREPARED BY: IPHC SECRETARIAT (A. HICKS & I. STEWART; 22 SEPT, 16 OCT 2018)

1 PURPOSE

To provide an update on the progress of the IPHC Management Strategy Evaluation process to investigate fishing intensity, and to present results of the closed-loop simulations (as of 16 October 2018).

NOTE: In this latest revision, Appendix A has been added to provide updated results on some long-term performance metrics for some runs requested at MSAB011 (IPHC-2018-MSAB011-R). Some short-term performance metrics area also reported for those same runs, following direction received from the Commission on 4 October 2018, as follows

> *The Commission* ***RECOMMENDED*** *that the MSAB:*
>
> - *While it is recognized that the MSAB has spent considerable time and effort in developing objectives for evaluating management procedures, for the purpose of expediting a recommendation on the level of the coast-wide fishing intensity, and noting SRB11–Rec.02 to develop an objectives hierarchy, the MSAB is requested to evaluate management procedure performance against objectives that prioritize long-term conservation over short-/medium-term (e.g., 3-8 years) catch performance. Where helpful in accelerating progress on scale, the MSAB is requested to constrain objectives to (1) maintain biomass above a limit to avoid critical stock sizes, (2) maintain a minimum average catch, and (3) limit catch variability.*

2 INTRODUCTION

At the 2017 Annual Meeting (AM093) Commissioners supported a revised harvest policy that separates the scale and distribution of fishing mortality (Figure 1). Furthermore, the Commission identified an interim “hand-rail” or reference for harvest advice based on a status-quo SPR, which uses the average estimated coastwide SPR for the years 2014–16 from the 2016 stock assessment, resulting in an SPR of 46%. The justification for using an average SPR from recent years is that this corresponds to fishing intensities that have resulted in a stable or slightly increasing stock, indicating that, in the short-term, this may provide an appropriate fishing intensity that will result in a stable or increasing female spawning biomass.

The 2017 stock assessment updated the population estimates and determined that the SPR resulting from actual total mortality from all sources in 2017 was 40%, instead of the 45% adopted by Commissioners at AM093. This was an example of estimation error and something that is inherent in the process due to uncertainty in the data. The SPR of 40% was well within the confidence bounds for SPR reported in the 2017 stock assessment (30-59%) and was most likely less than the adopted SPR because of the updated estimation of recent below average recruitment. The estimation may easily go either way (above or below the adopted value).

This document (IPHC-2018-SRB013-07 focuses on the coastwide simulations and includes the following topics:

1. changes to the simulation framework, and
2. preliminary closed-loop simulation results for the evaluation of the harvest control rule. This includes values of SPR and the fishery trigger in the control rule. Final results will be provided before and at MSAB012.

Appropriate background or reference to documents is provided, when needed. Useful documents to reference are IPHC-2018-MSAB012-06 for a description of objectives, and IPHC-2018-MSAB011-08 for a description of the simulation framework. The MSAB011 report (IPHC-2018-MSAB011-R) provides a summary of the outcomes of that meeting. Additionally, documents IPHC-2018-SRB012-08 and IPHC-2018-SRB012-R provide background to SRB discussions in June 2018.

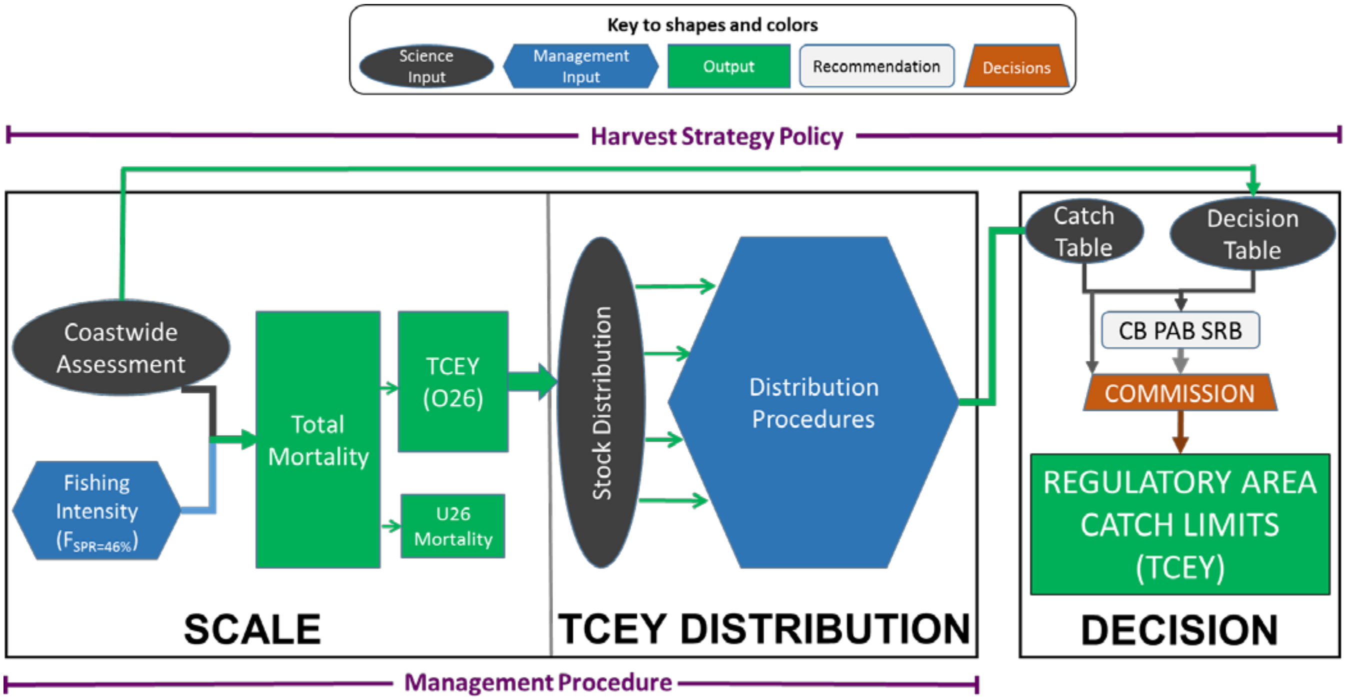


Fig. 1. A pictorial description of the interim IPHC harvest strategy policy showing the separation of scale and distribution of fishing mortality. The “decision step” is when policy and decision making (not a procedure) influences the final mortality limits.

3 CLOSED-LOOP SIMULATION FRAMEWORK

The framework of the closed-loop simulations is a map to how the simulations will be performed (Figure 2). There are four main modules to the framework:

1. The **Operating Model (OM)** is a representation of the population and the fishery. It produces the numbers-at-age, accounting for mortality and any other important processes. It also incorporates uncertainty in the processes and may be composed of multiple models to account for structural uncertainty.
2. **Management Procedure**
   a. **Monitoring (data generation)** is the code that simulates the data from the operating model that is used by the estimation model. It can introduce variability, bias, and any other properties that are desired.

b. The **Estimation Model (EM)** is analogous to the stock assessment and simulates estimation error in the process. Using the data generated, it produces an annual estimate of stock size and status and provides the advice for setting the catch levels for the next time step. However, simplifications may be necessary to keep simulation times within a reasonable time.

c. **Harvest Rule** is the application of the estimation model output along with the scale and distribution management procedures (Figure 1) to produce the catch limit for that year.

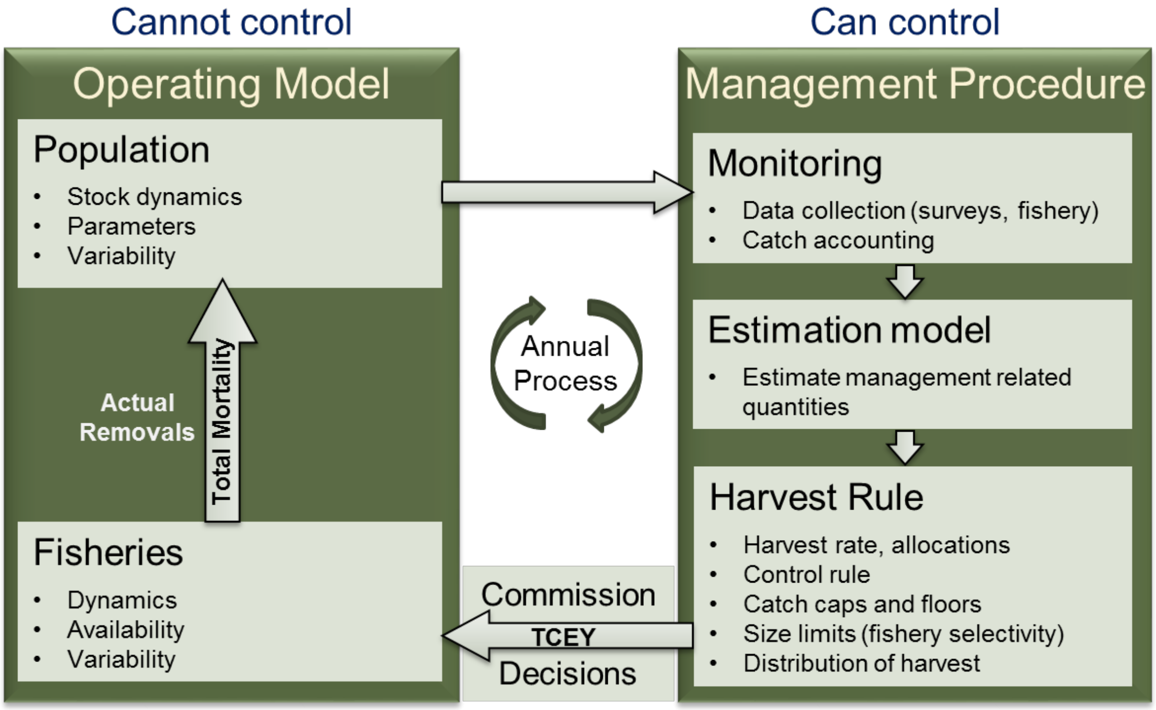


Fig. 2. Diagram of the relationship between the four modules in the framework. The simulations run each module on an annual time-step, producing output that is used in the next time-step. See text for a description of operating model, monitoring, estimation model, and harvest rule.

3.1 OPERATING MODEL

For the simulations to investigate a coastwide fishing intensity, the stock synthesis (Methot and Wetzel 2013) assessment software was used as an operating model. This platform is currently used for the stock assessment, and the operating model was comprised of the two coastwide assessment models (short and long time-series) currently used in the ensemble. For future MSE evaluations (in particular, investigating the Distribution component of the harvest policy) a more complex operating model will be developed that can provide outputs by defined areas or regions and can account for migration between these areas. This model has been referred to as a multi-area model.

The current stock assessment ensemble, composed of four different assessment models, includes a cross between coastwide or fleets-as-areas structuring of the data, and the length of the time series. Using an areas-as-fleets model would require generating data and distributing catch to four areas of the coast, which would involve many assumptions. In addition, without a multi-area model, there would not be feedback from migration and productivity of harvesting in different areas. Therefore, only the two coastwide models were used, but with additional variability. These models are structured to use five general sources of removals (these are aggregated for modelling purposes and do not necessarily correspond to specific fisheries or sectors): the directed commercial halibut fishery (including research landings), commercial discard mortality (previously known as wastage), bycatch (from non-halibut-target fisheries), recreational, and subsistence. The TCEY was distributed to each source in an ad hoc manner using current available information (see below).

3.1.1 Conditioning the Operating Model

The operating model (OM) should be a reasonable depiction of reality with an appropriate level of uncertainty, which is accomplished through a process called conditioning. The operating model (OM) consists of two Stock Synthesis, or SS (Methot and Wetzel 2013), models parameterized similarly to the short and long coastwide assessment models for Pacific halibut (Stewart 2015 appendix of RARA). Each SS model is conditioned by fitting to the same data used in the 2017 stock assessment (Stewart & Hicks 2018, documents 08-10). In order to evaluate and choose management procedures that are robust to uncertainty in the population, many assumptions in the assessment model were freed up to characterize a wider range of possibilities in the future. Table 1 shows the parameters that were different from the assessment models. Estimating natural mortality in both models and estimating steepness were the only processes changed from the assessment model when conditioning.

Table 1. Parameter estimation in the assessment and operating model.

Parameter	**Assessment**	**OM**
Natural Mortality (M)	Some estimated	All estimated without priors
Recruitment (lognormal devs)	Variability fixed at 0.6 (long) 0.9 (short)	Same as assessment
Steepness (h)	Fixed at 0.75	Estimated variability based on long model centered around 0.75 for both.

3.1.1.1 Characterizing Variability in Stock and Fishery Dynamics

Variability was characterized by the estimated variance-covariance matrix estimated automatically by inverting the Hessian within ADMB (http://www.admb-project.org/), which is the optimization software that SS uses. This provides the uncertainty for each estimated parameter, and its correlation with other parameters, given the data and assumptions. Using this variance-covariance matrix, sets of parameters were randomly generated from a truncated multivariate normal distribution. The truncation of parameter bounds was determined from the bounds entered in the SS model files. Some bounds (e.g. dev parameters) were infinite.

An alternative approach for characterizing variability is to design a grid over which different parameter values and assumptions are used. For example, different values of steepness could be chosen and simulations use those fixed values of steepness. Then, the simulations are combined across grid points. We are using the Hessian approach to integrate over a range of parameter values and account for correlation between parameters.

To ensure that parametrically sampling from using a multivariate normal distribution and the inverted Hessian produced similar results as the assessment SS models (the current best information for the historical trajectory), 1000 samples of the parameters estimated in the assessment models were generated from a multivariate normal distribution. Estimated recruitment deviations were bias-corrected by their corresponding estimated variances before sampling from the multivariate normal distribution. The mean spawning biomass trajectory and 95%

confidence interval around that trajectory were compared to the assessment results and the long coastwide model showed an increased density of low spawning biomass compared to the assessment model (Figure 3). Trajectories with a maximum F greater than 0.4 were not within the 95% confidence interval determined from the inverted Hessian in assessment model, thus the sampling from the multivariate normal was limited to trajectories that had a maximum fishing mortality rate less than 0.4.

