IPHC Management Strategy Evaluation Update

PREPARED BY: IPHC SECRETARIAT (A. HICKS; 18 OCTOBER 2017)

PURPOSE

To provide an update on the progress of the IPHC Management Strategy Evaluation process and seek recommendations from the MSAB related to the Management Strategy Evaluation simulation framework. Also, to present results from the simulations for evaluation by the MSAB.

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 2014–2016 from the stock assessment. 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 spawning biomass.

The stock assessment predicted a 68% chance that the spawning biomass will decline in 2017 and a 6% chance that it will decline more than 5% with the status quo SPR fishing intensity (Table 4 in Stewart and Hicks (2017)). The greater than 50% chance of decline, although a slight decline, indicates that the status quo SPR may not determine a fishing intensity that will meet the long-term goals and objectives defined by the MSAB. Therefore, an evaluation of fishing intensities, through simulation, should be done. A brief description of the framework and components of these simulations is given below, followed by details of the framework that were not presented at MSAB09 in May 2017. Details of the framework that were presented at SRB10 in June 2017 are provided in Appendix A.

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. **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.
3. The **Estimation Model (EM)** is analogous to the stock assessment. 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.
4. **Management Procedure** 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.

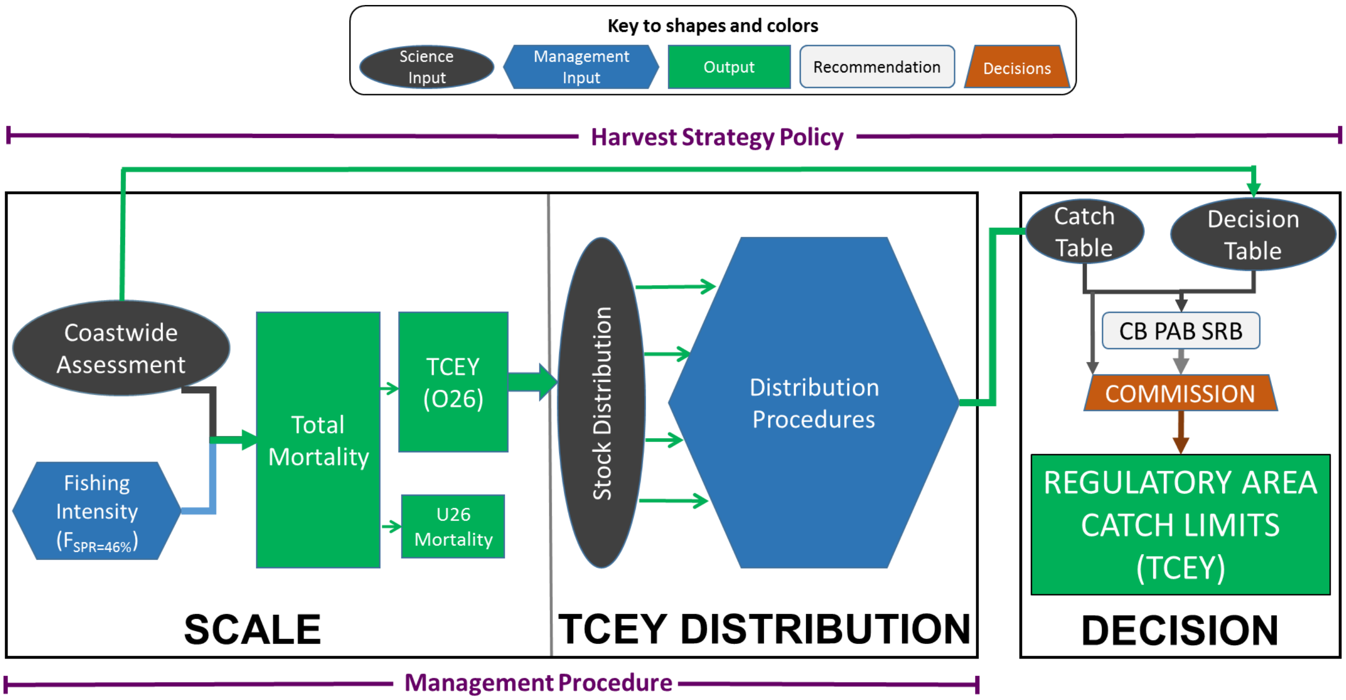

Figure 1: A pictorial description of the interim IPHC harvest strategy policy showing the separation of scale and distribution of fishing mortality. The "decision step" is when policy and decision making (not a procedure) influences the final mortality limits.

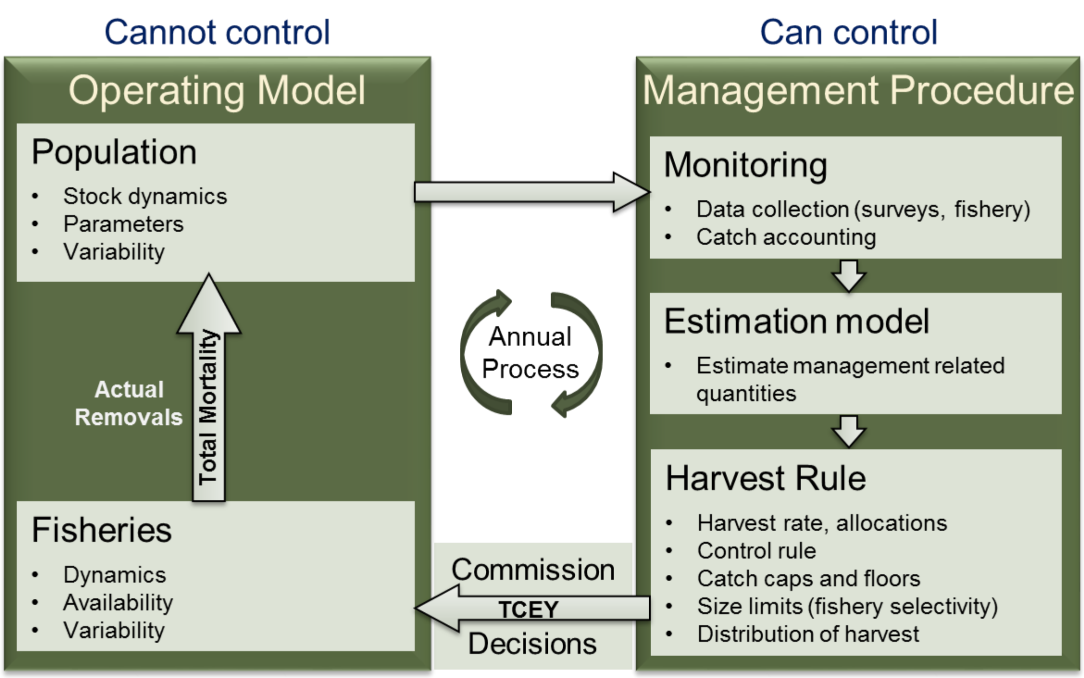

Figure 2: Diagram of the relationship between the four modules in the framework. The simulations run each module on an annual time-step, producing output that is used in the next time-step. See text for a description of operating model, monitoring, estimation model, and harvest rule.

Operating Model

For the simulations to investigate a coastwide fishing intensity, the Stock Synthesis (Methot and Wetzel 2013) assessment program was used as an operating model. This platform is currently used for the stock assessment, and the operating model was composed 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 a fleets-as-areas 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 use five fisheries: commercial, discards (wastage), bycatch, sport, and personal use, and the TCEY was distributed to each fleet in an ad hoc manner (see Scenarios). The survey is also included as a fleet without catch.

Monitoring (Data Generation)

A simplified estimation model was used due to time constraints; thus no data were generated.

Estimation Model

Of the four options presented in Appendix A, the Perfect Information option was used in these simulations. Perfect Information assumes that the population values needed to apply the management procedure are exactly known (e.g., spawning biomass) and is useful as a reference to the performance without uncertainty in an estimation model. The other options will be considered for future simulations, but due to time-constraints, were not used here.

Management Procedure

The 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 AM094 in 2018. In addition to F_{SPR}, a control rule was used to adjust the fishing intensity at low stock status. This is discussed below in the Management Procedures section. For these simulations, I used the two coastwide models, and only needed to distribute the catch to the five coastwide fleets.

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) Uncertainty incorporated through parameter uncertainty and model uncertainty. See Scenarios.

2) Management Procedure
    a) A coastwide fishing intensity (F_{SPR})
    b) A control rule
    c) Catch assigned to sectors based on historical information (with variability)
3) Estimation Models
    a) Perfect Information (as a reference if we knew population values exactly).
    b) Simulate error from the simulated time-series to mimic a stock assessment.
4) Data Generation
    a) Not needed at this time.

SCENARIOS

Scenarios are alternative states of nature in the operating model, which are represented by parameter and model uncertainty, as described in Appendix A. These alternative states of nature integrate over the uncertainty in the system that we cannot, or choose not to, control. The scenarios for the MSE simulations include uncertainty in the operating model processes as described in Table 1.

Table 1: Processes and associated uncertainty in the operating model (OM) to potentially include as scenarios in the simulations. 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

ALLOCATING SIMULATED TOTAL MORTALITY TO SECTORS

The simulated management strategy returns a coastwide recommended TCEY, which is then allocated to each of the five sectors, with variability. In reality, there is a slight difference between the Total Mortality (TM) and the TCEY because of shortfalls and overages, but those should be dealt with on a sector basis. The MSAB09 meeting in May 2017 noted that catch history, in conjunction with uncertainties and sensitivities, can be used to allocate TM to each sector. Recent sector-specific mortality or proportions of TM for each sector were used to guide the allocation using relationships between the sector specific mortality or proportions to the TM. For example, at low TM the bycatch is likely a larger proportion. Figure 3 shows the percentage of TM attributed for each sector for the past 40 years.

The mortality for the personal use sector has been between 1.1 Mlbs and 1.5 Mlbs for the last ten years, and a consistent 1.20 Mlbs for the last three years. Therefore, simply randomly drawing a mortality for the personal use sector from a lognormal distribution with a median of 1.2 Mlbs and a CV of 15% would provide 5th and 95th percent quantiles of approximately 0.9 and 1.5 Mlbs. This randomly generated poundage is the mortality for the personal use sector. However, it is set to a minimum of 0.5 Mlbs if generated to be below that, and if generated to be more than one-half the total mortality, then is generated from a lognormal distribution with a median equal to one-half the total mortality. Figure 4 shows the lognormal distribution along with observed personal use mortality since 1998.

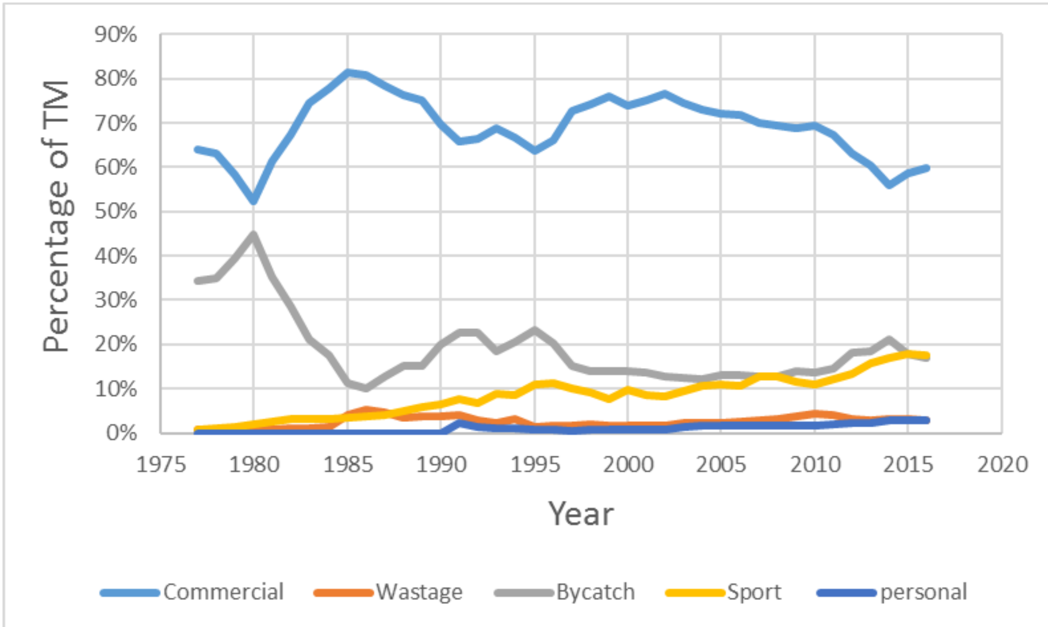

Figure 3: Percentage of Total Mortality (TM) for each sector used in the assessment model from 1976 to 2016.

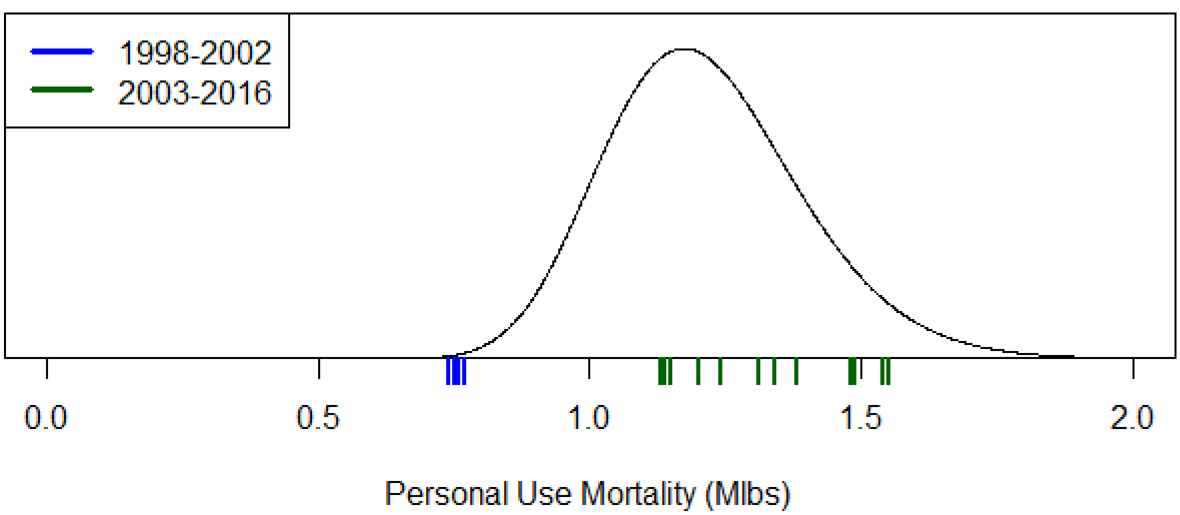

Figure 4: The lognormal distribution used to randomly generate personal use. Shown as blue and green tick lines at the bottom are the observed personal use mortality (Mlbs) from 1998-2002 (blue) and 2003-2016 (green).

Bycatch mortality is a non-directed fishery component of the TM, and is subtracted from the TCEY (along with some other components depending on the Regulatory Area) leaving the remaining amount as the FCEY. It is often managed as a static amount (e.g., fixed PSC limits), but the bycatch mortality has typically been less than the caps and has been declining over the last decade. For simplicity, the bycatch is randomly drawn from a lognormal distribution with a median of 7 Mlbs and a CV of 20%. Figure 5 shows the lognormal distribution. The 5th and 95th percentiles are approximately 5 Mlbs and 9.7Mlbs, respectively. This will provide a reasonable bycatch mortality that when integrated into the simulations will provide for evaluations that are robust to a reasonable range of bycatch mortality amounts.

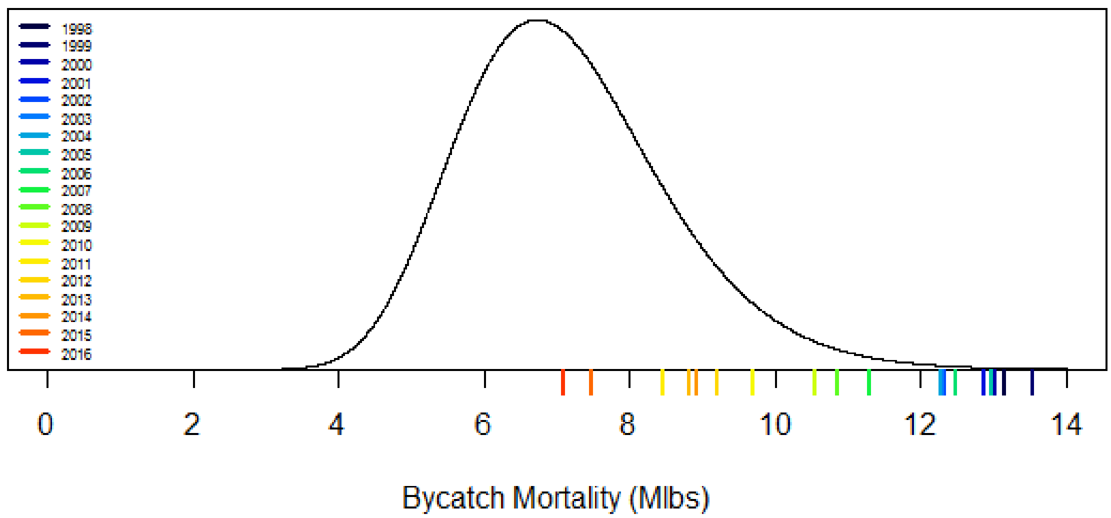

Figure 5: The lognormal distribution from which bycatch mortality is randomly drawn along with observed bycatch mortality since 1998. The colors represent years of the observations, starting with dark blue for 1998 moving to red in 2016.

Sport fishery mortality was related to total mortality, but it appears that the percentage of sport mortality was consistently around 11% or distributed around 7.6 Mlbs during the early 2000's when the TM was above 50 Mlbs (Figure 6). However, since 2011, and when the TM has been less than 50 Mlbs, the percentage of sport mortality has been larger than 15%, but consistently near 7 Mlbs. There were significant changes to the catch sharing plans in the last few years, which could help explain this increase in the percentage of sport mortality. Therefore, the MSAB suggested (at MSAB09) to look at more recent years to determine how to simulate sport mortality.

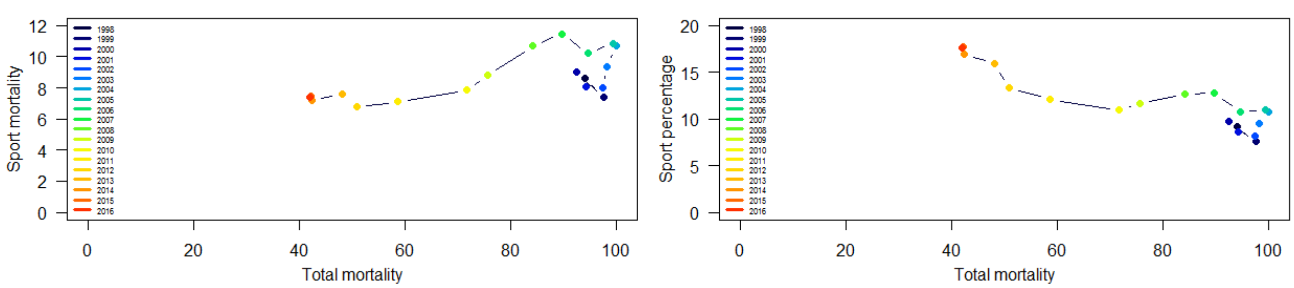

Figure 6: Sport mortality (left) and percentage of sport mortality (right) plotted against Total Mortality (TM) for the years 1998-2016. Colors indicate the year of that observation.

It is difficult to characterize what sport mortality may be at higher total mortality when recent years occurred during a time when total mortality was less than 50 Mlbs. Therefore, the sport mortality was treated differently above and below 57 Mlbs. Sport mortality is generated from a lognormal distribution with a median of 7.682 Mlbs and a CV of 20% when the total mortality is greater than 57 Mlbs. At values of total mortality less than 57 Mlbs, the percentage of sport mortality (as a function of total mortality) is a linear relationship with y-intercept 0.2698 and slope -2.369×10^{-06}. The slope and intercept parameters had a slight amount of variability determined from the predicted uncertainty from the linear regression.

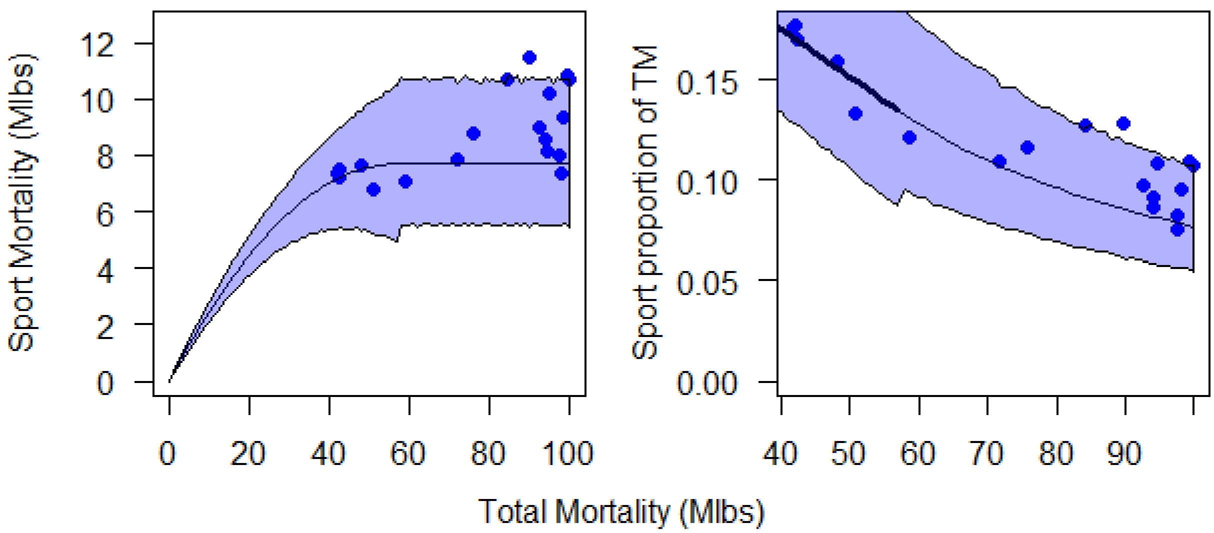

Figure 7: Simulated and observed sport mortality (left) and the sport proportion of the total mortality (right) with the area between the 5th and 95th quantiles shown in light blue.

Once personal use, bycatch, and sport mortalities are determined, the commercial plus discard (wastage) mortality makes up the remaining amount. The discard component represents the mortality of halibut caught in the directed fishery but not landed (i.e., under-sized fish, lost gear, etc.) When more fishing occurs (higher commercial+discards mortality) then the discards may increase, and when the fish are smaller the discards should also increase. Therefore, the discards were modelled as a function of the commercial+discards mortality and the size at age for an age 8 male halibut. When the commercial+discards mortality goes up, the wastage also increases, and when age 8 males are small, the wastage increases. Figure 8 shows the discard proportion of the TM plotted against commercial+discards and also against male weight at age 8. It shows that there is an interplay between these two variables and they are both important to how much discard mortality there is.