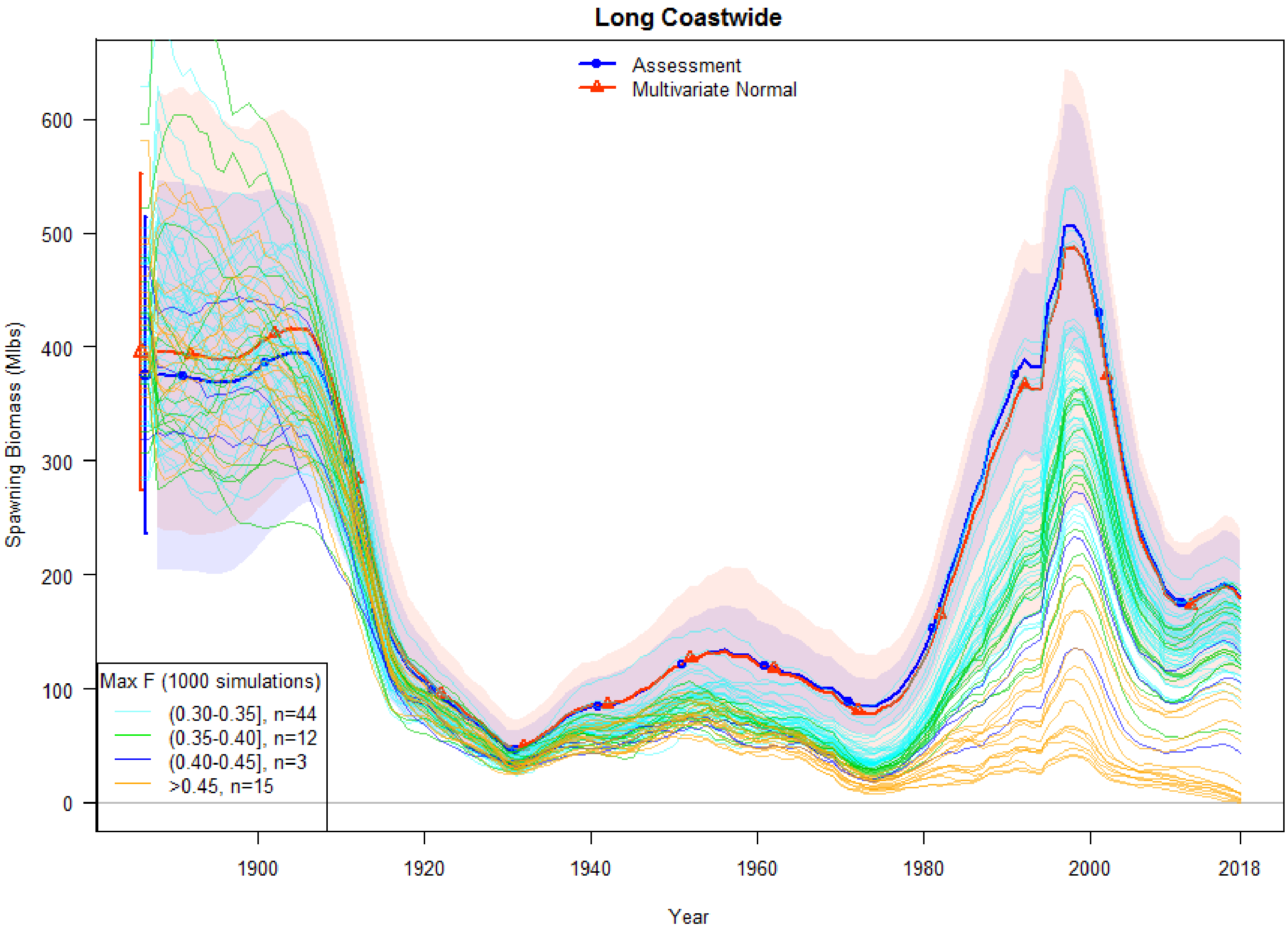


Fig. 3. Mean spawning biomass trajectories from the long coastwide assessment model with 95% confidence range (blue) and the mean and 95% confidence range of 1000 samples from a multivariate normal using the parameter estimates and inverted Hessian from the long coastwide assessment model (red). Individual trajectories from specific samples that produced large maximum F values are also plotted with the number of trajectories for various ranges of F listed in the legend.

Implementing a maximum F of 0.4 when sampling from the multivariate normal distribution (only the long coastwide was limited as short coastwide showed fishing mortality rates lower than 0.2), the assessment was mimicked reasonably well by the sampled trajectories for the long and short coastwide models (Figure 4).

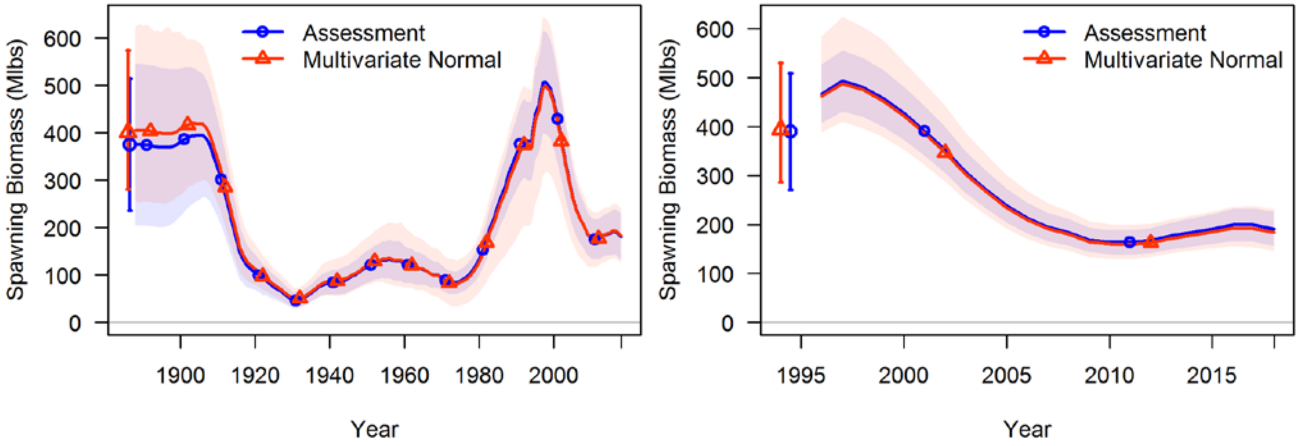


Fig. 4. Median spawning biomass trajectories from the long coastwide (left) and short coastwide (right) assessment models with a 95% confidence range (blue) and the median and 95% confidence range of 1000 samples from a multivariate normal using the parameter estimates and inverted Hessian from each assessment model (red).

Estimating parameters that were fixed in the assessment may produce stock dynamics that are not consistent with the assessment. To condition the OM to match the assessment, but introduce additional variability, the following steps were performed.

1. Allow for the estimation of the additional parameters in the assessment models. For the long coastwide model, steepness was estimated without a prior. For the short coastwide model, female M was estimated without a prior (and the upper bounds on female and male M's were increased to 0.45) and steepness was estimated with a prior created from the results of the long coastwide model and assuming a normal distribution. A prior on steepness was used to keep steepness within a reasonable range and force the estimated standard deviation for the short coastwide OM to be similar to the standard deviation in the long coastwide OM (i.e., both operating models are sampling from the same steepness distribution). Without a prior, the estimated variability in steepness resulted in a nearly uniform distribution between 0.2 and 1.0. The prior is centered around 0.75 with a standard deviation of 0.084 (2.5th and 97.5th percentiles equal to 0.59 and 0.91, respectively). See Figure 5 and the following steps.
2. Use the estimated covariance from the models with the extra parameters estimated (full model), the variances from the assessment model, and the variance of the additional estimated parameters from the full model to build a covariance matrix. Use the point estimates from the assessment model with that covariance matrix to sample from a multivariate normal distribution. This keeps the full model's predictions near the assessment model, but introduces extra variability accounting for correlation between estimated parameters.
3. Run the SS model using the sampled parameters, but without estimation to predict the historical population dynamics.
4. Eliminate the simulation if the maximum exploitation rate is greater than 0.4 in any year, or if the spawning biomass drops below 100 pounds in any year.
5. Repeat 2 through 4 as many times as necessary to create 1000 simulated trajectories.

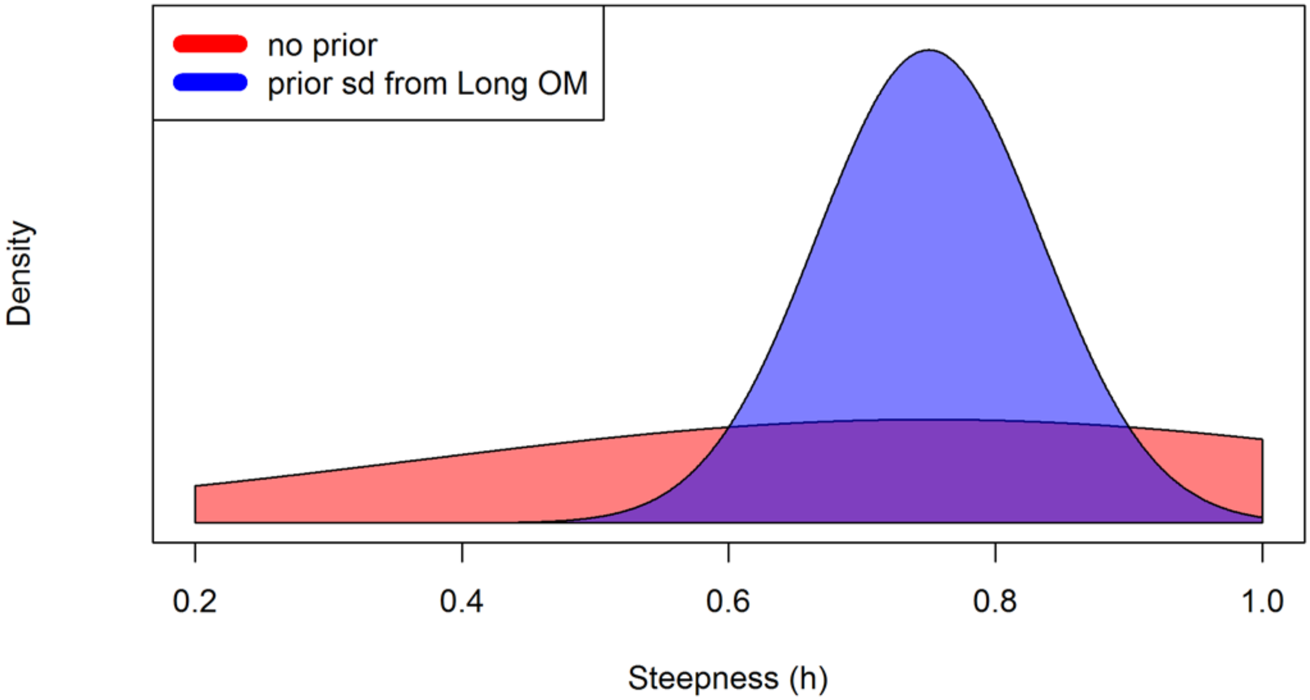


Fig. 5. Steepness Normal distributions centered around 0.75 using the standard deviations estimated without a prior in the short coastwide model (red) and with a prior determined from the long coastwide operating model (blue).

3.1.1.2 Long coastwide operating model

Steepness was the only additional parameter in the long coastwide operating model, compared to the assessment, that had variability. Steepness was centered on 0.75, as in the assessment, even though the estimated value of steepness was 0.9463, but the estimated variance (standard deviation = 0.08376) and covariances were used. The normal distribution of steepness, from which values were sampled, can be seen as the blue curve in Figure 5, and the estimated value (0.9463) is the 88th percentile in this distribution.

The parameters, including steepness centered around 0.75, were sampled from a multivariate normal distribution to create 1000 parameter vectors, each used to create a population trajectory. Trajectories that showed a maximum exploitation rate greater than 0.4 at any point in the time series were eliminated and parameters were re-sampled until 1000 acceptable parameter vectors were found. In total, 399 parameter draws were eliminated in the process. The final 1000 trajectories of historical spawning biomass from the operating model are compared to the assessment in Figure 6.

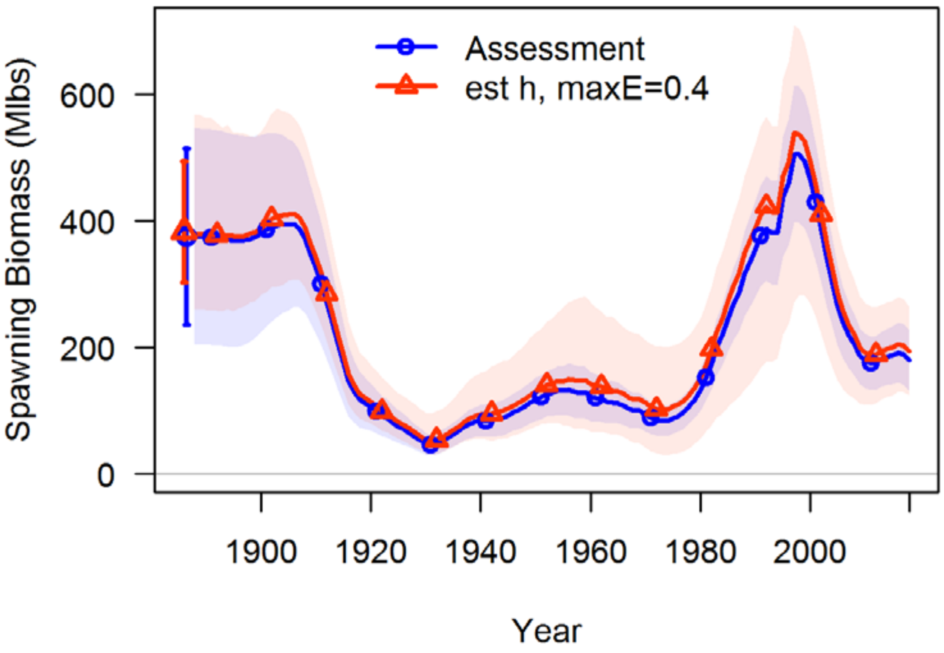


Fig. 6. Predicted median biomass trajectories with 95% confidence intervals for the long coastwide assessment model (blue) and the long coastwide operating model (red).

The median spawning biomass in the operating model is slightly greater than the assessment model. This is an effect of using a parametric bootstrap and adding the variability on steepness, even though the distribution of steepness was centered on the assessment value of 0.75. There are a number of reasons that the median of the operating model is slightly greater than the assessment model.