Using commercial+discards and male weight at age 8 as explanatory variables, four different models were tested to model the discard proportion of commercial+discards mortality.

$$p_d = \begin{array}{ll} \alpha_1 + \beta_{\{1,1\}}\mathrm{CommDis} & (1) \\ \alpha_2 + \beta_{\{1,2\}}\mathrm{maleWt8} & (2) \\ \alpha_3 + \beta_{\{1,3\}}\mathrm{CommDis} \quad + \beta_{\{2,3\}}\mathrm{maleWt8} & (3) \\ \alpha_4 + \beta_{\{1,4\}}\log(\mathrm{CommDis}) + \beta_{\{2,4\}}\mathrm{maleWt8} & (4) \end{array}$$

The residuals for these four models are plotted in Figure 9 along with the sum of squared residuals, which was lowest for model (4). Including the log of commercial+discards and the male weight at age 8 improved the pattern in residuals and was better at explaining the 2014–2016 observations. None of the above models explained 2012 and 2013 very well. Model (4) was used to determine the discard proportion of the commercial+discards (Figure 10), and overall, the proportion of discards is small.

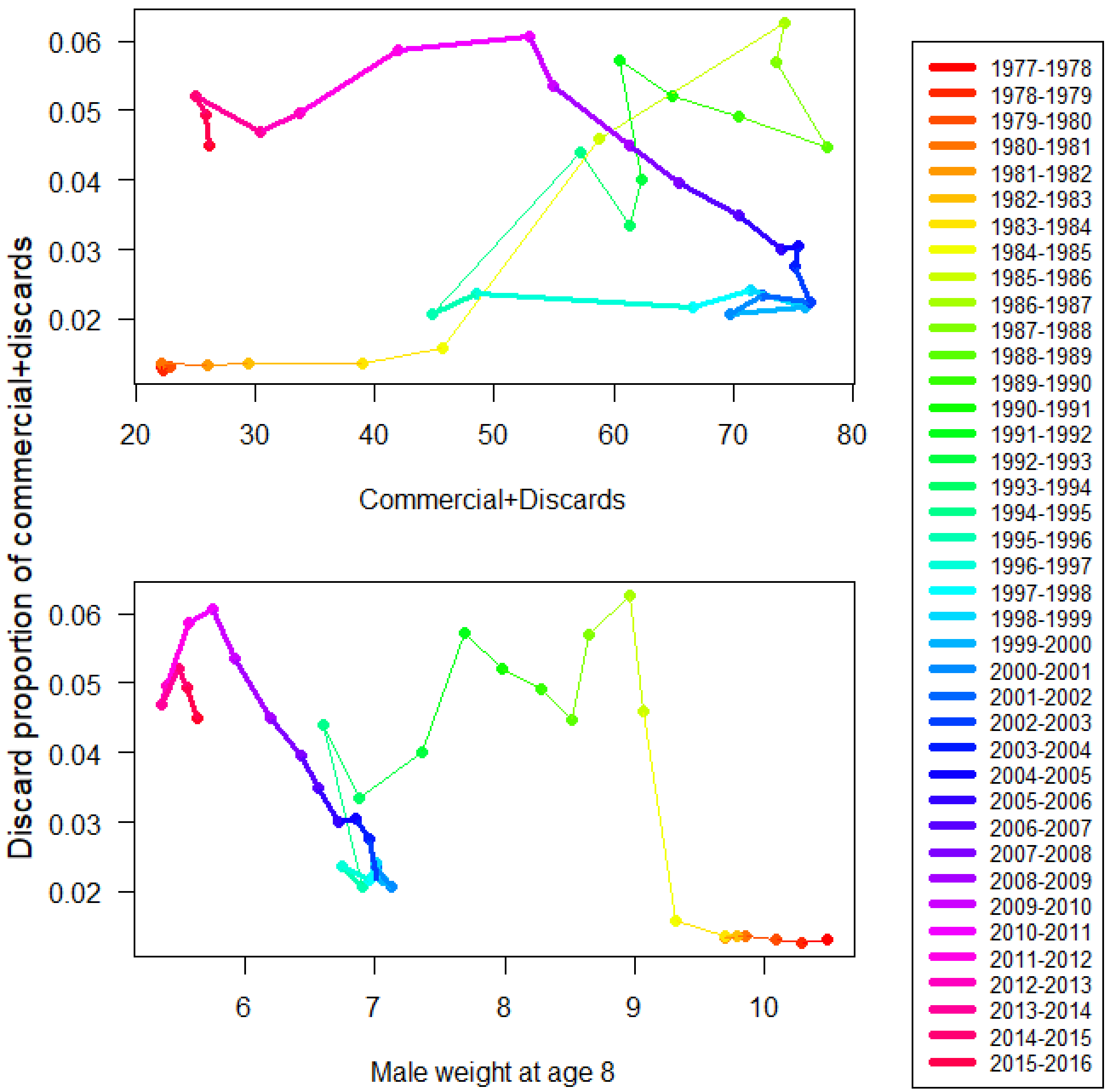

Figure 8: The discard proportion of the total mortality plotted against the commercial+discard mortality (top) and the male weight at age 8 (bottom). The colors indicate the year, and thicker lines are plotted for 1995 and later.

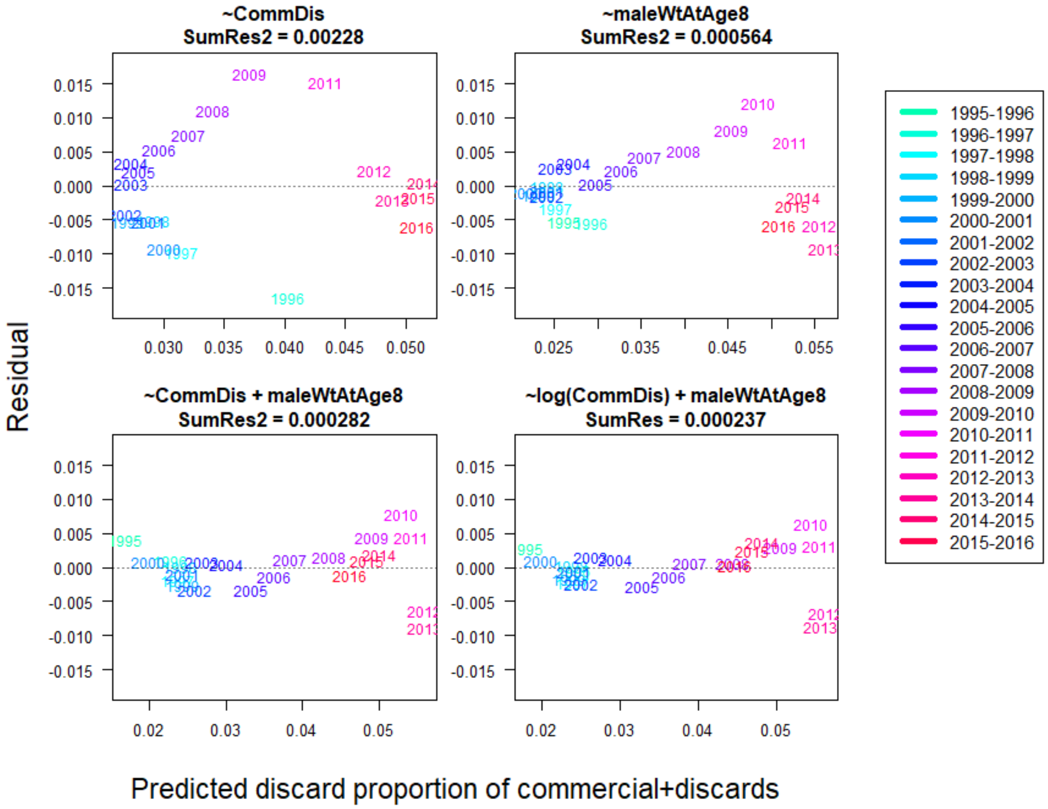

Figure 9: Residuals from four different models (labeled in titles) plotted against the predicted discard proportion of commercial+discards, colored by year. The sum of squared residuals (SumRes2) is also shown in the title. The model p_d=log(CommDis)+maleWtAtAge8 (lower left) was used to determine the discards.

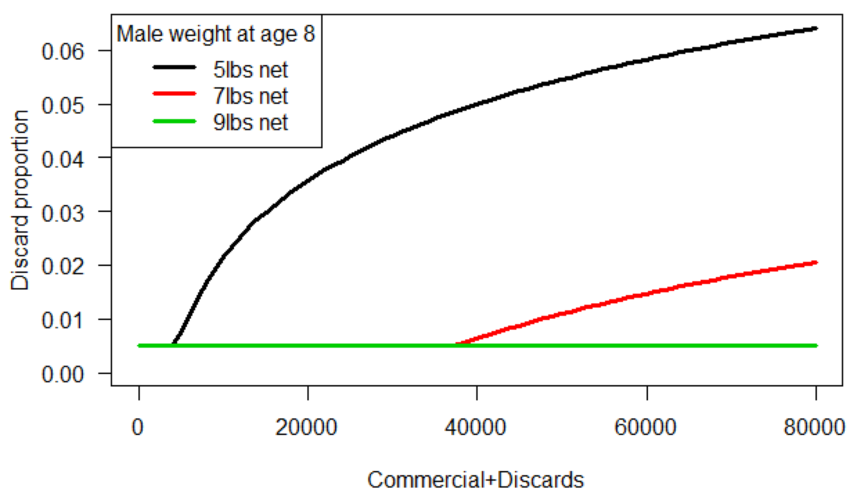

Figure 10: The discard proportion used to allocate discards as a function of commercial+discards and three different values of male weight at age 8.

The commercial mortality is the remainder of the total mortality after subtracting the subsistence, bycatch, sport, and discard components. Due to minimum levels of mortality and random variability, it is possible that, at low levels of total mortality, there is no commercial mortality and that the actual total mortality exceeds the mortality determined from the management procedure. Expected values of the mortality and proportion by sector plotted against Total Mortality is shown in Figure 11.

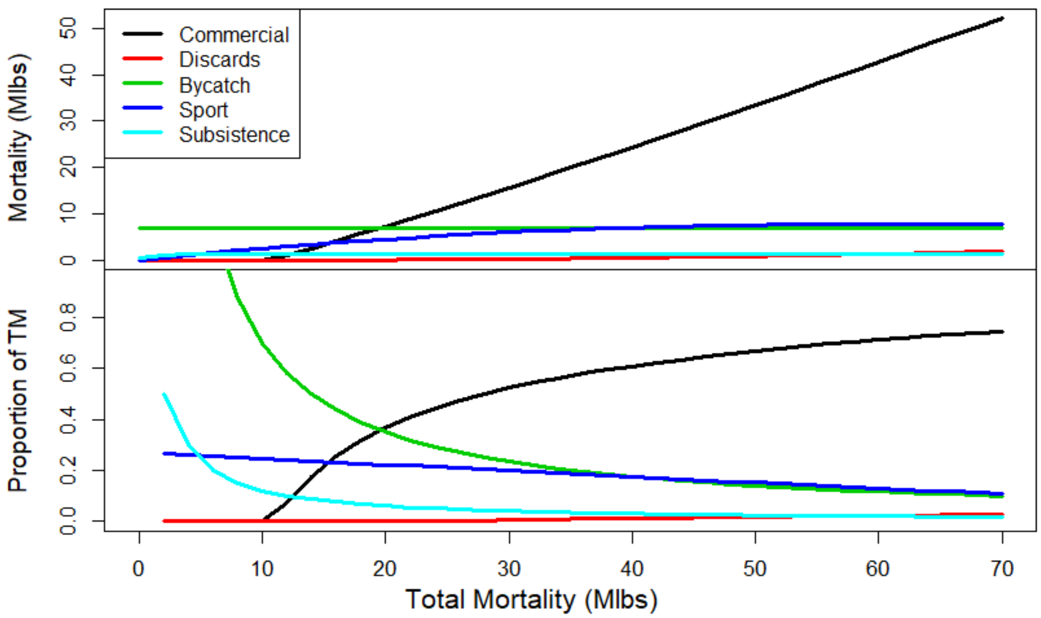

Figure 11: Average sector specific mortality (top, Mlbs) and the sector-specific proportion of Total Mortality (TM) plotted against TM. Age 8 males are 6 lbs and random variability is not present.