1. The distribution of spawning biomass from the operating model is broader and not necessarily symmetric, whereas the assessment model uses a point estimate (maximum likelihood) and an assumption that the variability in spawning biomass is characterized by a normal distribution.
2. The threshold maximum exploitation rate of 0.4 eliminates some low trajectories.
3. The covariances in the variance-covariance matrix used to characterize the normal distribution are from the full model (with steepness estimated) and are different than the covariances estimated in the assessment model. The variances of the parameters estimated in the assessment model are from the assessment model in the variance-covariance matrix used for sampling. Even setting the variance and covariances of the steepness parameter to zero in the variance-covariance matrix for sampling resulted in a median spawning biomass trajectory slightly above the assessment for most of the time-series, although it was similar to the assessment in recent years.

The 2018 point-estimate of spawning biomass from the assessment is the 36th percentile of the distribution of 2018 spawning biomass in the operating model (see Figure 7).

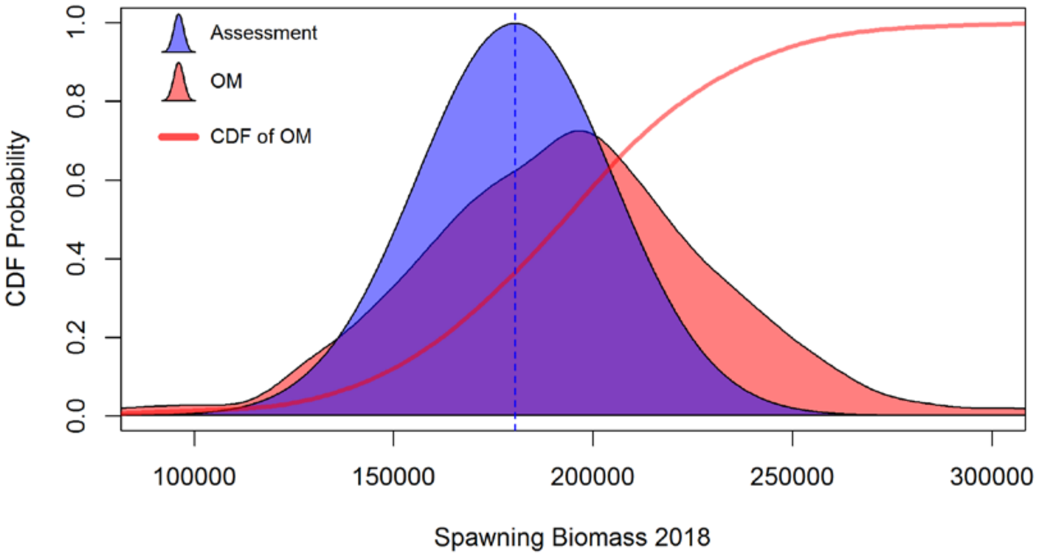


Fig. 7. Predicted distributions of 2018 spawning biomass for the long coastwide assessment model (blue) and the long coastwide operating model (OM, red). The cumulative distribution function (CDF) of the OM distribution and the median of the assessment 2018 spawning biomass (dashed blue line) are also shown.

3.1.1.3 Short coastwide operating model

Steepness and female natural mortality were the additional parameters in the full short coastwide model, compared to the assessment, that had variability. Steepness was centered on 0.75, as in the assessment. A prior was put on the steepness parameter (normal with a mean of 0.75 and a standard deviation of 0.08376, from the long coastwide model estimate of steepness), as discussed above, to make it have a similar distribution as the long coastwide model (see Figure 5). Female natural mortality was estimated without a prior, but the upper bound was extended to 0.45 because the estimate was 0.35. The upper bound on male natural mortality was also extended to 0.45 and its estimate was 0.26.

The estimated variances and covariances of steepness and female natural mortality were used, along with estimated variances and covariances from the assessment model for other parameters, to characterize the variance-covariance matrix used in the multivariate normal distribution from which parameters were sampled. The estimated standard deviations for steepness and female natural mortality were 0.08399 and 0.00864, respectively. The means for the multivariate normal distribution were the estimated or fixed values from the assessment (i.e., $h = 0.75$ and female $M = 0.15$).

The parameters, including steepness, were sampled from a multivariate normal distribution to create 1000 parameter vectors, each used to create a population trajectory. Trajectories that showed a maximum exploitation rate greater than 0.4 at any point in the time series were eliminated until 1000 parameter vectors were obtained. In total, 68 parameter draws were eliminated. The final 1000 trajectories of historical spawning biomass from the operating model are compared to the assessment in Figure 8.

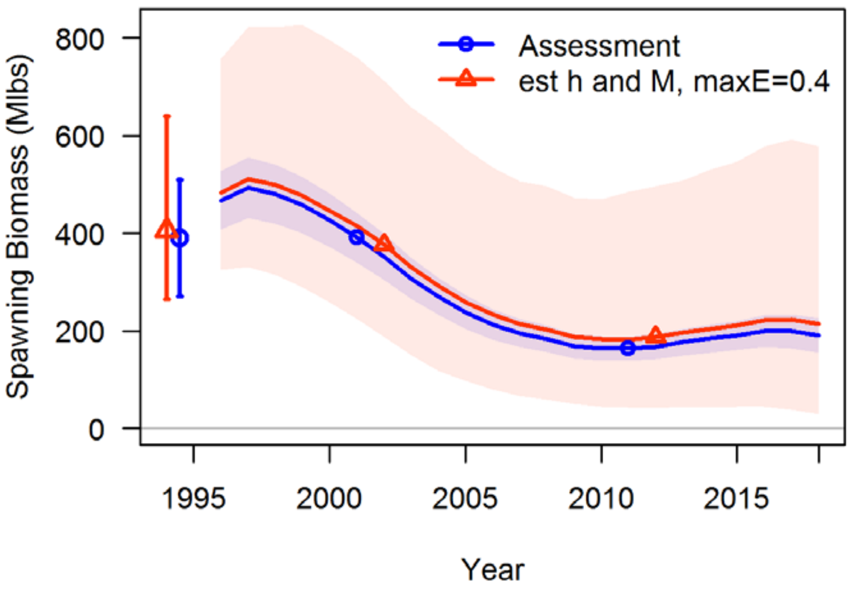


Fig. 8. Predicted median biomass trajectories with 95% confidence intervals for the short coastwide assessment model (blue) and the short coastwide operating model (red).

The median spawning biomass in the operating model is slightly greater than the assessment model. This is an effect of using a parametric bootstrap and adding the variability on steepness and female natural mortality, even though the distributions of these parameters were centered on the assessment values. This occurs for a number of reasons, as outlined above when discussing the long coastwide model.

The 2018 point estimate of spawning biomass from the assessment is the 44th percentile of the distribution of 2018 spawning biomass in the operating model (see Figure 9).

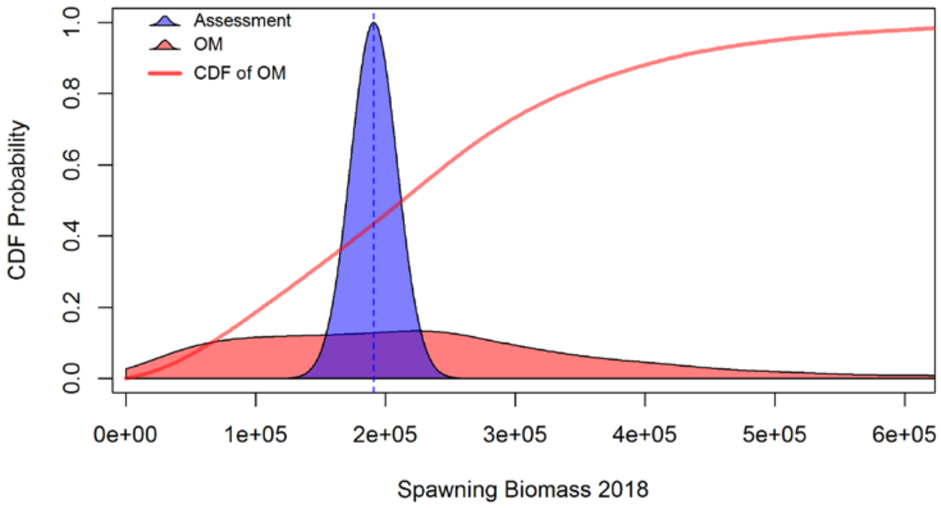


Fig. 9. Predicted distributions of 2018 spawning biomass for the short coastwide assessment (blue) and the short coastwide operating model (OM, red). The cumulative distribution function (CDF) of the OM distribution and the median of the assessment 2018 spawning biomass are also shown.

3.1.1.4 Summary of conditioned operating models

Overall, the individual operating models mimic the assessment well, but with additional uncertainty. The presence of a slightly higher median spawning biomass in the individual operating models is not a concern because the MSE is focused on ranking procedures and is not meant to predict the exact quantities. The most important aspect is to characterize variability and the dynamics of the stock. The variability in the short coastwide model is much greater than in the long coastwide model, and is a large contributor to the overall variability, in recent years, of the operating model consisting of the combination of the two individual models (Figure 10). When comparing the combined operating model to the ensemble assessment, the median spawning biomass trajectories are similar, but the variability in the operating model is much greater than the ensemble assessment (Figure 10).

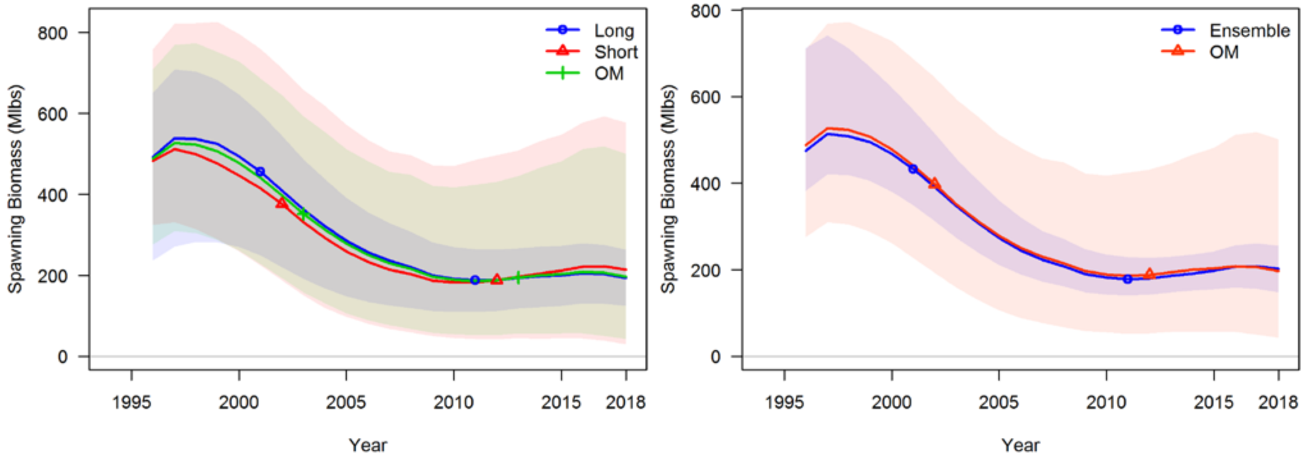
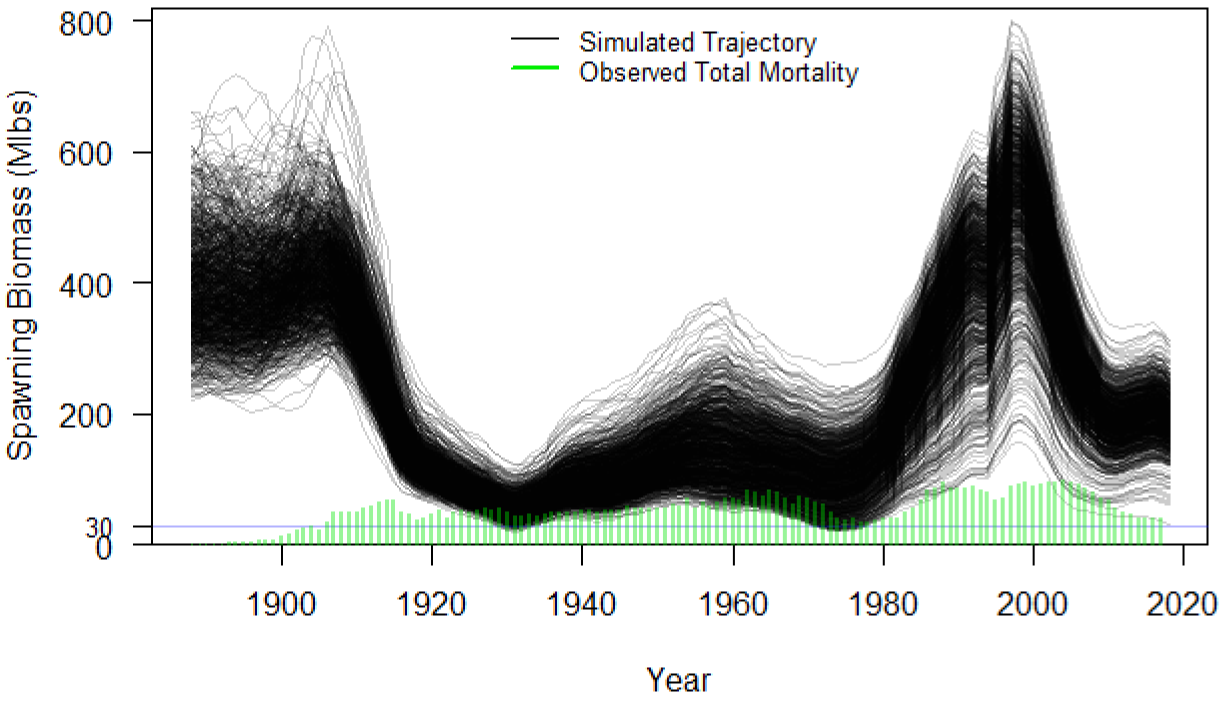


Fig. 10. The conditioned operating model (red) compared to the stock assessment ensemble (blue) with 95% confidence intervals on each.

The historical simulated trajectories were examined for evidence of “quasi-extinction”, which can be defined as a trajectory that reaches a value low enough that it would unlikely recover (in reality). That low value is not defined, so we compared simulated trajectories of spawning biomass to observed total mortality from all fisheries (Figure 11). The spawning biomass was generally low from around 1920 to 1980, and again in recent years. Especially low spawning biomass occurred near 1930 and 1975, and in recent years in the short coastwide model. The observed total mortality from fishing overlaps the lower trajectories around these low points, even with a maximum exploitation rate of 0.4. This can occur because the fishing mortality is partially composed of immature, young fish. Overall, some spawning biomass trajectories are surprisingly low, but it does not appear that quasi-extinction is apparent.



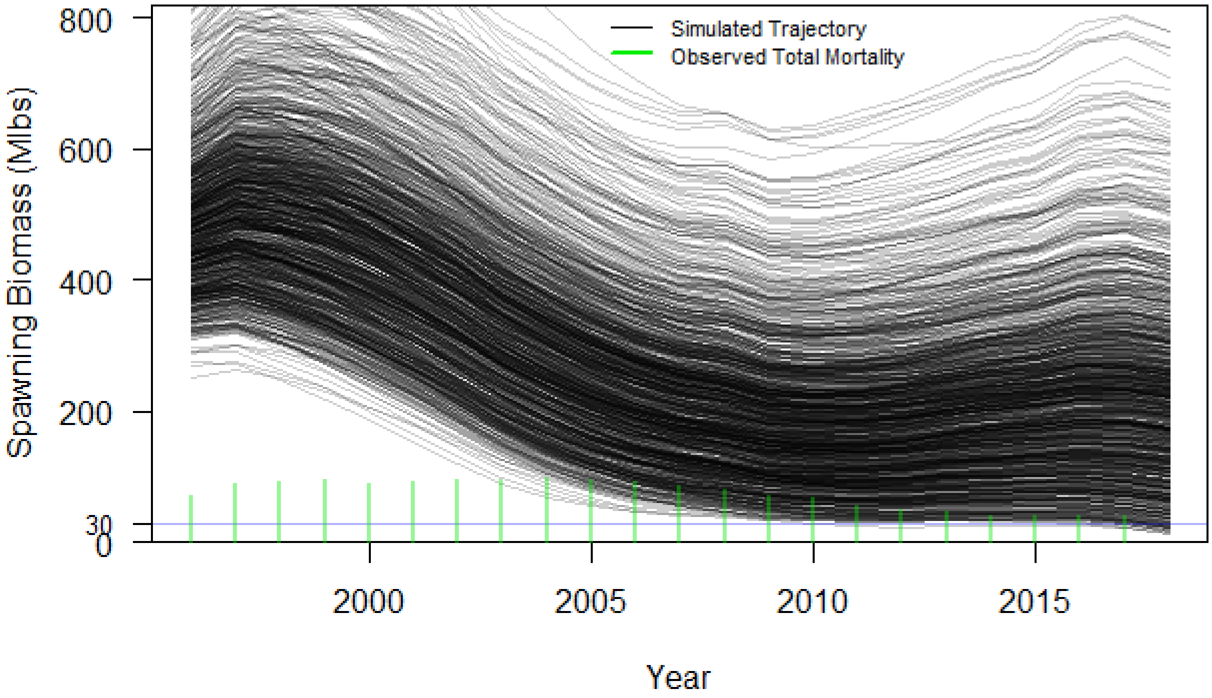


Fig. 11. Historical simulated trajectories of spawning biomass (M lbs) from the long coastwide operating model (top) and the short coastwide operating model (bottom). Observed total mortality (M lbs) from all fisheries is shown by the green histogram bars. A horizontal line at 30 million pounds is drawn for reference.

3.1.2 Simulating Forward with the Operating Model

The short and long coastwide models make up the operating model and incorporate variability associated with estimated parameters describing stock and fishery dynamics. Variability from other sources (e.g., weight-at-age, recruitment regimes, and allocation to fishery sectors) was introduced when projecting into the future. Descriptions of these procedures are provided in IPHC-2017-MSAB010-09 Rev1, and updates to the procedures are described here. An overview of major sources of variability are shown in Table 2.

Table 2. Processes and associated variability in the operating model (OM). TM refers to total mortality.

Process	**Uncertainty**
Natural Mortality (M)	Estimate appropriate uncertainty when conditioning OM
Recruitment	Random, lognormal deviations
Size-at-age	Annual and cohort deviations in size-at-age with bounds
Steepness	Estimate appropriate uncertainty when conditioning OM
Regime Shifts	Autocorrelated indicator based on properties of the PDO for regime shift
TM to sectors	See section on allocating TM to sectors
Proportion of TCEY	Sector specific. Sum of mortality across sectors may not equal coastwide TM

3.1.2.1 Allocating the Total Mortality to Fishery Sectors

There are five fishing sectors in simulations, as is defined in the coastwide assessment models. These are a commercial fishery, a discard mortality from the commercial fishery, a recreational fishery, bycatch mortality, and a subsistence fishery. The changes to the methods used to allocate total mortality to these five sectors are described below.

Bycatch Mortality

Bycatch mortality across all IPHC Regulatory Areas (Figure 12) has been declining since a peak in 1992 of 20 million pounds (~9,000 t). In 2017, bycatch mortality was estimated to be 6.0 million pounds (~2,700 t), which is due to industry measures to reduce bycatch as well as reductions in the Pacific halibut stock.

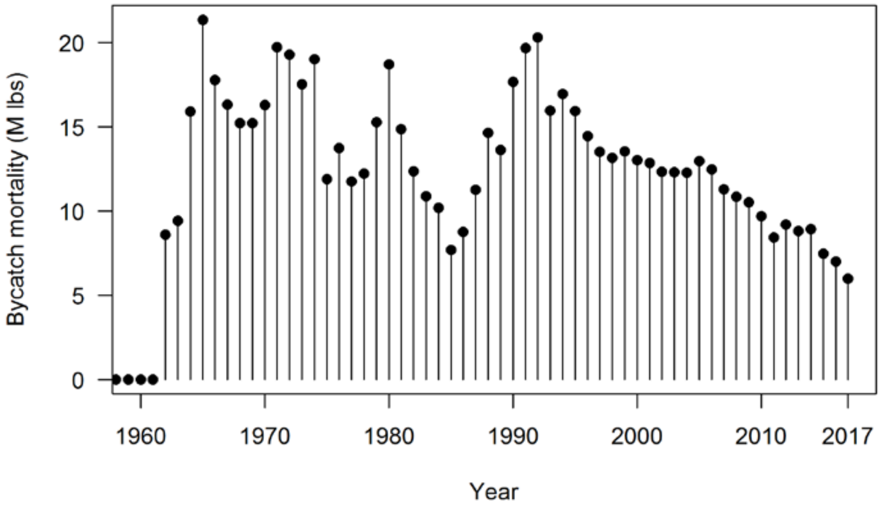


Fig. 12. Observed bycatch mortality.

A look at the historical relationship between bycatch mortality and total biomass was done to predict how bycatch may change with changes in Pacific halibut biomass. Before 1997 bycatch increased greatly with little change to total biomass (Figure 13) and after 2014 the bycatch dropped substantially with little change in total biomass (likely due to the industry specified protocols to reduce bycatch, such as deck sorting in the Amendment 80 trawl fleet). Therefore, using bycatch mortality from 1997 to 2014 and estimating the relationship with total biomass, the predicted slope of the line is 0.004. This is interpreted as each pound increase in total biomass results in a 0.4% increase in bycatch mortality. However, in the past three years, the bycatch mortality has declined from approximately 9 million pounds (4,000 t) to 6 million pounds (2,700 t) with little change in total biomass, thus the prediction line should reflect the efforts to reduce bycatch mortality, and the intercept was shifted to match the 2017 observations of bycatch mortality and total biomass (Figure 13). The predicted total biomass in 2017 was 848 million pounds (385 thousand t) which shifts the line downward by 3.4 million pounds to current bycatch levels but retains the relationship (change in bycatch) with total biomass.

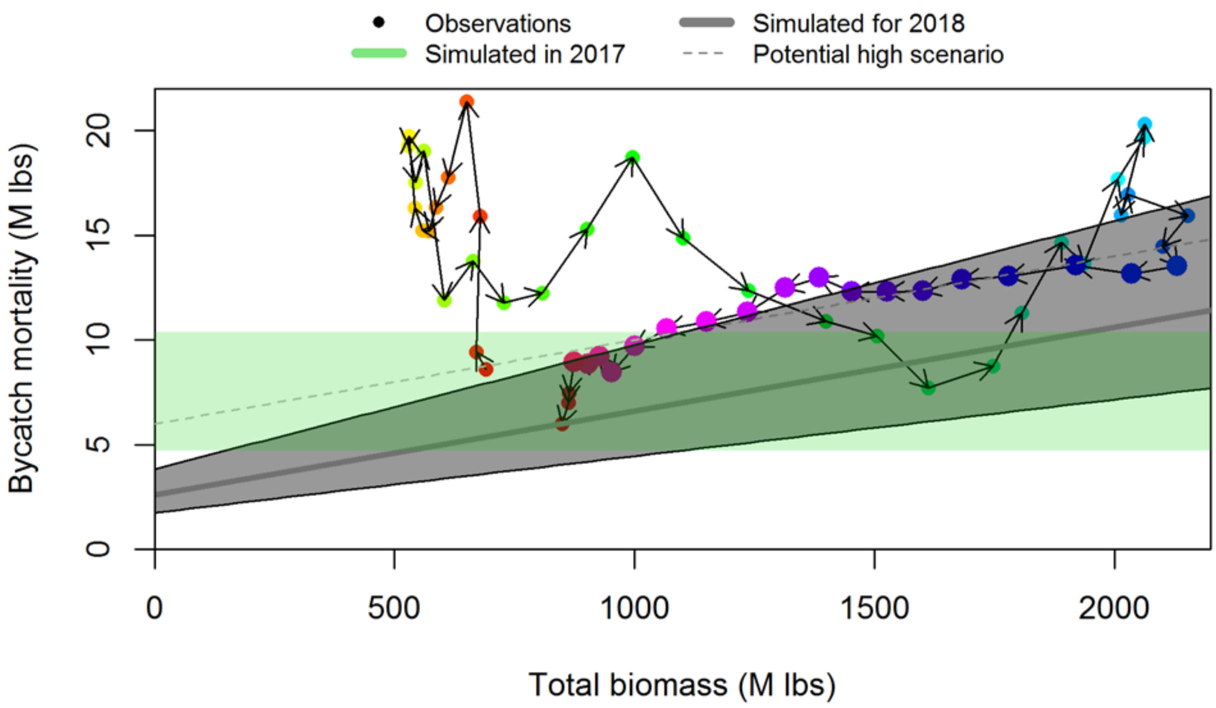


Fig. 13. Bycatch mortality (colored dots) plotted against estimated total biomass from the 2017 stock assessment. Arrows and colors show the sequence of time. The years 1997 to 2014 are shown by larger dots. The light green area shows the range of bycatch that was simulated from a lognormal distribution for 2017 MSE results, and did not change with total biomass. The grey areas shows the updated lognormal distribution for simulated bycatch that is a function of total biomass. The dashed line shows the mean of a potential high scenario for simulating bycatch.

A potential high bycatch scenario would be to use the original intercept of 6, which creates a line passing through the 1997-2014 observations (Figure 2, dashed line).

The previous CV on bycatch was 0.2 with a constant mean bycatch regardless of total biomass. This CV was kept to maintain the unpredictability of bycatch in the future.

Recreational mortality

A recommendation from MSAB012 was to modify the recreational allocation so that it kept increasing as the biomass (or TCEY) increased (REF to paragraph). Therefore, recreational mortality was investigated, and a constant proportion of the total mortality was used for allocation. To determine the proportion, the last five years (2013-2017) were used to determine the mean proportion, which was 0.18. The error on the proportion was set to capture the range of proportions observed over the past five years, resulting in a CV of 0.01. Figure 14 shows the recreational mortality and the proportion of recreational mortality plotted against the total mortality, as well as the simulated mean and range.

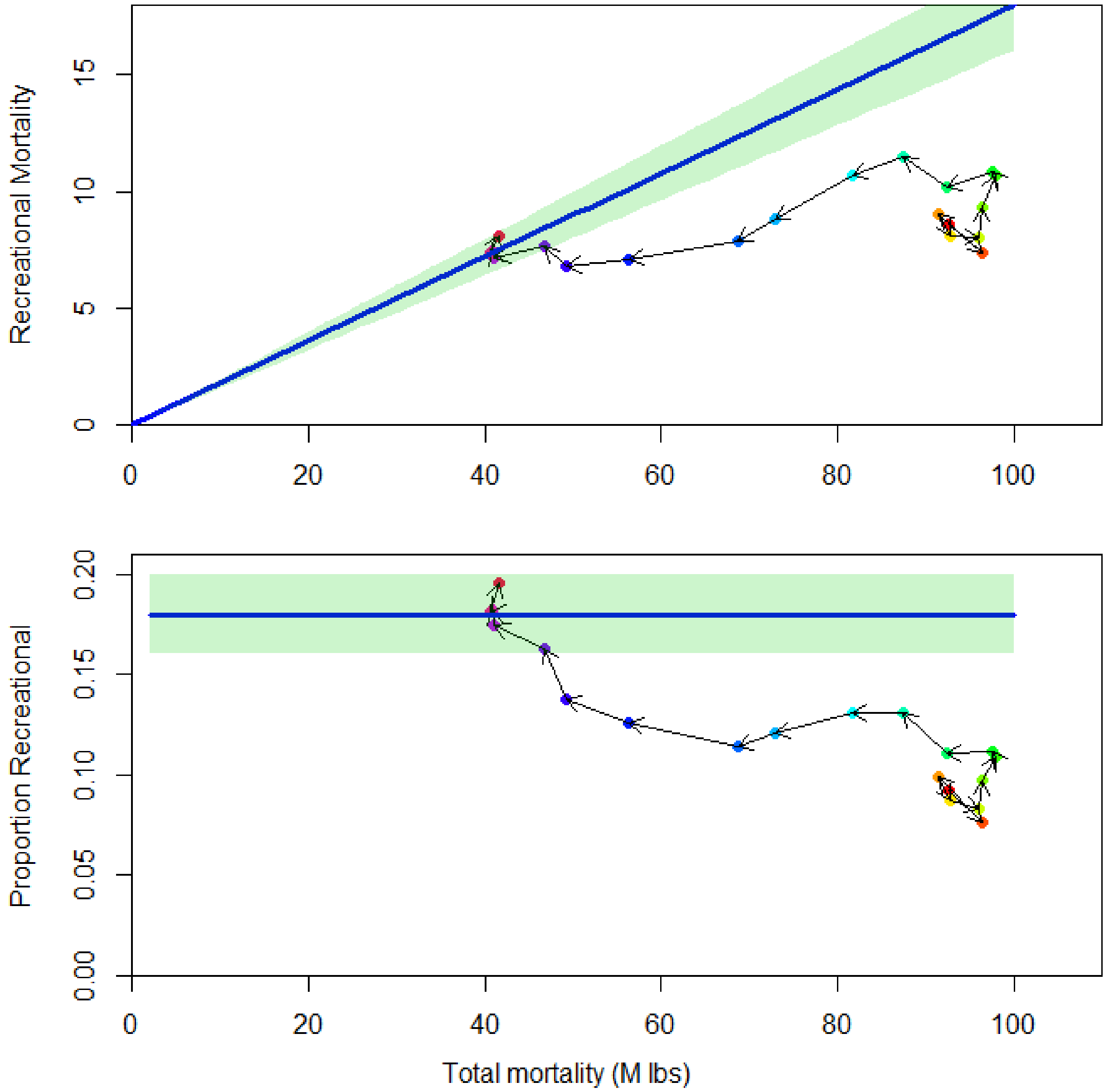


Fig. 14.: Recreational mortality (top) and the proportion of recreational mortality (bottom) plotted against the total mortality, as well as the simulated mean (blue line) and range (green area). Arrows show the sequence of time.