SIMULATING WEIGHT-AT-AGE

It is important to simulate time-varying weight-at-age because it is an influential trait on the yield and status of Pacific halibut. There are 83 years of weight-at-age observations in the long time-series assessment models, with an observed wide range over the years (Figure 12 and **Figure 13**). Many years of these data have been estimated from sparse data, and the entire time-series has been smoothed to eliminate large deviations from year to year.

Many methods were trialed to simulate future weight-at-age with the goal of mimicking the general behavior of the historical weight-at-age. Important behaviors of the historical time-series are

1. the age-specific weights-at-ages tend to increase and decrease in the same year (little evidence of lags for a cohort),
2. the time-series appears to be similar to a random walk with smooth trends and few large jumps in observations (partly due to the smoothing that was done), and
3. there appears to be some ages that do not follow the general trend (evident at the end of the time series where the sampling was likely greater).

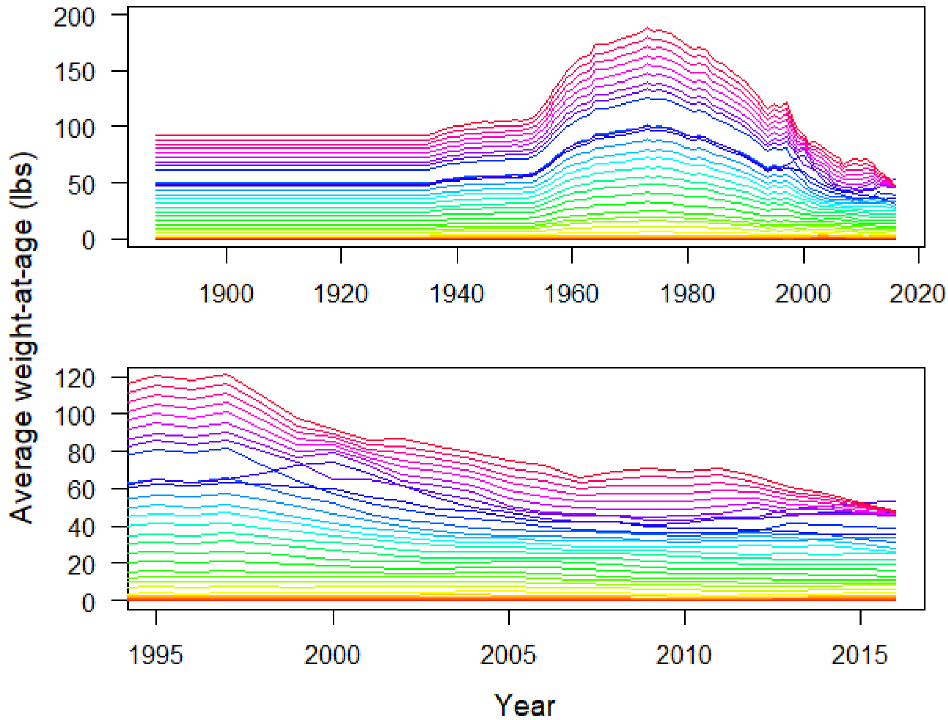

Figure 12: Historical weight-at-age as used in the long time-series assessment models. Note that the observations are smoothed over years to reduce spurious observations.

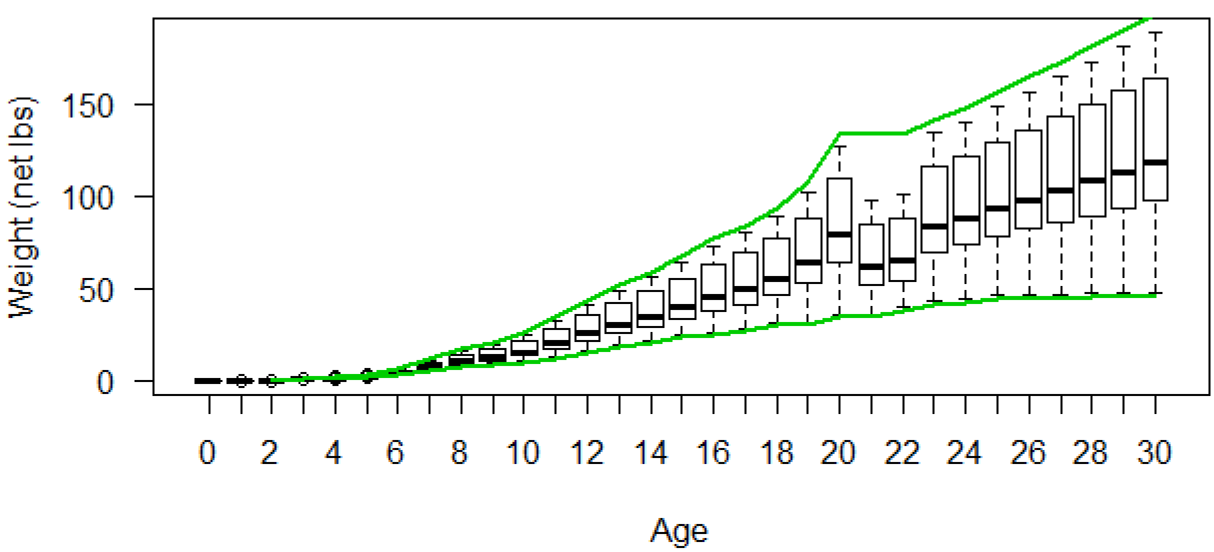

Figure 13: Boxplots of weight at ages 0 to 30 over all historical years. The green line shows the lower and upper bounds used in the simulations.

The method used to simulate weight-at-age addressed each of these behaviors in the following ways.

1. A single deviation was generated from a normal distribution with a constant standard deviation (0.05), and was a multiplier on the current years weight-at-age to determine the weight-at-age in the next year. This made all weights for each age increase or decrease similarly.
2. A random walk was used where the weight-at-age in the next year was generated from the weight-at-age in the current year. The deviation in (1) was also correlated with past deviations to simulate longer periods of similar trends (ρ=0.5).
3. Deviations for each age 6 and greater were generated from a normal distribution with a constant coefficient of variation for each age (0.01), resulting in standard deviations scaled by the mean weight-at-age observed over all historical years with observations (Figure 14). This allows for larger deviations for older fish and provides a mechanism for the mean weight of a specific age to depart from the overall trend (simulated in step 1).

$$W_{a,t} = W_{a,t-1} e^{\varepsilon_{t,1} - 0.5 * \sigma_1^2} + \varepsilon_{a,t,2}$$

$$\begin{array}{lll} \varepsilon_{t,1} & \sim & \rho \varepsilon_{t,1} + \sqrt{1-\rho^2} \times N(\mu = 0, sd = 0.1) \\ \varepsilon_{a,t,2} & \sim & N(\mu = 0, cv = 0.02) \qquad a \geq 6 \end{array}$$

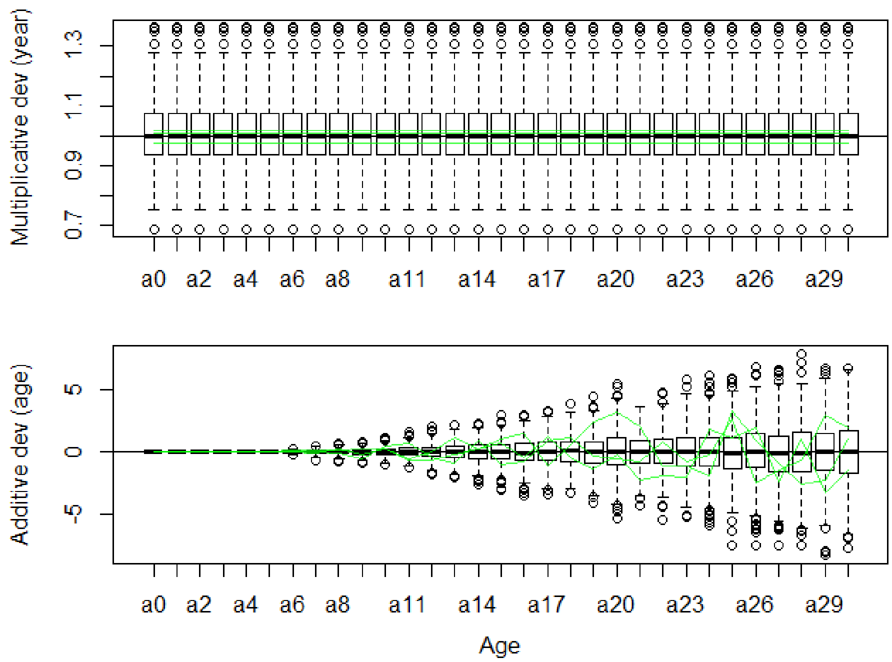

Figure 14: Randomly drawn deviations for 1000 simulations. The multiplicative deviations are $e^{\varepsilon_{t,1} - 0.5 * \sigma_1^2}$ and the additive deviations are $\varepsilon_{a,t,2}$. Green lines show deviations for individual simulations.

The random walk could potentially walk to extremely high values or low values (obviously negative weight-at-age is not an option). Therefore, boundary conditions were set to limit the range over which weight-at-age could vary. The boundary limits were determined from the observed range of weight at each age, and expanded 5% beyond the minimum and maximum weight at each age observed. Two upper boundaries (ages 21 and 22) were expanded further to equal the upper boundary of age 20 (**Figure 13**). The random walk simulations remained within the bounds by applying the following algorithm.

1. If a weight-at-age was simulated to be beyond the bounds, $\varepsilon_{t,1}$ and $\varepsilon_{a,t,2}$ for only the ages where the age-specific bounds were exceeded were reduced by one-half and applied again to determine if it still exceeded the bounds.
2. Repeat step (1) until no age-specific bounds were exceeded.

The historical weight-at-age was smoothed across years (Figure 12) and the simulation procedure above produced annual weight-at-age that was less smooth (Figure 15). It is possible to smooth the entire time-series so that the 'spikiness' of the simulated weight-at-age is reduced, but the entire time-series would need to be smoothed or else a disconnect would occur between the non-smoothed and smoothed series (e.g., historical and simulated). Changing the historical time-series would modify the conditioning of the OM, but smoothing the simulated time-series would not have much effect on the simulation results since it is the large-scale changes that are likely most influential. Therefore, the above algorithm was used without additional smoothing.

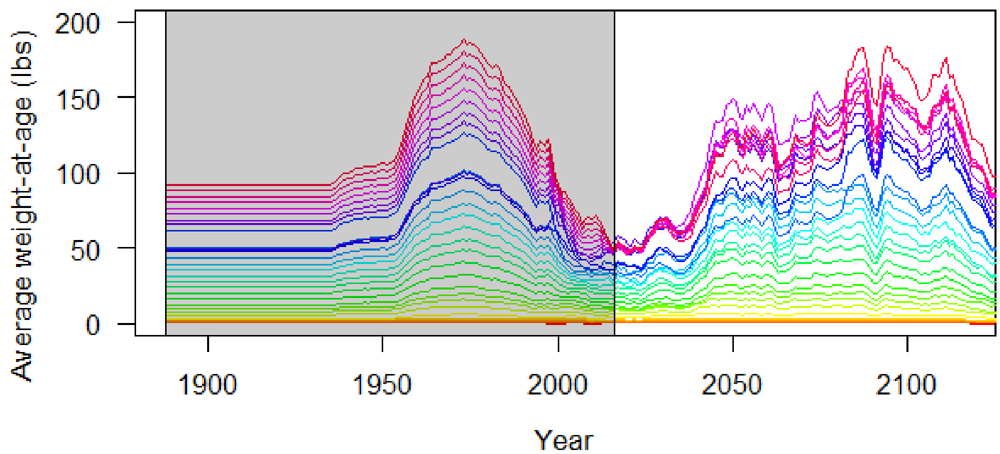

Figure 15: One potential simulated female weight at age in the historical period (1888-2016, shaded) and the simulated period (2017-2116).