The resulting average allocations are shown in Figure 15.

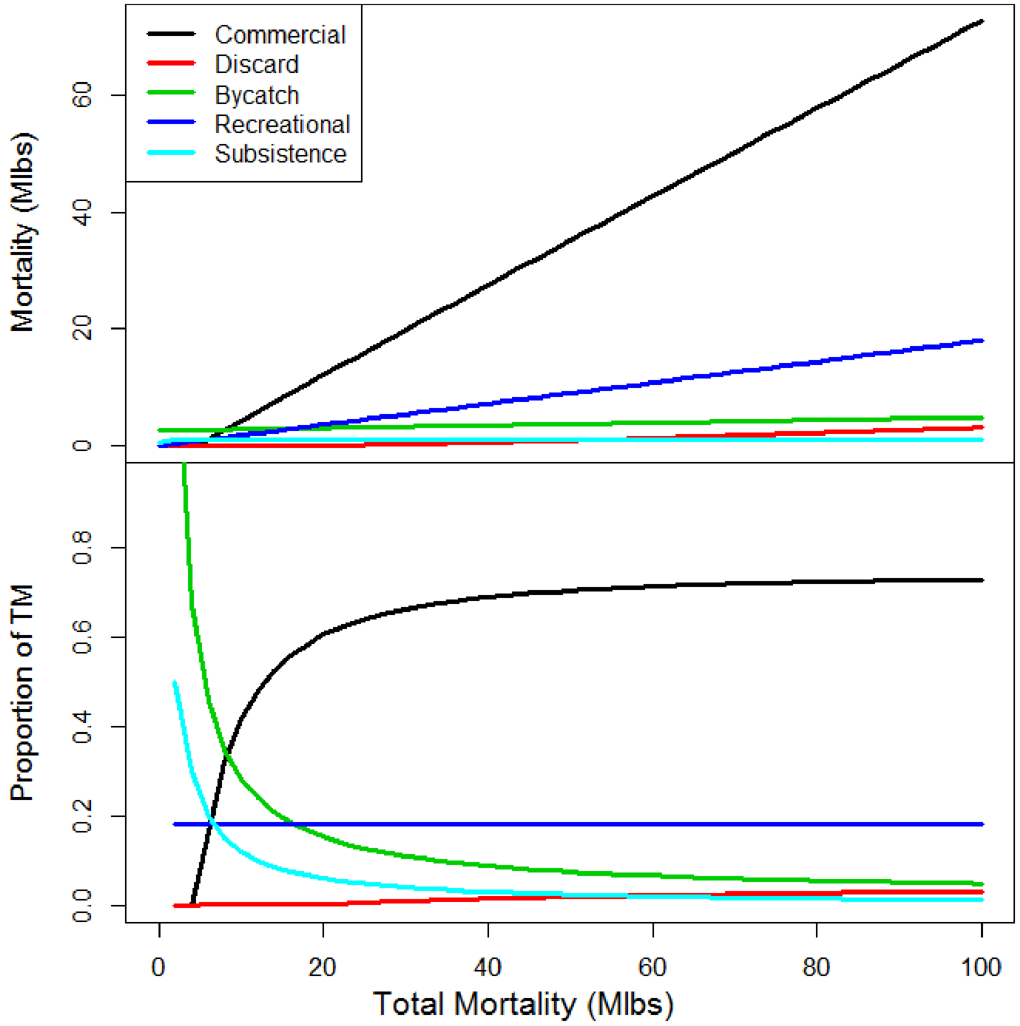


Fig. 15. Average allocations in terms of mortality (top) and proportion (bottom) for the five fishing sectors. Bycatch allocation is a function of total biomass, and it was assumed that total mortality is 17.5% of total biomass (based on estimates from 1998–2017).

3.1.2.2 Variability in Commercial Selectivity

Selectivity-at-age for the commercial sector is modeled in the long and coastwide models with a double-normal formulation. However, the descending width parameters are fixed such that the function in monotonic and asymptotes at one (i.e., full selectivity at older ages), and only two parameters are estimated: the ascending width (controlling how steep the ogive is) and the peak parameter (controlling where the ogive reaches a value of one. These two parameters are time-varying and result in year-specific selectivity ogives. Annual deviates are estimated and the changes in the parameters are a random walk from the previous year.

$$\mathrm{param}_{\{y\}} = \mathrm{param}_{\{y-1\}} + \mathrm{dev}_{y}$$

The estimated selectivity ogives for the commercial sector from the long coastwide model are shown in Figure 16.

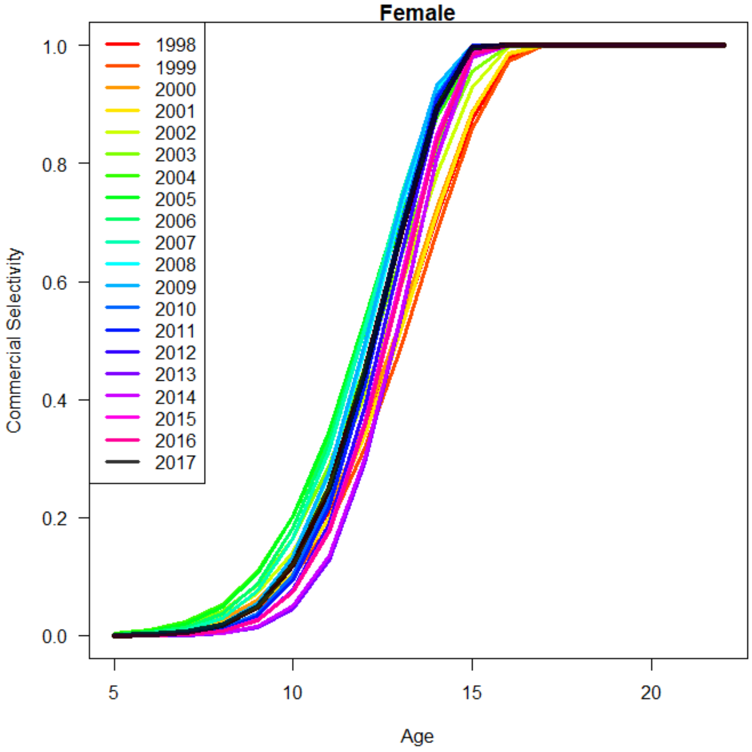


Fig. 16. Estimated commercial selectivity from the long coastwide model for the years 1998-2017.

Changes in selectivity may be related to changes in weight-at-age because weight-at-age is a proxy for changes in size. Given that the selectivity parameters are a random walk from the previous year's adjusted parameter, simply modeling the deviates as a function of weight-at-age is not clear, but modeling the adjusted parameter estimates as a function of weight-at-age is reasonable. There are likely many other factors affecting selectivity, such as economic conditions, bycatch, and other fisheries, thus only recent observations of weight-at-age and estimates of parameters were used. The current design of the survey began in 1998, which gives twenty years of observations with a large amount of data collected coastwide to inform the weight-at-age. Figure 17 shows that the selectivity parameters and weight at age 9 are correlated to some degree.

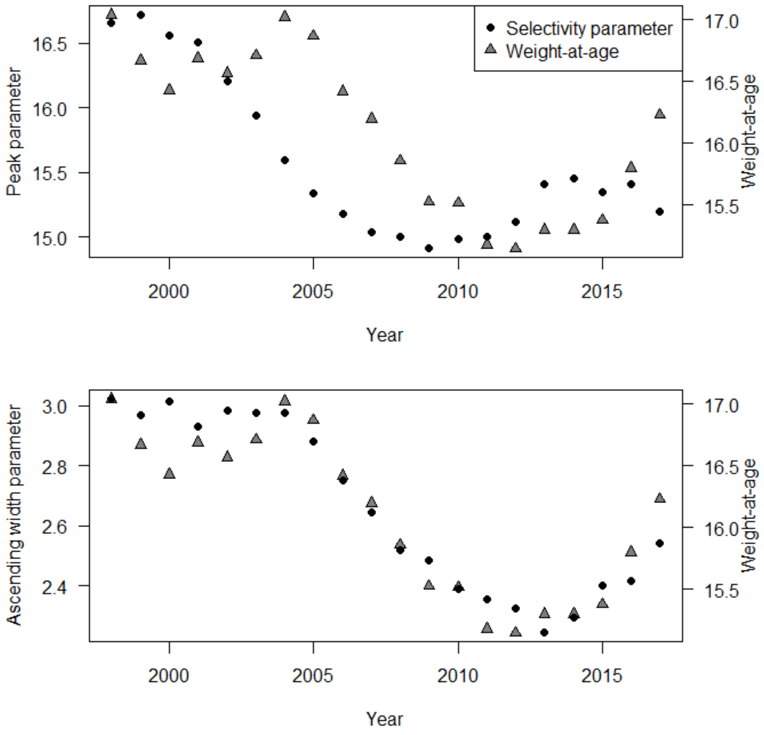


Fig. 17. Estimates of the peak (top plot) and ascending width (bottom plot) parameters for the years 1998-2017 (circles). Also shown are the observations of weight at age 9 (triangles) for those same years.

The estimates of the peak and ascending width parameters are highly correlated (Pearson correlation coefficient = 0.75) and show a positive relationship with weight at age 9 (Figure 18). When fish are growing to a larger size, the peak is shifted to the right and the ascending width is larger, resulting in an ogive that is less steep and the increasing portion is spread out more ages. It may seem counter intuitive that the peak is shifted to the right (older ages) when the fish are growing faster (i.e., they should be selected at an earlier age if the process is truly size-based). However, the ascending width parameter is increasing, and we believe it does that to select more of the fish that are not fully selected (Figure 16, ages 8-12, green lines). Then, because of inflexibility in the two-parameter approach, the peak parameter is shifted to older ages to accommodate the informative data that occurs at the younger ages.

Therefore, it appears that the ascending width parameter is driven by weight at age 9, and the peak parameter is related to the ascending width parameter estimate. The linear regression line for the relationship between the ascending width parameters and weight at age 9 had a R^2 value of 0.9 and showed a positive slope (Figure 18). The relationship of the peak parameter to the ascending width parameter seems to be two phases: a small peak parameter with small variability when the ascending width parameter is small, and a higher peak parameter with a larger variability when ascending width parameter is large. This was simulated with two states of the peak parameter, with a linear connection between ascending width values of 2.7 and 2.9 (Figure 18). The relationship was captured when relating the peak parameter to weight at age 9, even though weight at age 9 was not used directly to predict it. The correlation between the peak and ascending width parameters was 0.88, without extra variability.

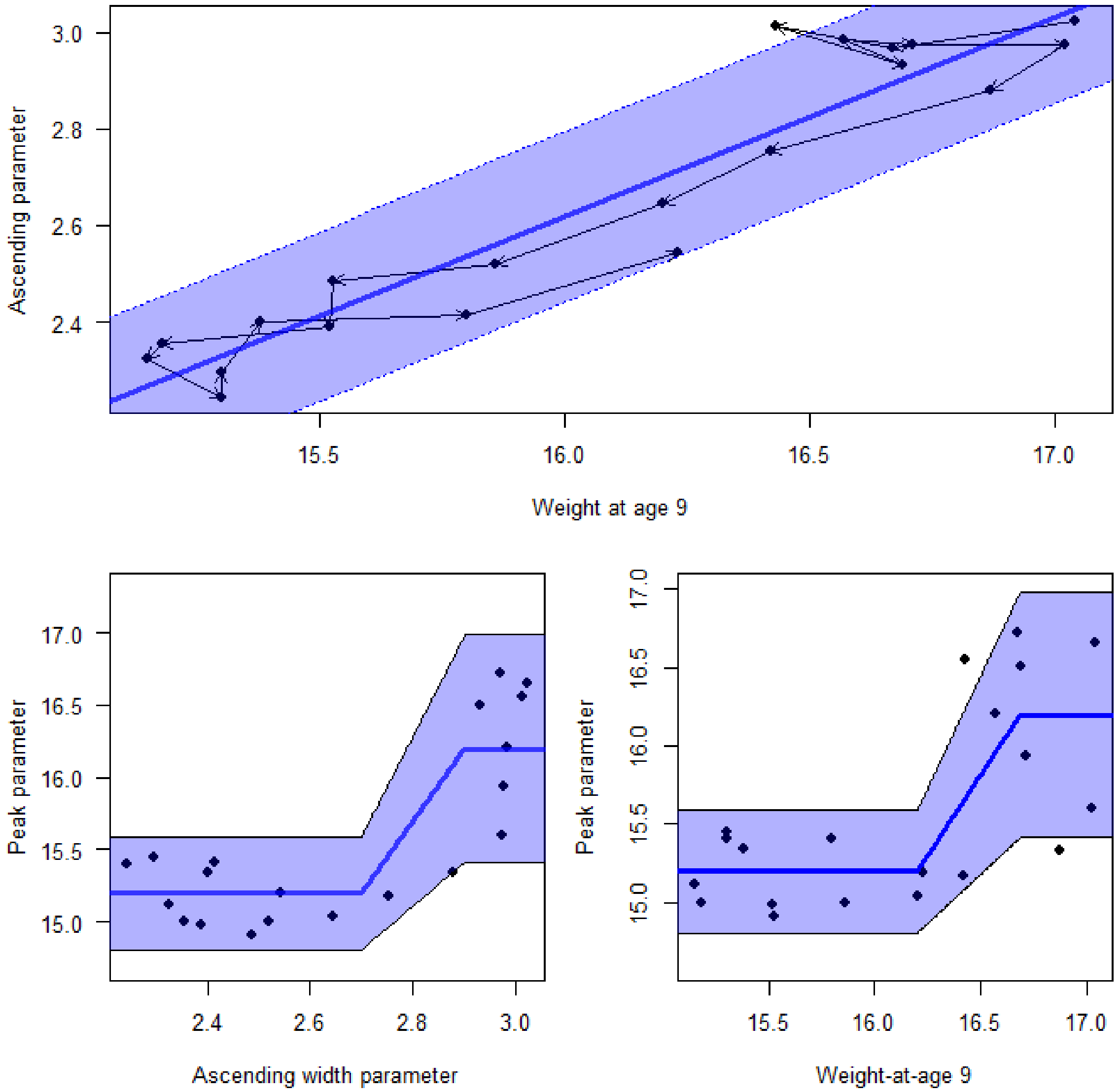


Fig. 18. Estimates of the ascending width parameters for the years 1998-2017 plotted against weight at age 9 (top). The blue line is the fitted regression line. The bottom row shows the peak parameter plotted against the ascending width parameter and against the weight at age 9. The blue shaded area is the 95% interval for the simulated values.

With the short model, weight at age 9 had a high correlation with the ascending width parameter (R^2=0.60. Figure 19). The peak parameter had very little variation (ranged between 15.57 and 15.78), thus was considered to not be time-varying. The male asymptote was time-varying in the short coastwide model, but there was not clear relationship with any weight-at-age. Figure 20 shows the relationship with weight at age 9 and seems well correlated in early part of the time series, but varies just as much with little change the weight at age in more recent years. Therefore, the male asymptote is simulated as a random walk.

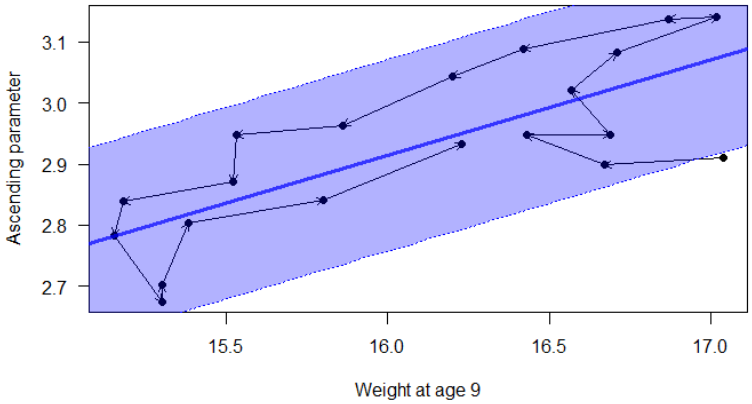


Fig. 19. Estimates of the ascending width parameter of the female selectivity ogive in the short coastwide model for the years 1998-2017 plotted against weight at age 9. The blue line is the fitted regression line.

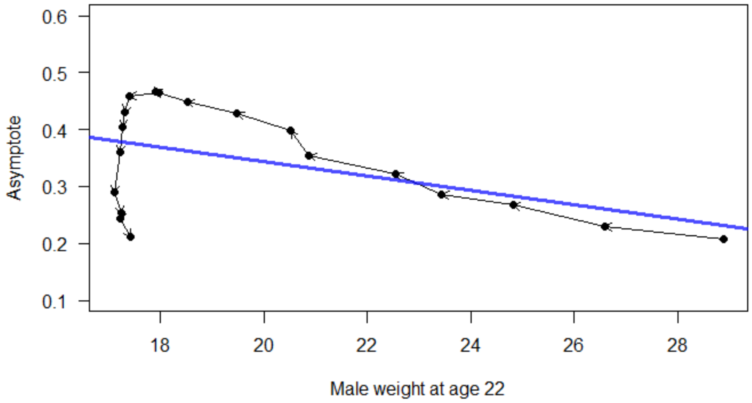


Fig. 20. Estimates of the asymptote of the male selectivity ogive for the years 1998-2016 plotted against weight at age 22. The blue line is the fitted regression line.

An example of simulated selectivity at age 10 and age 15 is shown in Figure 21. The parameters were bounded so that they did not traverse to values outside of the estimates for the last two decades. Overall, the selectivity shows a randomness that is linked to weight-at-age but not completely driven by weight-at-age. This is likely due to the spatial availability of specific year-classes as the distribution of landings has changed over time (Stewart and Martell 2014).

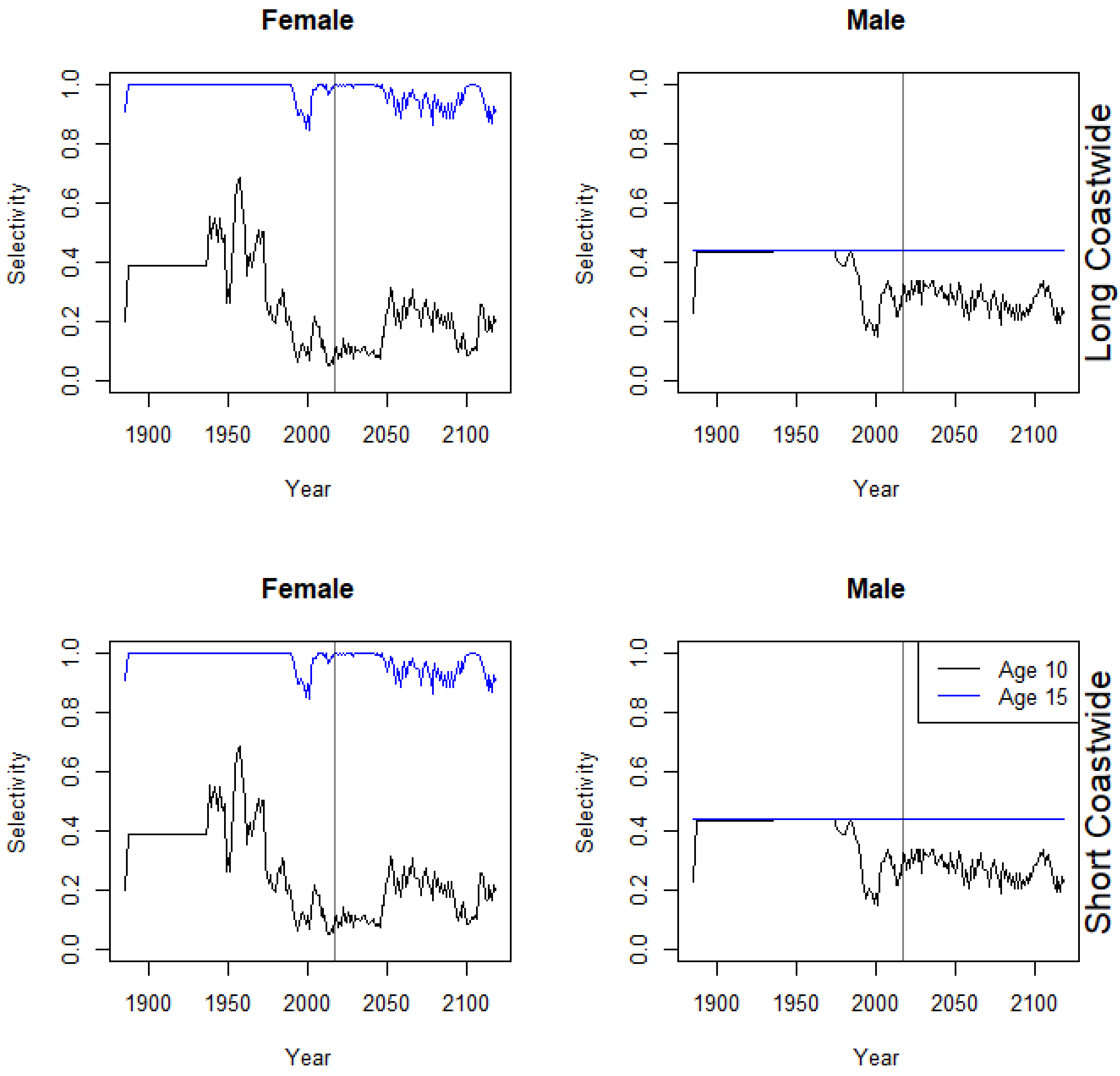


Fig. 21. Example of simulated commercial selectivity at age 10 and age 15 for the long coastwide model (top) and the short coastwide model (bottom). The vertical grey line is at the year 2018.

3.2 MANAGEMENT PROCEDURE

The elements of the management procedure are described in reverse order because it is easier to understand the decisions made for modelling them since they are dependent on each other. Therefore, the harvest rule is presented first, followed by the estimation model, and finishing with monitoring.

3.2.1 Harvest Rule

The generalized management procedure to evaluate is shown in Figure 1, but the focus will be on the Scale portion to produce results for the MSAB to evaluate before AM095 in 2019. Specifically, the portion of the management procedure being evaluated is a harvest control rule (Figure 22) that is responsive to stock status and consists of a procedural SPR determining fishing intensity, a fishery trigger based on stock status that determines when the fishing intensity begins to be linearly reduced (note that this may differ from the biological threshold), and a fishery limit that determines when there is theoretically no fishing intensity (this may differ from the biological limit). For these simulations, the two coastwide models were used, thus mortality only needed to be distributed to the five coastwide sources of mortality (directed commercial, discard mortality, bycatch mortality, recreational, and subsistence).

Simulations have been used in the past to evaluate a range of SPR values from 25% to 60% and trigger values of 30% and 40% (IPHC-2017-MSAB10-09 Rev 1). Those simulations provided insight into how those different levels of SPR would meet the objectives defined by the MSAB, but few values of SPR below 40% were tested. Future simulations will use a finer resolution of SPR values ranging from 30% to 56% and fishery trigger points of 30% and 40% (with the addition of 45% if time allows).

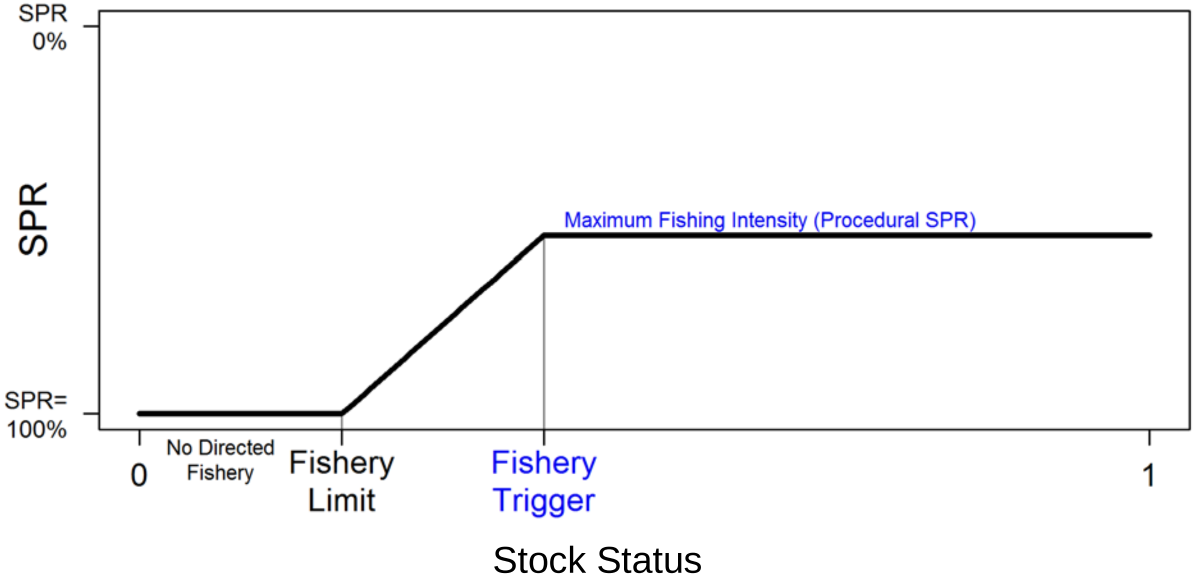


Fig. 22: A harvest control rule responsive to stock status that is based on Spawning Potential Ratio (SPR) to determine fishing intensity, a fishery trigger level of stock status that determines when the fishing intensity begins to be linearly reduced, and a fishery limit based on stock status that determines when there is theoretically no fishing intensity (SPR=100%). In reality, it is likely that only the directed fishery would cease. The Procedural SPR and the Fishery Trigger (in blue) are the two values to be evaluated.

3.2.2 Estimation Model

Two options to simulate an estimation model will be used: the No Estimation Model (previously called Perfect Information) option, as was used in past simulations, and the Simulate Error option. The No Estimation Model method assumes that the population values needed to apply the management procedure are exactly known (e.g. spawning biomass). This option is useful as a reference to better understand the performance with and without uncertainty in an estimation model. Due to time constraints, the only other option to be considered for simulations in 2018 is the Simulate Error option, which will be suitable to understand the effects of estimation error. This method is described below.

The harvest control rule contains two components that have estimation error. The first component is the estimated total mortality determined from the specified SPR. The second component is the estimated stock status that is used to reduce the fishing intensity when stock status is low (fishery trigger and fishery limit). These components are dependent on the estimated biomass, but it is more straightforward and computationally efficient to introduce error into these two components, rather than introducing error on the estimated biomass and then determining the resulting estimates of total mortality and stock status.

The 2017 stock assessment (Hicks & Stewart 2018) was used to determine a reasonable amount of variability in these two components and the correlation between them (see Section 4.2 of IPHC-2018-SRB012-08).

Autocorrelation is implemented by independently applying it to the deviation of the estimated stock status and the estimated total mortality. The correlated variability in these two quantities is applied and then the autocorrelation occurs independently using equation 1.

$$\left(\hat{X}_t - X_t\right) = \rho\left(\hat{X}_{t-1} - X_{t-1}\right) + \sqrt{1-\rho^2}\varepsilon_X \qquad (1)$$

Where $\left(\hat{X}_{t-1} - X_{t-1}\right)$ is the deviation for the quantity of interest (TM or stock status) in time step *t*, ρ is the autocorrelation parameter, and ε_X is a randomly generated deviation from a multivariate normal distribution for TM and stock status (as described above).

Overall, there are many assumptions in this incorporation of estimation error, but we are only trying to determine a reasonable amount of error for the simulations. Coefficients of variation on stock status and total mortality were fixed at 15% with a correlation of 0.5. Autocorrelation was fixed at 0.2. Other levels of error will likely be simulated to determine how sensitive the results are to the assumed estimation error.

3.2.3 Monitoring (Data Generation)

The simplified incorporation of estimation error will be used due to time constraints; thus no data are required to be generated. However, if a stock assessment were simulated, there would be many sources of data to generate.