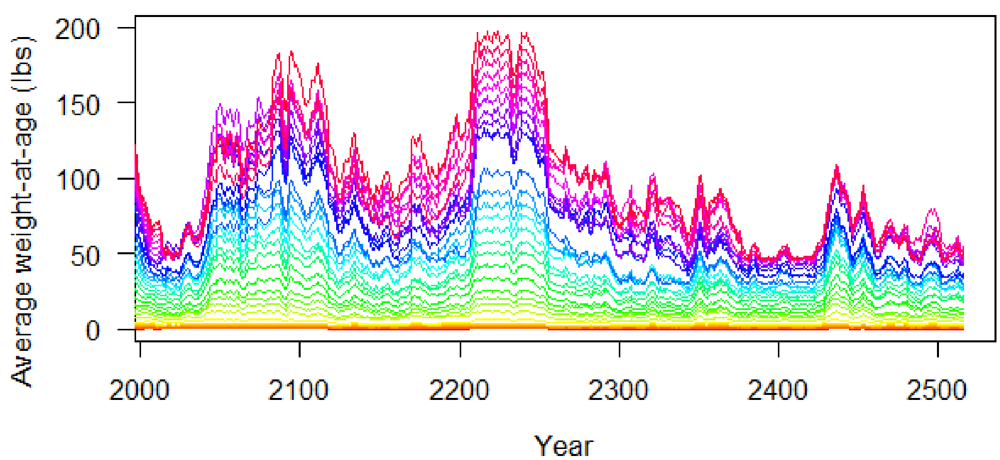

Figure 16: One potential simulation of female weight-at-age for 500 years.

SIMULATING REGIME SHIFTS

An environmental regime is used in the stock assessment to determine if average recruitment is high or low. This is based on the pacific Decadal Oscillation (PDO, http://research.jisao.washington.edu/pdo/, Mantua et al. 1997, **Figure 17**) and the value is 0 or 1 depending on classified cool or warm years, respectively (Figure 18).

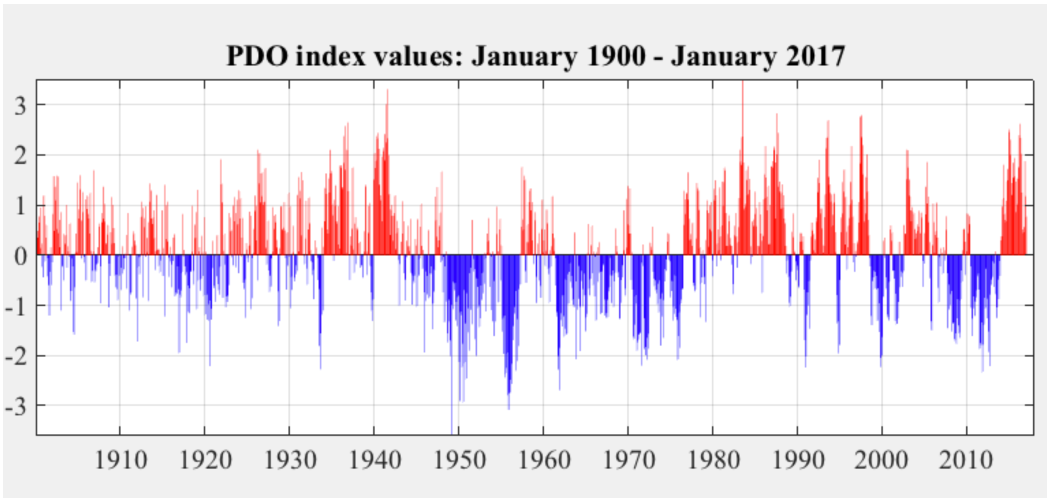

Figure 17: Pacific Decadal Oscillation (PDO) (figure from http://research.jisao.washington.edu/pdo/).

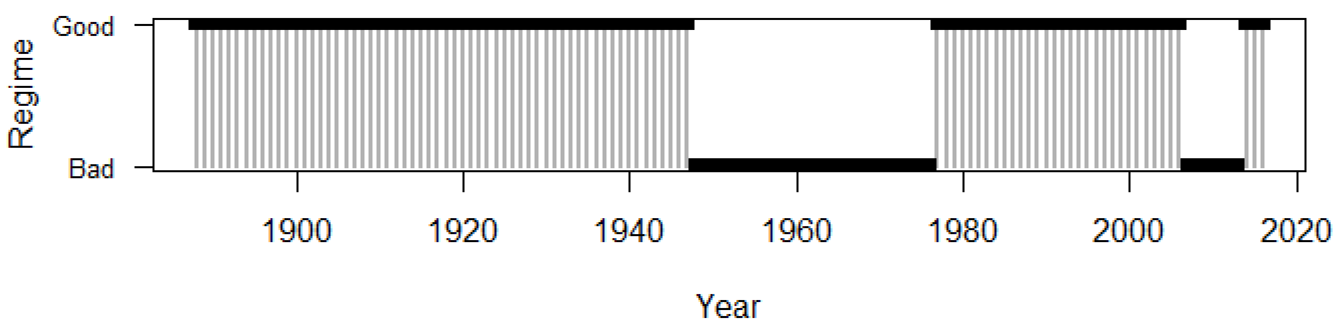

Figure 18: Good and bad regimes in the Pacific halibut stock assessment for 1888-2016.

The regime was simulated in the MSE simulations by generating a 0 or 1 to indicate the regime in that future year. To encourage runs of a regime between 15 and 30 years (an assumption of the common periodicity, although recent years have suggested less), the environmental index was simulated as a semi-Markov process, where the next year depends on the current year. However, the probability of changing to the opposite regime was a function of the length of the current regime (Figure 19) with a probability of 0.5 at 30 years.

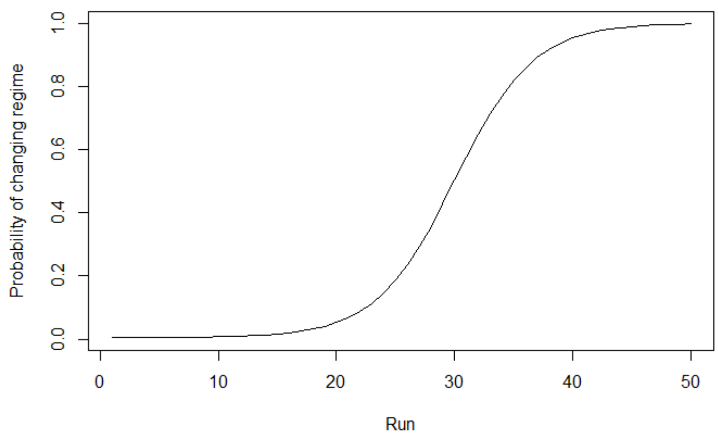

Figure 19: The probability of the regime changing given the length of the current regime (run).

An example simulation for 500 years is shown in Figure 20 and the histogram of run lengths is most often between 20 and 30 years, with occasional runs between 1 and 20 years. No run was longer than 35 years.

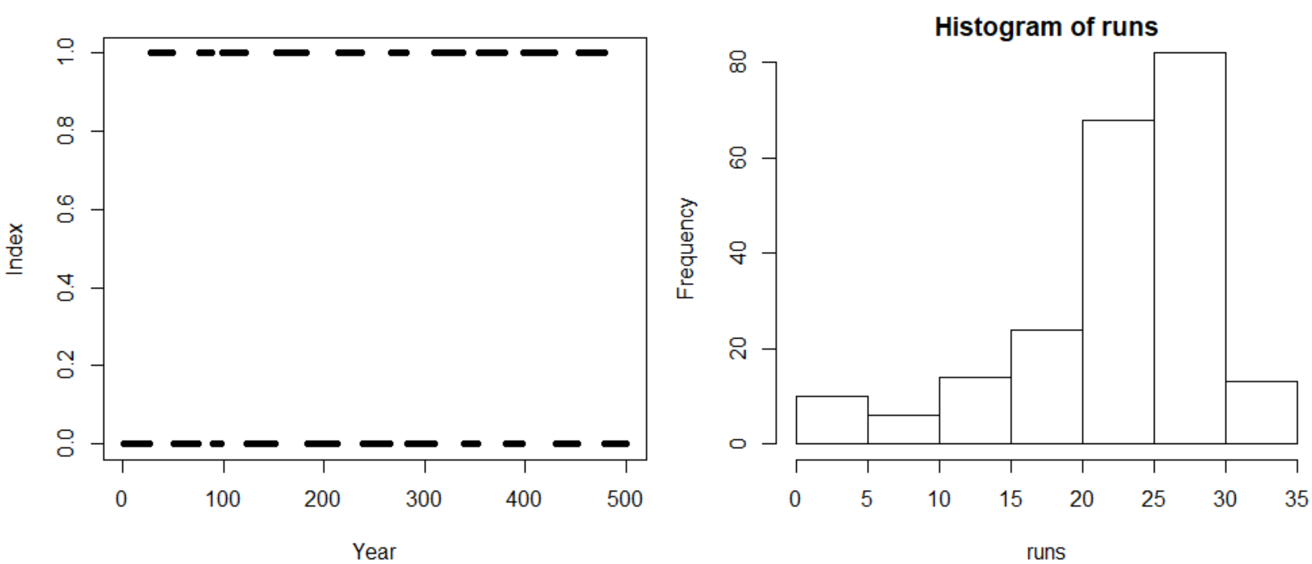

Figure 20: Simulated environmental index (500 years) and the histogram of run length (number of years with the same index) simulated for 10,000 years.

Some Additional Scenarios Not Currently Considered

Some scenarios that were not considered, but will likely be considered in the future are:

Selectivity: It may be desirable for the time-varying selectivity for at least commercial gears to be linked to changes in weight-at-age.

Migration: Migration will require a multi-area model and hypotheses about movement. A multi-area model is being developed with four regions. Migration hypotheses will be informed by tagging data as well as other observations from various fisheries and surveys.

CONDITIONING THE OPERATING MODEL

The operating model should be the best possible depiction of reality with an appropriate level of uncertainty (as described by the scenarios). 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 2016 stock assessment (Stewart & Hicks 2017). 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. Estimating natural mortality for both sexes in both models and estimating steepness were the only processes changed from the assessment model when conditioning.

Variability was characterized by the estimated variance-covariance matrix estimated automatically by inverting the Hessian within ADMB (Fournier et al. 2012), 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.

To ensure that the SS models incorporating this variability had a mean spawning biomass trajectory similar to the assessment 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 (Figure 21).

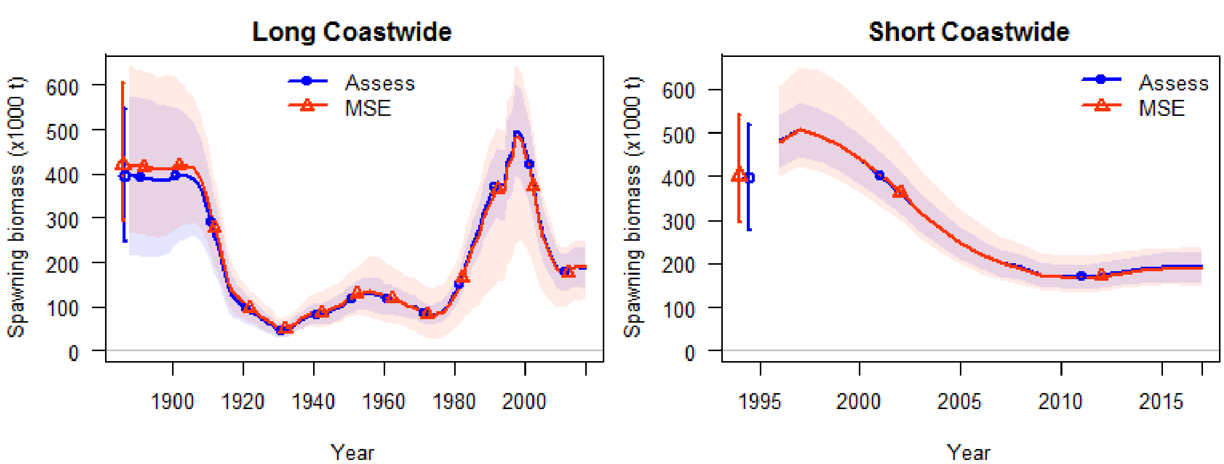

Figure 21: Comparison of mean spawning biomass trajectory and the 95% confidence interval from the assessment models (Assess) and the closed-loop simulations (MSE) for the long coastwide and short coastwide SS models.

The mean spawning biomass trajectories of the assessment and MSE (sampled) methods were equivalent, but the sampling of parameters in the MSE procedure introduced additional variability, resulting in a wider confidence interval. Part of this additional variability was due to the exploitation rate exceeding the maximum exploitation rate (catch divided by age 8+ biomass) derived from the assessment models. The long coastwide assessment model showed that the maximum exploitation rate was highly unlikely to be greater than 0.3 (99^{th} percentile of the highest estimated exploitation rate using the estimated Hessian), and the short coastwide model was much less. Other portions of the additional variability are a result of runs of recruitment deviations that were unlikely given the data (e.g., a series of low recruitments when the stock was at a low point may actually cause the population to go extinct), selectivity patterns that were restrictive (particularly for the survey), and catchability parameters. Using a multivariate normal distribution for the distribution of the parameter uncertainty is an assumption and even though correlations between parameters are accounted for in the generation of random samples, and the parameters are truncated to reasonable bounds, the distribution may over- or under-characterize some of the actual distribution (see Stewart et al 2013). Overall, though, the variability is slightly greater than the assessment model and in a reasonable range, which is desired for MSE simulations.

Two methods were used to keep the closed-loop simulation models within a reasonable range when sampling parameters. First, a truncated multivariate normal distribution was employed using parameter bounds set in the SS control file. This limited the range of some parameters, but was most useful in keeping steepness within a range of 0.2 to 1. The second method was to reject any samples that produced an exploitation rate greater than 0.5 anywhere in the time series. It is very unlikely that the historic exploitation rate exceeded 0.5 at any point in the time series given that this is a coastwide measure, and this measure was chosen allow for additional variability beyond what the assessment estimates (0.3 is the 99^{th} percentile for maximum estimated exploitation in the long coastwide assessment model). An additional measure that may be useful to identify highly unlikely trajectories is minimum spawning biomass. However, only the maximum exploitation rate was used for these results.

After making sure that the MSE procedure produced results similar to the current assessment, additional parameters were assigned variability so that the management procedures could be evaluated against a broad range of uncertainty (i.e. scenarios). For example, steepness was not estimated in either of the SS models, but was assigned variability when sampling parameters for the closed-loop simulations. Natural mortality was also assigned variability.

Estimating the parameters that were originally fixed in the assessment models typically resulted in a considerable change to the model outcomes due to a change in those parameters. For example, estimating steepness in the long coastwide model resulted in a steepness value near 0.97, and a different population trajectory. To maintain the current accepted historical trajectory, but introduce variability in these parameters that were fixed in the assessment, the fixed parameter values were used as the mean of a marginal normal distribution to generate values from, but the variability and correlation with other parameters was estimated in a model with the fixed parameters freed for estimation, without priors. These estimated correlations and variances were then used to draw a full set of parameter values from a multivariate normal distribution.

Steps to condition a SS model used in the operating model and generate realizations of parameters

1. Check that the SS model using the MSE method mimics the assessment when generating samples of the parameters estimated in the assessment model (see Figure 21). Call this model the “MSE model.”
2. Estimate additional parameters without priors and make sure that a Hessian matrix is inverted to get a covariance matrix. Call this model the “Full model”.
3. Use the covariance matrix from the Full model, but insert the estimated standard deviations for the estimated parameters from the MSE model to characterize them appropriately. This also helps maintain a positive-definite covariance matrix that is needed for the random number generator because the correlations are all estimated from the Full model.
4. Use the parameter values from the MSE model for the mean of the normal distribution. This includes the fixed values for the parameters estimated in the Full model, and ensures that the MSE model has an average that is similar to the assessment.
5. Generate random realizations of the parameters from a truncated multivariate normal distribution via rejection sampling.

A description of this process and how it relates to steepness (h) and female natural mortality (M) is given below.

Steepness was not estimated in either the long or short assessment models, but was fixed at 0.75. Therefore, it was estimated in the Full models to allow for that additional uncertainty. When estimating steepness only in the long coastwide model, it increased to 0.97 and the estimated standard error was 0.08. The small standard error may be a result of the steepness estimate being near the upper bound of one. In the short coastwide model, the estimated steepness value was 0.73 with an estimated standard error of 0.75, which is nearly a uniform distribution between 0.2 and 1.

Natural mortality was estimated for both sexes in the long coastwide model, but only for males in the short coastwide model. Steepness and natural mortality for females was estimated simultaneously in the Full short model. The estimate for female M is the short coastwide model was 0.34, considerably higher than expected. However, the fixed value (as used in the assessment) was 0.15. The estimated standard error for the female M was 0.01, which is slightly higher than the standard error for the estimated male M in the MSE model.

There are a few differences with generating parameter realizations compared to the assessment that are worth mentioning. As noted above, some parameters are given variability externally (steepness, female M for the short coastwide model). Second, randomly generated recruitment deviations for MSE in the long coastwide model are not forced to be zero-centered for the main period. The short coastwide assessment model does not zero center recruitment. Lastly, the bias adjustment on recruitment deviates is determined from the estimated variance of the deviates in the realizations, whereas the assessment model used a user defined bias adjustment ramp. Bias correcting the realizations with the ramp produced very similar results.

The long and short Full models each have a wide range of variability (Figure 22). The conditioned OM has a considerable amount of extra variability compared to the ensemble stock assessment (Figure 23). The assessment ensemble contains four individual models while the OM contains only two, which is why the trend at the end of the time series is slightly different, although well within the uncertainty.

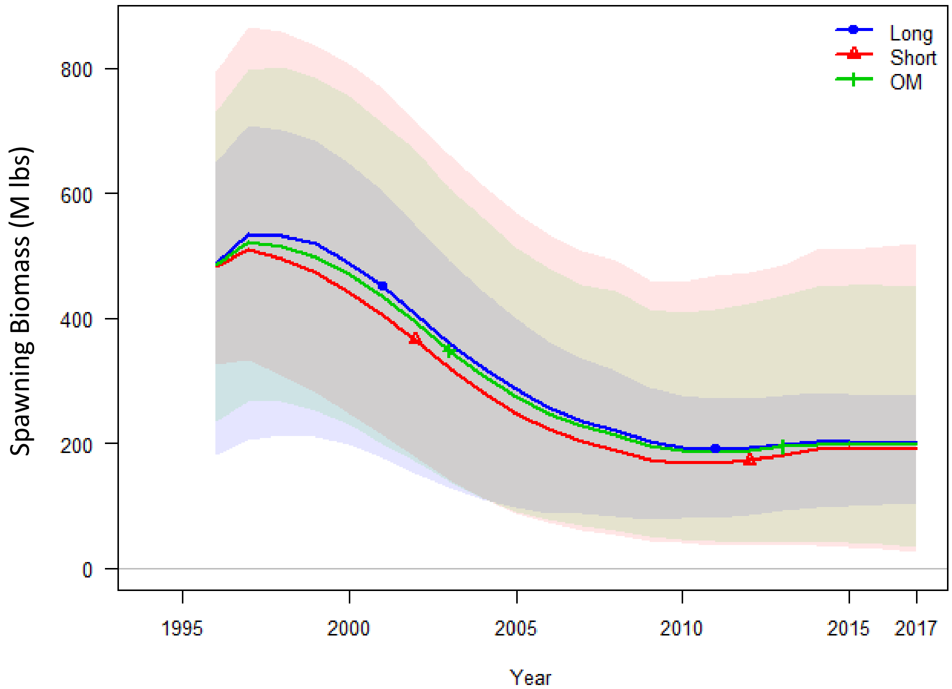

Figure 22: The long (blue) and short (red) models with additional variability estimated (natural mortality and steepness) shown as shaded 95% confidence intervals. The Operating Model (combination of these two models, weighted equally) is shown in green.

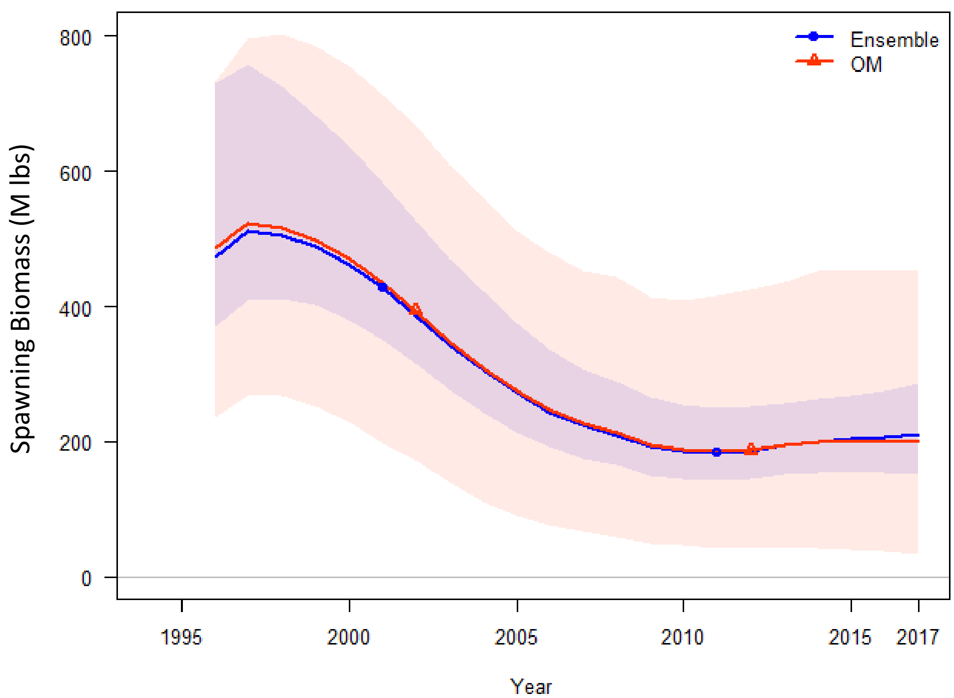

Figure 23: The conditioned operating model (red) compared to the stock assessment ensemble (blue) with 95% confidence intervals.

A potential issue highlighted at SRB11 was that starting the OM in 2017 with such a wide range of uncertainty will not adequately characterize our best knowledge of the near future (short-term) and also the medium-term. However, the long-term results are representative since the current state would not have an effect, and the wide range of uncertainty would be appropriate. One solution would be to start from the assessment model and its uncertainty (the blue shaded region in Figure 23). However, this may still result in large deviations at the start of the simulation due to the wider range of uncertainty in the parameters, and still may not be indicative of our best predictions for the short-term or medium-term. Instead, we present results for the long-term to identify a small set of management procedures that meet goals and objectives. This small set can then be further investigated using short-term predictions from the assessment model (1-3 years) to identify how they may affect the fishery now. For example, the decision table already presents risk factors for various SPR values, and these results can be used to evaluate the immediate consequences to the fishery of a change in the harvest policy. Additionally, transitory behavior from the short-term to the long-term can be highlighted. This would be describing the trends of various trajectories (e.g., catch or spawning biomass) between the short-term or long-term. For example, the short-term may indicate low catches with a higher catch on average in the long-term, but to get there, it appears that catches may be low for a short time before increasing.