3.3 SUMMARY OF THE FRAMEWORK

A summary of the major specifications for each component is provided below, with the components listed in a specific order where the next component is dependent on the decisions for the previous components.

1) Operating Model
    a) Stock synthesis, based on coastwide assessment models (short and long models).
    b) Five fleets, as in the assessment models (commercial, discards, bycatch, sport, personal use).
    c) Fishing mortality assigned to sectors based on historical information (with variability).
    d) Uncertainty incorporated through parameter uncertainty, model uncertainty, a simulated variability in future weight-at-age and recruitment.
2) Management Procedure
    a) Estimation Models

    i) Perfect Information (as a reference if we knew population values exactly when applying the harvest rule).
    ii) Simulate error in total mortality (cv=0.15) and spawning biomass (cv=0.15), with autocorrelation (0.2), from the simulated time-series to mimic an unbiased stock assessment.
b) Data Generation
    i) Not needed at this time.
c) Harvest Rule
    i) Coastwide fishing intensity (F_{SPR}) using a procedural SPR (to be evaluated).
    ii) A fishing trigger to reduce the fishing intensity (increase SPR) when stock status is below a specified level (to be evaluated).
    iii) A fishing limit to cease directed fishing when the stock status is less than a specified value (20%).

4 PERFORMANCE METRICS

Defining goals and objectives is a necessary part of a management strategy evaluation (MSE) which should be revisited often to make sure that they are inclusive and relevant. The MSAB is currently refining goal and objectives (see IPHC-2018-MSAB012-06), which are translated into performance metrics. Many performance metrics have been developed by defining a measurable outcome, a probability (i.e. level of risk), and time-frame over which it is desired to achieve that outcome. Management procedures can then be evaluated by determining which ones meet various objectives (via the performance metrics). Some performance metrics have been defined by the MSAB that are called statistics of interest, and even though they are associated with various objectives, they are secondary to the evaluation of the management procedure. Some of the primary performance metrics and statistics of interest being reported are described in Table 3.

5 SIMULATION RESULTS

Using the simulation framework described above and in previous documents, test cases were first investigated to better understand the dynamics of the simulations. The simulations were done with no directed fishing, but with bycatch and subsistence fishing (approximately ranging from 4.5 million pounds to 12 million pounds), to investigate the nature of the projections and the presence, if any, of quasi-extinction. Additionally, projections with constant levels of weight-at-age and recruitment (low/high combinations) were done.

Figure 23 shows forward simulation results for the no directed fishing case with simulated variability in weight-at-age and simulated recruitment regimes. Only one-hundred trajectories were simulated, but it is clear that the entire range of variability is not captured until at least after 60 years. As also shown in the conditioning results, the short coastwide model had a wider range of variability. No simulated trajectory for the long coastwide model produced a spawning biomass less than 30 million pounds, and the minimum spawning biomass from all long coastwide model trajectories was near 60 million pounds, which occurred at time step 2. The short coastwide model produced four (out of 100) trajectories that had a spawning biomass less than 30 million pounds. Of these four, three of them started at a spawning biomass less than 30 million pounds, and all three recovered to levels above that. One trajectory started above 30 million pounds, but eventually crashed to zero.

Table 3. Performance metrics and statistic of interest for the long-term to evaluate the management procedures. Primary metrics are the main performance metrics for the evaluation and the statistics of interest are intended to supplement and inform that evaluation.

Primary Metrics	
Performance metric	**Description**
P(SB > SB_{Lim})	Times out of 100 that the stock biomass (status) is above the limit. The limit is defined as 20% of the biomass if no fishing had occurred.
P(AAV > 15%)	Times out of 100 that the average annual variability (AAV) is greater than 15%. AAV can be thought of as the average change in the Total Mortality quota (TMq) from year to year.
P(TM < TM_{min})	Times out of 100 that the Total Mortality quota (TMq) would be set below a minimum value. The minimum TMq has not been determined, and is currently an *ad hoc* value of 34 Mlbs, which is the minimum Total Mortality observed (TM) since 1906.
Secondary Metrics	
Statistic of interest	**Description**
Median SB	The median biomass expected in the long-term
Median # females	The median number of females expected in the long term.
AAV	The Average Annual Variability, which can be thought of as the average change in the TM from year to year.
P(↓TM > 15%)	Times out of 100 that the TMq decreases by more than 15% compared to the previous year.
AAV|SB<SB_{Trig}	The average annual variability when the stock status is below the fishery trigger (often referred to as 'on the ramp').
$P\left(\widehat{SB} < SB_{Trig}\right)$	Times out of 100 that the estimated spawning biomass (status) is less than the fishery trigger, thus invoking 'the ramp' and reducing fishing intensity.
Median TMq	Median coastwide TMq. The TMq is greater than this value in half of the simulations.
P(TMq < 54)	Times out of 100 that the TMq is less than 54 Mlbs, which is 70% of the average TM from 1993 to 2012.
5th & 75th TMq	The 5th and 75th percentiles of the Total Mortality quota from the simulations. This means that 5 out of 100 are less than or equal the 5th percentile and 25 out of 100 are greater than or equal to the 75th percentile.

Specific states of weight-at-age and recruitment regimes were simulated to investigate how these factors, and the combination of them, affect the simulated population trajectories. Low and high recruitment regimes were simulated by fixing the regime in the model at its low or high value since it is modeled as discrete low or high. Changes in weight-at-age are continuous, thus specific states had to be determined. Low, medium, and high states are determined by calculating the 15th, 50th, and 85th percentiles of the historical weight-at-age (1935-2017) for each age, running a loess smoother through the specific quantile-at-ages, and then making sure it increases monotonically over age by predicting weight (from the loess model) for any ages that had a weight less than the weight at a younger age (Figure 24).

Using the low and high states of weight-at-age, crossed with the low and high recruitment regimes, and keeping them static for the entire simulation allowed for the investigation of these different factors as well as testing to make sure that they produced reasonable results. Figure 23 shows the simulated trajectories using the long coastwide model and the short coastwide model for the four different combinations. The long coastwide model was most influenced by weight-at-age, and each combination produced a well-defined band of trajectories. The short coastwide model showed more influence from recruitment with the high weight low recruitment scenario showing similar trajectories as the low weight high recruitment scenario. Some trajectories in the low weight low recruitment scenario showed quasi-extinction. In both models, the high recruitment regime resulted in more variability.

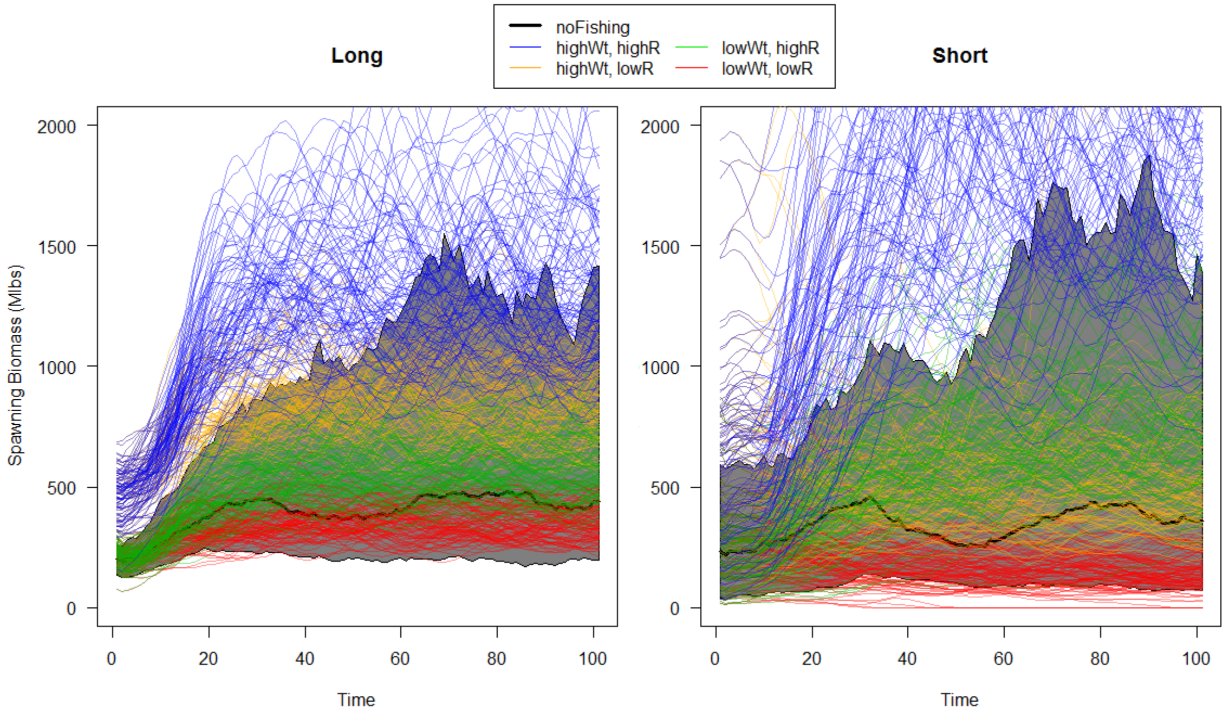


Fig. 23. One-hundred forward simulated trajectories of spawning biomass without directed fishing. Bycatch mortality and subsistence mortality occurred (note, bycatch is simulated as a constant level with error for these trajectories). The gray area shows the range of simulations between the 2.5th and 97.5th percentiles with no fishing, but with simulated weight-at-age and simulated recruitment regimes. The individual lines of different colors show individual simulated trajectories with specific constant levels of weight-at-age and recruitment.

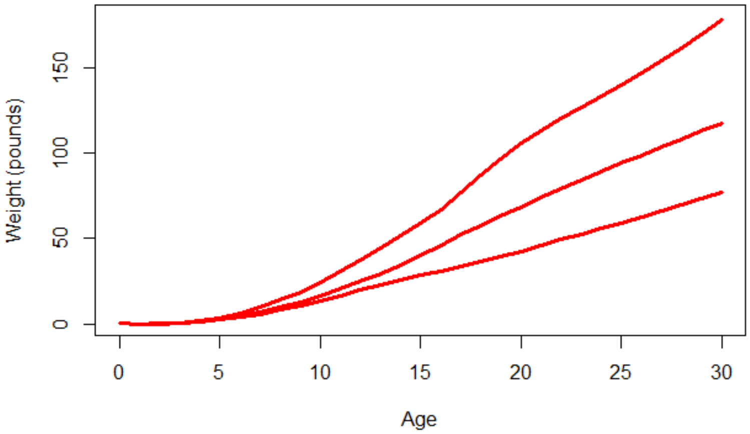


Fig. 24. Plot of the low, medium, and high states of weight-at-age for testing.

Table 4. Performance metrics for four different management procedures. The additional two columns of performance metrics show the effect when estimation error and autocorrelation are introduced. (*Note the results in this table are superseded by the results presented in Appendix A*).

Control Rule	**30:20**	**30:20**		**40:20**	**40:20**		**30:20**	**30:20**
SPR	**0.40**	**0.46**		**0.40**	**0.46**		**0.46**	**0.46**
Est Error	**None**	**None**		**None**	**None**		**0.1**	**0.1**
Autocorrelation	**NA**	**NA**		**NA**	**NA**		**0.0**	**0.2**
Metric								
P(SB < 20%)								
P(AAV > 15%)								
P(TM < 34)								
Median SB								
Median # females								
P(SB < 30%)								
AAV								
P(↓TM > 15%)								
Median TMq								
5th & 75th TMq								

Table 4 presents a small sample of results that will be shown at MSAB012. Additional results and alternative ways to view those results will be discussed at the meeting.

6 RECOMMENDATIONS

That the MSAB:

1) **NOTE** paper IPHC-2018-MSAB012-07 which provides the MSAB with an update on the MSE framework and presents a small subset of results.
2) **NOTE** the simulation framework and improvements to the simulation framework
3) **NOTE** the results of simulating forward in time with no fishing and the influence of weight-at-age and recruitment regimes.
4) **NOTE** the performance metrics reported for various management procedures.
5) **RECOMMEND** additional ways to present the results and examine trade-offs between objectives.
6) **RECOMMEND** a management procedure that meets the goals and objectives defined by the MSAB.

7 APPENDICES

Appendix A: Updated results – Long- and short-term performance metrics

Appendix A: Updated results

Tables A1 and A2 show some long-term performance metrics for some runs requested at MSAB011 (IPHC-2018-MSAB011-R). Tables A3 and A4 show some short-term performance metrics for those same runs. For long-term results with a control rule (Figure A1), the probability that the stock is below 20% of the dynamic unfished equilibrium biomass is less than 1% for all cases. This is a result of the control rule limiting the fishing intensity as the stock approaches this threshold, even with estimation error present. It is rare that the estimation persists such that fishing intensity remains high and the stock falls below the 20% threshold. The outcome of this can be seen in the average annual variability (AAV), which is a measure of the change in the quota from year to year. At fishing intensities greater than that associated with an SPR 0f 40% (i.e., SPR values less than 40%) the probability that the AAV is greater than 15% is more than 0.90. This probability declines to 0.61 at an SPR of 56%. The median AAV's range from 16% to 42% when using a 30:20 control rule (Table A1) and from 21% to 46% when using a 40:20 control rule (Table A2). The 40:20 showed higher variability in the quota. The absolute value of the Total Mortality quota ranged from 34% to 42% and was highly variable for a given SPR (Figure A1). In summary, long-term performance metrics showed little risk of falling below the 20% threshold, high variability in catches that increased with higher fishing intensities (i.e., lower SPR), and median Total Mortality quotas that increased slightly with greater fishing intensity.

Many more results and sensitivities will be shown at MSAB012.

Table A1. Long-term performance metrics for an estimation error CV of 0.15, autocorrelation of 0.4, a 30:20 control rule, and a range of input SPRs.