The reason that it is difficult to quantify medium-term results is that we have very little predictive power for that time period. In the short-term, we have an idea of where we currently are and what may occur in the next few years (e.g., we have some data indicating recruitment and weight-at-age). And, in the long-term, we are summarizing statistics over a wide range of uncertainty (we do not need to know anything about the current state of the population). But, that uncertainty is not well known in the medium-term because it is partially dependent on the current state, but also affected by the wide range of possibilities.

MSE RESULTS

Preliminary results from the simulations are provided below, and additional results will be presented at the MSAB10 Meeting.

Outputs of the closed-loop simulations, summarized using various performance metrics developed from defined objectives, are shown in Table 2. The first set of performance metrics are related to biological sustainability and report various statistics summarizing relative spawning biomass (dRSB). The second set of performance metrics are related to fishery sustainability and report various aspects of yield. Other objectives defined by the MSAB are related to discards, bycatch, consumer needs, and preserving biocomplexity, but performance metrics have only been defined for the discards objectives. However, the discard performance metrics are not useful to report here since discards are defined as a function of total mortality (and weight-at-age), and are not modelled specifically as a process. Therefore, they would not be entirely meaningful to base decisions on.

All of the performance metrics are calculated using the last ten years of a 100-year simulation, and summarized over all simulations. For example, the median average total mortality calculates the average total mortality over the last ten years of each simulation, and then determines the median value from all of the simulations. The probabilities are simply the probability of the event occurring over all ten years and simulations, which is effectively the same as the probability of the event happening in year 100 (but has better precision since more observations, from the ten years, are included). The AAV is the average annual variability and is measure of how much the yield varies annually, on average.

The performance metrics related to biological sustainability show higher relative spawning biomass as SPR increases (i.e., fishing intensity decreases; Table 2 and Figure 24 panel c). The effect of the control rule is seen with higher RSB at lower SPR when using the 40:20 control rule, but similar RSB at higher SPR's since the control rule would rarely be invoked at these higher SPR's (lower fishing intensities). The effect of the control rule can also be seen in the median average SPR, which is the effective SPR given reductions in SPR when invoking the control rule. At low values of target SPR, the median average SPR is greater than the target SPR, while high target SPR's have a similar median average SPR. With a control rule, as the target SPR declines, the effective SPR levels off at a minimum value (Table 2 and Figure 24, panel a).

Performance metrics related to fishery sustainability show that yield in terms of total mortality (TM), directed mortality (DM; all mortality minus bycatch), and commercial (only commercial directed fishery landings) is about level between an SPR of 30% and an SPR of 40%, and declines at higher values of SPR (i.e., lower fishing intensity; Figure 25). The maximum median average total mortality is around 40 million pounds, but the variability from the simulations show that total mortality commonly ranges from less than 20 million pounds to over 80 million pounds. The variability in total mortality comes from the uncertainty in the population parameters (e.g., natural mortality) as well as variation in weight-at-age and recruitment.

The average DM from the years 1993–2012 (70.53 M lbs) was used as a benchmark for various performance metrics (Table 2). The median average DM was less than 34 M lbs in the range of SPR's and control rules simulated, thus the probabilities of being less than 70% of this benchmark are greater than 65%. The range of years used for this benchmark represent a period of time with high weight-at-age and some extremely large recruitment events, and is atypical of average conditions.

The variability in yield is represented with a number of performance metrics (Table 2). A decrease or increase in the total mortality was equally common, but increases greater than 15% were slightly more common than decreases greater than 15% (a difference in the probability of about 2%). The average annual variability (AAV) showed that the minimum variability that is likely to be achieved with these management procedures is around 6%. However, the trade-off between variability and yield becomes apparent at an SPR of 30%, where a slight increase in yield results in a nearly doubling of yield variability (Figure 24, panel d).

Table 2: Performance metrics determined from outputs of the closed-loop simulations for two different control rules and various fishing intensities indicated by Spawning Potential Ratio (SPR).

	30:20 Control Rule					40:20 Control Rule				
Target SPR	**30%**	**40%**	**46%**	**50%**	**60%**	**30%**	**40%**	**46%**	**50%**	**60%**
Median average SPR	39%	42%	47%	51%	61%	44%	45%	48%	51%	61%
Biological Sustainability										
Median average dRSB	29%	34%	41%	45%	56%	36%	39%	42%	45%	56%
P(dRSB<20%)	2%	3%	2%	2%	1%	1%	1%	2%	2%	1%
P(dRSB<30%)	63%	19%	7%	5%	2%	3%	3%	3%	3%	2%
Fishery Sustainability										
Median average Total Mortality (M lbs)	40.06	39.91	36.37	35.51	32.72	39.22	40.64	34.68	34.51	29.27
Median average DM (M lbs)	32.78	32.72	29.23	28.14	25.38	32.21	33.46	27.56	27.26	22.31
Median average Commercial (M lbs)	24.74	24.47	21.24	20.09	17.70	24.04	25.05	19.69	19.59	15.17
P(No Commercial)	8%	8%	8%	8%	10%	8%	7%	8%	8%	10%
P(DM < 70% average 1993-2012)	65%	68%	72%	73%	79%	67%	66%	72%	72%	80%
P(DM < 90% average 1993-2012)	73%	77%	80%	82%	87%	76%	76%	81%	80%	86%
P(DM < 110% average 1993-2012)	79%	83%	86%	88%	92%	84%	83%	86%	86%	91%
P(a decrease in TM)	49%	50%	51%	51%	51%	48%	48%	48%	51%	51%
P(decrease TM > 15%)	17%	6%	5%	4%	3%	12%	7%	5%	4%	3%
P(increase TM > 15%)	19%	7%	6%	5%	5%	15%	10%	7%	5%	5%
Median catch variability (AAV)	13%	7%	6%	6%	6%	10%	8%	6%	6%	6%

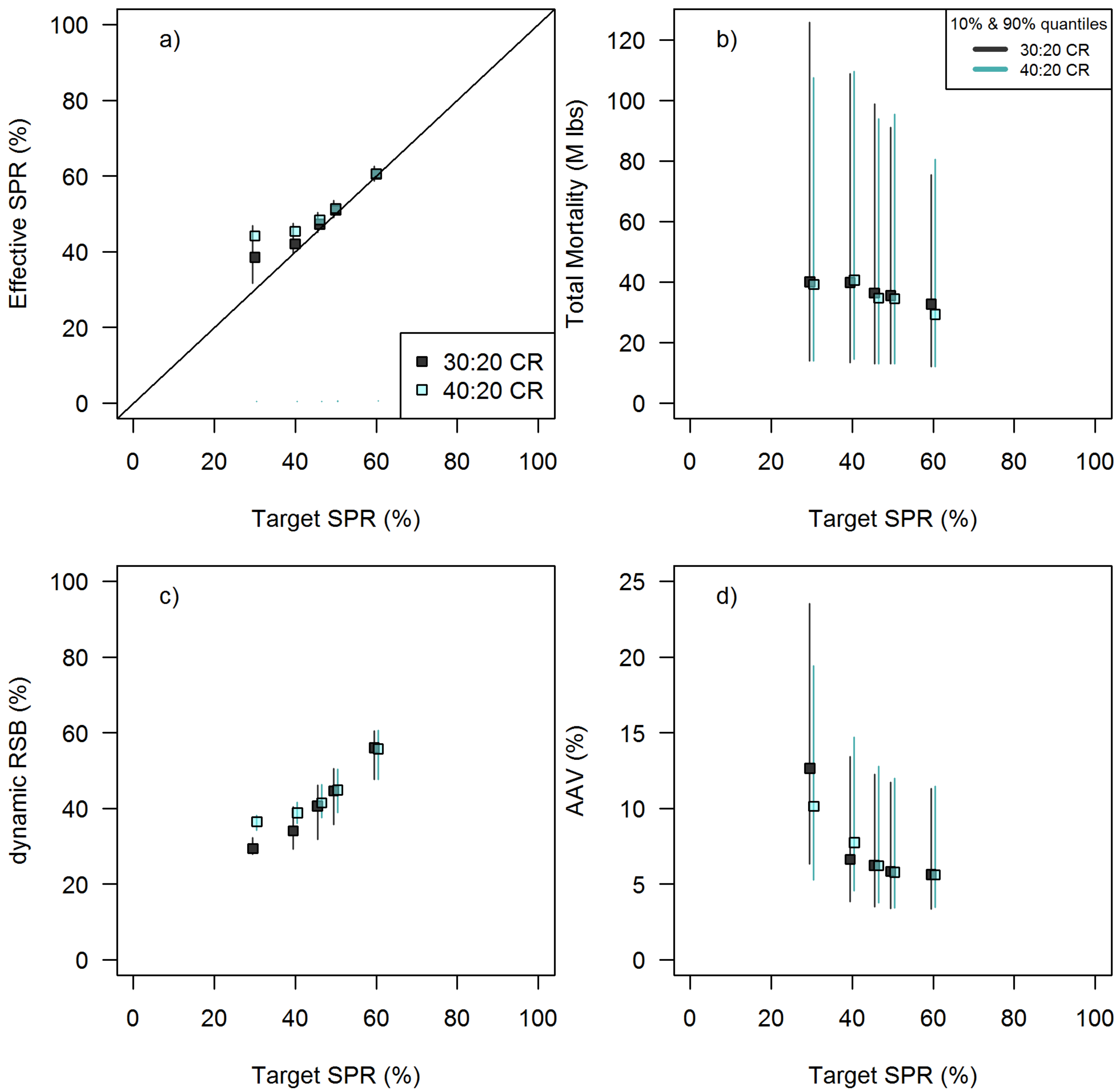

Figure 24: Various metrics plotted against the target SPR for different control rules.

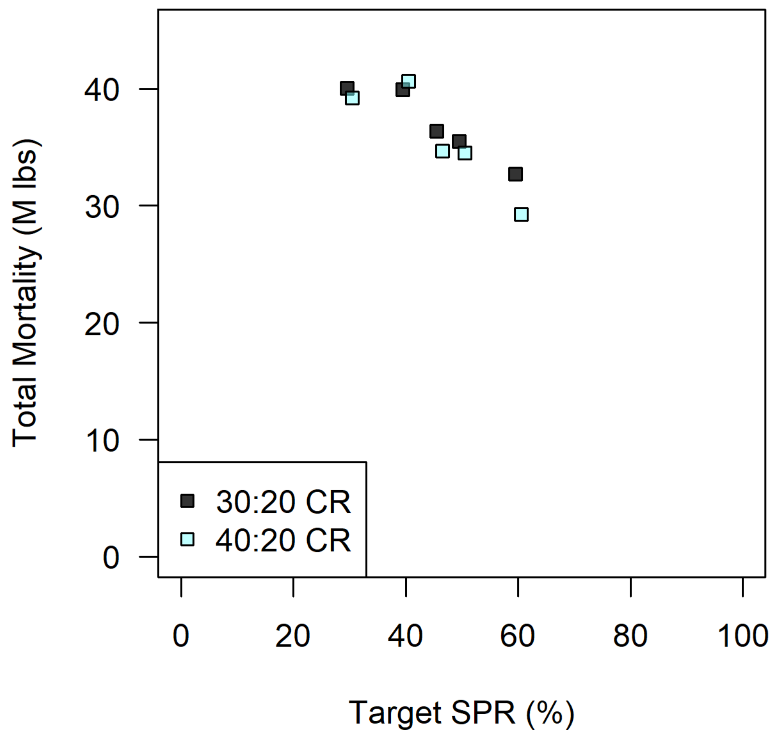

Figure 25: Total mortality plotted against target SPR for two different control rules (CR).