Input Est Error	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**
Input Autocorrelation	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**
Input Control Rule	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**
Input SPR	**30%**	**32%**	**34%**	**36%**	**38%**	**40%**	**42%**	**44%**	**46%**	**48%**	**56%**
Median SPR	42.6%	42.4%	42.4%	42.5%	42.7%	43.5%	44.5%	45.9%	47.4%	49.0%	56.3%
Biological Sustainability											
Median average dRSB	30.4%	31.0%	31.7%	32.9%	33.9%	35.0%	36.5%	37.9%	39.7%	41.6%	50.2%
P(all dRSB<20%)	0.004	0.004	0.006	0.005	0.002	0.002	0.003	0.004	0.003	0.002	0.002
P(any dRSB_y<20%)	0.011	0.008	0.009	0.006	0.002	0.002	0.004	0.005	0.004	0.003	0.002
P(all dRSB<30%)	0.470	0.405	0.338	0.253	0.191	0.142	0.094	0.065	0.031	0.023	0.002
P(any dRSB_y<30%)	0.867	0.789	0.676	0.545	0.402	0.307	0.202	0.149	0.07	0.044	0.003
Fishery Sustainability											
P(all AAV > 15%)	0.993	0.988	0.958	0.927	0.905	0.847	0.813	0.771	0.722	0.689	0.606
P(all TM < 34 Mlbs)	0.465	0.458	0.457	0.439	0.425	0.432	0.426	0.436	0.448	0.455	0.507
P(any TM < 34 Mlbs)	0.891	0.862	0.81	0.758	0.718	0.681	0.661	0.641	0.633	0.627	0.662
Median average TM	42.06	41.84	39.64	40.6	41.12	39.57	39.82	38.48	37.97	37.39	33.95
P(all decrease TM > 15%)	0.365	0.352	0.336	0.319	0.302	0.285	0.273	0.261	0.244	0.236	0.221
P(any decrease TM > 15%)	0.997	0.992	0.992	0.982	0.974	0.967	0.958	0.946	0.94	0.932	0.921
median AAV TM	41.8%	37.3%	33.1%	30.2%	26.8%	23.9%	21.1%	19.4%	18.4%	17.5%	16.3%
Rankings (lower is better)											
P(<20%)[1]	11	10	9	8	7	6	5	4	3	2	1
P(AAV > 15%)[2]	11	10	9	8	7	6	5	4	3	2	1
Maximum catch (TM)[3]	1	2	3	4	5	6	7	8	9	10	11

[1] This ranking is determined using P(any dRSB < 20%) and the objective to maintain RSB above 20% at least 90% of the time. Note that all procedures meet this objective.

[2] This ranking is determined using P(aall AAV >15%) and the objective to maintain AAV below 15%.at least 75% of the time. Note that no procedures meet this objective.

[3] This ranking is determined using a smoothed relationship for Median average TM to account for variability in the simulations. Note that the highest fishing intensity meets this objective.

Table A2. Long-term performance metrics for an estimation error CV of 0.15, autocorrelation of 0.4, a 40:20 control rule, and a range of input SPRs.

Input Est Error	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**
Input Autocorrelation	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**
Input Control Rule	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**
Input SPR	**30%**	**32%**	**34%**	**36%**	**38%**	**40%**	**42%**	**44%**	**46%**	**48%**	**56%**
Median SPR	47.7%	47.9%	47.9%	48.1%	48.3%	48.6%	49.1%	49.6%	50.4%	51.3%	55.4%
Biological Sustainability											
Median average dRSB	35.8%	36.4%	37.1%	37.8%	38.6%	39.5%	40.4%	41.5%	42.6%	43.9%	47.2%
P(all dRSB<20%)	0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001
P(any dRSB_y<20%)	0.003	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001
P(all dRSB<30%)	0.083	0.059	0.044	0.028	0.018	0.014	0.007	0.007	0.008	0.006	0.008
P(any dRSB_y<30%)	0.309	0.214	0.16	0.102	0.052	0.036	0.022	0.015	0.011	0.007	0.011
Fishery Sustainability											
P(all AAV > 15%)	0.998	0.996	0.994	0.994	0.986	0.985	0.974	0.948	0.921	0.88	0.788
P(all TM < 34 Mlbs)	0.495	0.488	0.479	0.476	0.470	0.468	0.465	0.463	0.460	0.459	0.483
P(any TM < 34 Mlbs)	0.889	0.869	0.856	0.836	0.819	0.801	0.778	0.756	0.735	0.711	0.693
Median average TM	39.71	39.6	39.97	39.59	39.19	38.79	38	37.73	37.6	37.27	35.56
P(all decrease TM > 15%)	0.390	0.386	0.381	0.372	0.362	0.349	0.337	0.326	0.310	0.289	0.275
P(any decrease TM > 15%)	0.999	0.998	0.998	0.998	0.998	0.997	0.996	0.994	0.981	0.973	0.953
median AAV TM	46.2%	43.6%	41.9%	39.3%	36.0%	33.5%	30.9%	28.2%	25.9%	23.2%	21.1%
Rankings (lower is better)											
P(<20%)[1]	11	10	9	8	7	6	5	4	3	2	1
P(AAV > 15%)[2]	11	10	9	8	7	6	5	4	3	2	1
Maximum catch (TM)[3]	1	2	3	4	5	6	7	8	9	10	11

[1] This ranking is determined using P(any dRSB < 20%) and the objective to maintain RSB above 20% at least 90% of the time. Note that all procedures meet this objective.

[2] This ranking is determined using P(aall AAV >15%) and the objective to maintain AAV below 15%.at least 75% of the time. Note that no procedures meet this objective.

[3] This ranking is determined using a smoothed relationship for Median average TM to account for variability in the simulations. Note that the highest fishing intensity meets this objective, although the yield curve appears flat at those low SPR values.

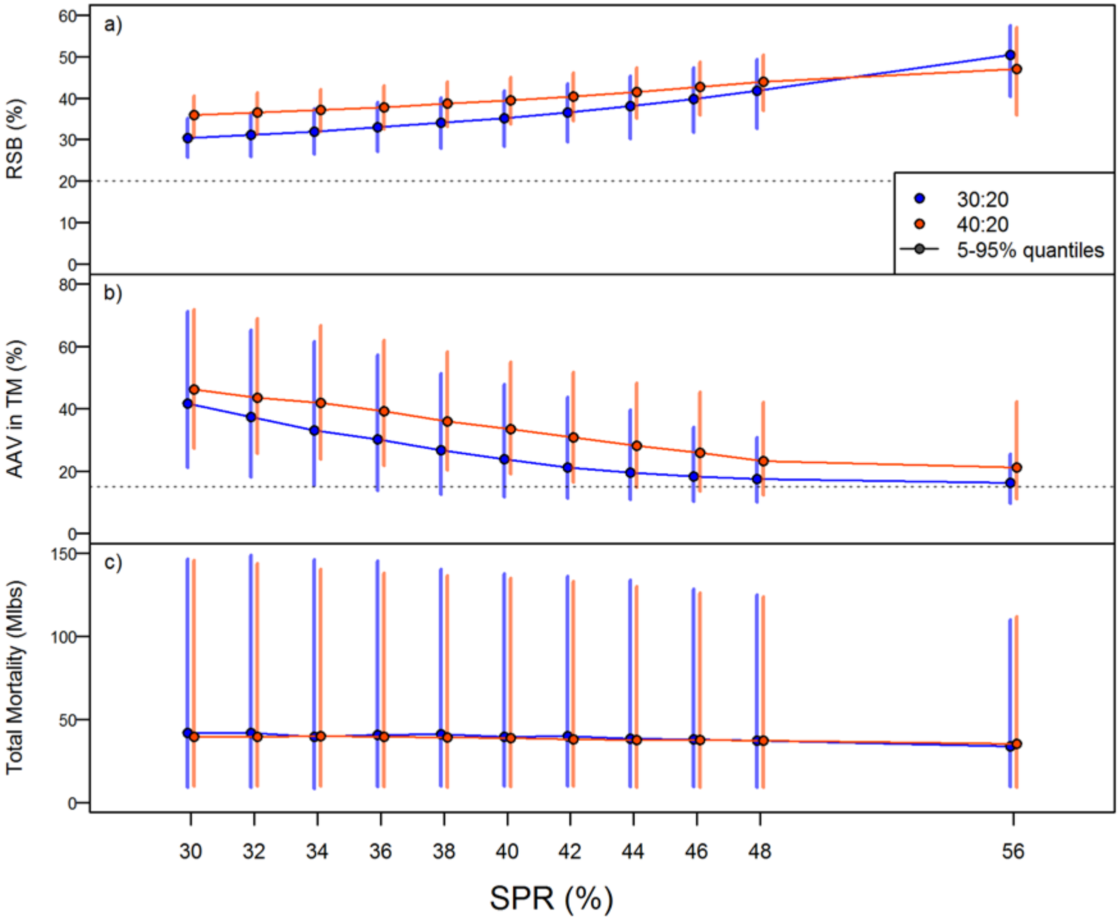


Figure A1. Long-term mdedian relative spawning biomass (a), median AAV for Total Mortality (b), and median Total Mortality (Mlbs, c) shown as points, with 90% confidence intervals shown as vertical lines for various SPRs and two control rules (30:20 and 40:20). The estimation error CV is 0.15 and autocorrelation is 0.4.

Table A3. Short-term (3-8 annual time-steps) performance metrics for an estimation error CV of 0.15, autocorrelation of 0.4, a 30:20 control rule, and a range of input SPRs.

Input Est Error	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**
Input Autocorrelation	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**
Input Control Rule	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**	**30:20**
Input SPR	**30%**	**32%**	**34%**	**36%**	**38%**	**40%**	**42%**	**44%**	**46%**	**48%**	**56%**
Median SPR	39.8%	39.6%	39.9%	40.6%	41.7%	43.2%	44.6%	46.4%	48.1%	49.9%	57.2%
Biological Sustainability											
Median average dRSB	31.1%	32.1%	33.0%	33.8%	34.6%	35.4%	36.2%	37.0%	37.8%	38.6%	41.4%
P(all dRSB<20%)	0.074	0.074	0.074	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
P(any dRSB_y<20%)	0.11	0.108	0.108	0.107	0.106	0.107	0.107	0.107	0.107	0.106	0.106
P(all dRSB<30%)	0.459	0.395	0.347	0.316	0.287	0.265	0.248	0.232	0.221	0.213	0.183
P(any dRSB_y<30%)	0.688	0.568	0.485	0.428	0.377	0.342	0.309	0.286	0.273	0.263	0.234
Fishery Sustainability											
P(all AAV > 15%)	0.893	0.866	0.832	0.81	0.796	0.78	0.751	0.734	0.722	0.713	0.677
P(all TM < 34 Mlbs)	0.377	0.356	0.353	0.354	0.360	0.378	0.399	0.426	0.460	0.494	0.683
P(any TM < 34 Mlbs)	0.732	0.664	0.629	0.594	0.594	0.59	0.608	0.637	0.67	0.708	0.902
Median average TM	46.81	46.4	45.62	44.26	42.89	41.53	40.02	38.61	37.25	35.61	29.41
P(all decrease TM > 15%)	0.398	0.382	0.359	0.341	0.324	0.312	0.300	0.291	0.284	0.273	0.249
P(any decrease TM > 15%)	0.943	0.937	0.919	0.91	0.893	0.885	0.876	0.867	0.856	0.842	0.811
median AAV TM	35.9%	32.1%	27.6%	25.6%	23.4%	22.6%	21.6%	20.9%	20.3%	20.0%	18.7%
Rankings (lower is better)											
P(<20%)[1]	11	9.5	9.5	6	2	6	6	6	6	2	2
P(AAV > 15%)[2]	11	10	9	8	7	6	5	4	3	2	1
Maximum catch (TM)[3]	1	2	3	4	5	6	7	8	9	10	11

[1] This ranking is determined using P(any dRSB < 20%) and the objective to maintain RSB above 20% at least 90% of the time. Note that no procedure meets this objective.

[2] This ranking is determined using P(aall AAV >15%) and the objective to maintain AAV below 15%.at least 75% of the time. Note that no procedures meet this objective.

[3] This ranking is determined using a smoothed relationship for Median average TM to account for variability in the simulations. Note that the highest fishing intensity meets this objective.

Table A4. Short-term (3-8 annual time-steps) performance metrics for an estimation error CV of 0.15, autocorrelation of 0.4, a 40:20 control rule, and a range of input SPRs.

Input Est Error	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**	**0.15**
Input Autocorrelation	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**	**0.4**
Input Control Rule	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**	**40:20**
Input SPR	**30%**	**32%**	**34%**	**36%**	**38%**	**40%**	**42%**	**44%**	**46%**	**48%**	**56%**
Median SPR	50.0%	49.7%	49.9%	50.3%	50.8%	51.5%	52.2%	53.1%	54.1%	55.3%	58.9%
Biological Sustainability											
Median average dRSB	34.5%	35.2%	35.8%	36.5%	37.1%	37.7%	38.2%	38.9%	39.5%	40.1%	41.5%
P(all dRSB<20%)	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
P(any dRSB_y<20%)	0.105	0.105	0.104	0.105	0.106	0.105	0.105	0.105	0.105	0.105	0.105
P(all dRSB<30%)	0.238	0.218	0.204	0.195	0.186	0.179	0.177	0.172	0.170	0.168	0.169
P(any dRSB_y<30%)	0.392	0.334	0.301	0.281	0.261	0.246	0.242	0.235	0.229	0.227	0.228
Fishery Sustainability											
P(all AAV > 15%)	0.963	0.956	0.957	0.945	0.94	0.923	0.904	0.89	0.872	0.859	0.819
P(all TM < 34 Mlbs)	0.550	0.539	0.530	0.523	0.518	0.517	0.523	0.532	0.544	0.560	0.646
P(any TM < 34 Mlbs)	0.941	0.93	0.905	0.885	0.868	0.852	0.846	0.845	0.848	0.848	0.912
Median average TM	38.16	37.42	37.13	37.03	36.3	35.39	34.54	33.66	32.77	31.82	28.86
P(all decrease TM > 15%)	0.413	0.404	0.393	0.382	0.367	0.356	0.343	0.328	0.316	0.309	0.282
P(any decrease TM > 15%)	0.94	0.941	0.936	0.932	0.926	0.922	0.92	0.912	0.904	0.89	0.857
median AAV TM	47.1%	43.9%	40.9%	38.9%	36.3%	34.2%	31.7%	29.9%	28.3%	27.1%	24.6%
Rankings (lower is better)											
P(<20%)[1]	6	6	1	6	11	6	6	6	6	6	6
P(AAV > 15%)[2]	11	9	10	8	7	6	5	4	3	2	1
Maximum catch (TM)[3]	1	2	3	4	5	6	7	8	9	10	11

[1] This ranking is determined using P(any dRSB < 20%) and the objective to maintain RSB above 20% at least 90% of the time. Note that no procedure meets this objective.

[2] This ranking is determined using P(aall AAV >15%) and the objective to maintain AAV below 15%.at least 75% of the time. Note that no procedures meet this objective.

[3] This ranking is determined using a smoothed relationship for Median average TM to account for variability in the simulations. Note that the highest fishing intensity meets this objective.