Along with tables and figures showing long-term results and trade-offs integrated over all of the uncertainty, some specific sensitivities may be reported (Table 3). Sensitivities are done by performing simulations with a specific state of a scenario. They will provide insight into the effects of a specific quantity, but should not be specifically used to choose a management procedure.

Table 3: Sensitivities that may be reported.

Sensitivity	**Values**
Size-at-age	High and low states
Recruitment	High and low states
Maximum bycatch	At per-area maximum regulatory bycatch
Bycatch selectivity	Shifted to a greater proportion of U26 fish
Uncertainty in total mortality	Unknown

RECOMMENDATION/S

That the Management Strategy Advisory Board:

1) **NOTE** paper IPHC-2017-MSAB10-09 which provided an overview of the simulation framework to evaluate the fishing intensity and harvest control rules in the IPHC harvest strategy policy.
2) **CONSIDER** the simulation framework and assumptions as described, including introducing variability to the OM, simulating weight-at-age and an environmental regime, and distribution of the Total Mortality to sectors.
3) **CONSIDER** the interpretation of short-term, medium-term, and long-term results.
4) **CONSIDER** the long-term results looking at the outcomes of various management procedures and the trade-offs between them.
5) **RECOMMEND** modifications to the simulation framework and assumptions.
6) **RECOMMEND** management procedures that would meet the goal and objectives defined by the MSAB (when results are available).
7) **RECOMMEND** a management procedure to update the IPHC interim harvest strategy, or to continue using the interim status quo harvest strategy.
8) **AGREE** on additional management procedures to evaluate in 2018.

ADDITIONAL DOCUMENTATION / REFERENCES

Clark WG. 1993. The Effect of Recruitment Variability on Choice of Spawning Biomass per Recruit. Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. Eds: G. Kruse, R. J. Marasco, C. Pautzke and T. J. Quinn. University of Alaska, Alaska Sea Grant College Program Report. 93-02: 233-246.

Harley SJ, Davies N, Tremblay-Boyer L, Hampton J. McKechnie S. 2015. Stock assessment for south Pacific albacore tuna. WCPFC-SC11-2015/SA-WP-06 Rev 1 (4 August 2015)

McCall AD, Klingbeil RA, and Methot RD. 1985. Recent increased abundance and potential productivity of Pacific mackerel (Scomber japonicus). CalCOFI Rep. 26: 119-129

Methot RD and Wetzel CR. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86-99.

Stewart IJ and Hicks AC. 2017. Assessment of the Pacific halibut stock at the end of 2016. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016: 365-394.

Stewart IJ, Hicks AC, Taylor IG, Thorson JT, Wetzel C, Kupschus S. 2013 A comparison of stock assessment uncertainty estimates using maximum likelihood and Bayesian methods implemented with the same model framework, Fisheries Research, Volume 142. pp 37-46. ISSN 0165-7836. http://dx.doi.org/10.1016/j.fishres.2012.07.003

APPENDIX A: DETAILS OF THE SIMULATION FRAMEWORK PRESENTED AT MSAB09 AND SRB10

OPERATING MODEL

The operating model represents an uncertain reality, or the states of nature. In other words, it is a computer program that simulates a population that one would normally not observe in its entirety. For example, this could be a model for the coastwide halibut population.

An operating model may be simple or complex, depending on the outcomes desired, and is designed to simulate a population given a set of pre-defined parameters. The scope of the parameter set depends on the defined complexity. These parameters may define natural mortality, recruitment, selectivity, or migration, or be related to any number of processes.

Uncertainty in the simulated population is introduced in two different ways: parameter uncertainty and model uncertainty. Parameter uncertainty involves changing the parameters to reflect the range of possible values for those parameters. The most straightforward way to introduce parameter uncertainty is to simply change the parameter from one value to another (e.g., change natural mortality from 0.15 to 0.2). This is more like a sensitivity analysis and does not best represent the range of estimated values for that parameter, although it could indicate the extremes or specific quantiles and be included as two separate values in two models (similar to how an ensemble of models works). A more complete and integrated approach is to sample parameter values from a joint probability distribution to apply to the operating model. This joint distribution may be determined from data.

Model uncertainty involves a change to the model structure or specification. For example, it may be adding a specific migration assumption, introducing density-dependence on recruitment or growth, or even changing the way that a fishery (e.g., bycatch) interacts with the population.

The simulations integrate over a range of scenarios (potential states of nature) by drawing parameters from a joint probability distribution and integrating over multiple model structures. This is similar to the ensemble modeling approach used for the Pacific halibut stock assessment, except that the scenarios could be more theoretical rather than practical. For the current stock assessment, it is not practical to fit a migration model to data and provide reasonable management advice until there is a better understanding of Pacific halibut migration. However, for the MSE simulations, supplying hypothetical, yet reasonable, migration models as scenarios will be useful to evaluate the potential effects of migration on the outcomes. In other words, it is an additional source of uncertainty in the state of nature. Potential scenarios are described in a separate section below.

Simulating processes like size at age, recruitment, recruitment regime, and others requires some assumptions about the temporal trends or changes in those processes. These processes cannot specifically be controlled by a management procedure, thus are used as scenarios in the operating model. These processes and other scenarios are discussed further in the Scenarios section below.

For the simulations to investigate a coastwide fishing intensity, I used the Stock Synthesis (Methot and Wetzel 2013) assessment program that is currently used for the stock assessment, and 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 need to 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 and is currently in development.

The current ensemble composed of four different assessment models includes a cross between coastwide or fleets-as-areas, and the length of the time series. Using a fleets-as-areas 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 the feedback from migration and productivity of harvesting in different areas. Therefore, I used the two coastwide models, with additional variability. These models use five fisheries: commercial, wastage, bycatch, sport, and personal use, and the TCEY was distributed to each fleet in an ad hoc manner (see Scenarios). The survey is also included as a fleet without catch.

MONITORING (DATA GENERATION)

The operating model represents the population and the data generation model simulates the process of collecting data from that population. There are many data collection programs in fisheries, but we will focus on the data needed in the estimation model (i.e., coastwide stock assessment model) and the harvest strategy (i.e., distribution). In the sequence of the simulation, data generation occurs between the operating model and estimation model. The data types to be generated are listed in Table 4.

Table 4: Data to generate and information about each.

Data	Type	Area	Frequency	Sexes	Probability Distribution	Bias[1]	Uncertainty
Survey total NPUE	Fishery-independent	Coastwide	Annual	Combined	Lognormal	No	From Assessment
Age composition for survey total numbers	Fishery-independent	Coastwide	Annual	Two sexes	Dirichlet	Possibly selectivity	From Assessment
WPUE for the directed commercial fishery	Fishery-dependent	Coastwide	Annual	Combined	Lognormal	Assumptions about catchability	From Assessment
Age composition from directed commercial fishery	Fishery-dependent	Coastwide	Annual	Combined	Dirichlet	Possibly selectivity	From Assessment
U26 age composition from the survey (proxy for wastage)	Fishery-independent	Coastwide	Annual	Two sexes	Dirichlet	Possibly selectivity	From Assessment
Age composition for the bycatch fisheries.	Fishery-dependent	Coastwide	Annual	Combined	Dirichlet	Possibly selectivity	From Assessment
Age composition for the sport fishery.	Fishery-dependent	Coastwide	Annual	Combined	Dirichlet	Possibly selectivity	From Assessment

[1]Bias is whether there is a difference between the generated data and the EM assumptions.

ESTIMATION MODEL

The estimation model in MSE simulations mimics the stock assessment, or the model and process used to estimate quantities needed for the harvest strategy from which to provide catch advice. At IPHC, the stock assessment is based on four models using the Stock Synthesis framework. The results from the four models are combined to produce uncertainty related to observation error and structural error, which is then translated into the decision table presented to Commission. One line in that decision table contains estimated catch levels consistent with the current harvest policy.

To capture the uncertainty of the stock assessment in the MSE simulations, the estimation quantities needed for catch advice are simulated with the estimation model. Simulating an annual stock assessment can be time-consuming, but is important to capture that source of variability. The following set of methods for simulating the stock assessment are useful for the MSE simulations.

- **Perfect Information**: This assumes that the quantities necessary for applying the management procedure and determining catch levels for the next year are known exactly and without estimation error. In other words, the Commissioners would have perfect information to guide them on their decision. It is not quite an omniscient Commission because they would only know the information for the current year and not all of the necessary information to maximize objectives in the future.
- **Simulate Error**: The method would simply take the simulated abundance/biomass from the operating model and apply variability to it. It is likely that this variability would be lognormal, which introduces a long tail for uncertain higher values, and forces the randomly generated value to be greater than zero. This is an approximation to the stock assessment with simplistic assumptions of constant error across all levels of abundance/biomass and no bias, but could be expanded to model variability as a function abundance/biomass and even introduce bias in some way. However, it will always be an approximation that does not take into account the types of data collected, the frequency of data collected, and the correlation of error given historic (unchanging) data.
- **Single Stock Assessment**: An actual stock assessment model can be run using the data generated during the Data Generation step. For example, one of the models used in the ensemble could be run to estimate the necessary inputs to the harvest strategy (i.e., catch at F_{SPR} and stock status). Running an age-structured stock assessment model like Stock Synthesis can be time-consuming, and doing that annually in the simulations can significantly increase the time for a run to complete. However, using the actual stock assessment model in the simulations captures the nuances of the combinations of data and will produce results that are more similar to what may be expected in practice.
- **Ensemble of Models**: An ensemble of four stock assessment models is currently used for short-term catch advice at IPHC. This process would be similar to running the single stock assessment model, except that it would run four models and then combine the results. Using parallel processing in modern day computers may not significantly increase the run time, except that the ensemble would be as slow as the slowest model. In addition, there are increased chances for a model not converging and causing further delays in the simulations.

MANAGEMENT PROCEDURE

The 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 AM094. In addition to fishing intensity, a control rule was used to adjust the fishing intensity at low stock status. This is discussed below in the Management

Procedures section. For these simulations, I used the two coastwide models, and only needed to distribute the catch to the five coastwide fleets.

MEASURES OF FISHING INTENSITY

Fishing intensity is a measure of how fishing is affecting the stock, and it is the management procedure in determining the scale of the current harvest policy shown in Figure 1. An intuitive measure of fishing intensity is an exploitation rate, which is simply the catch divided by the exploitable biomass. Less intuitive, but similar, is instantaneous fishing mortality, which is used in an exponential function, as is *M*. These are obvious measures of fishing intensity for a single fleet, but become very complicated when considering multiple fleets with different selectivities or annual changes in selectivity.

Measures of fishing intensity have been developed that are more holistic and provide a meaningful measure of fishing effort on the stock of fish, rather than a specific portion. Many of these metrics focus on the effect of fishing on the spawning biomass, and often measure the long-term effects after fishing consistently at the same intensity. The following are some of the desired properties of a fishing intensity metric (many from pers. comm., Owen Hamel, NWFSC).

- As fishing effort increases, the fishing intensity metric also increases appropriately.
- Applies to simple as well as complex (i.e., multiple areas and fleets) models.
- Metric changes with changes in selectivity, and captures systematic changes in selectivity.
- Easy to compute.
- A scale that is easy to understand.

A commonly used metric is the spawning potential ratio (SPR), which is a measure of the effect of fishing on the long-term reproductive potential of the stock. This metric is currently used in the IPHC interim harvest policy. Potential metrics to consider for evaluation are listed below along with descriptions. Table 5 compares the metrics.

U (exploitation rate): catch divided by a summary biomass (which may or may not be exploitable biomass). This metric ignores differences between fisheries and their impacts of different ages, sizes, and sexes. Changes in selectivity will not be captured by U unless selectivity is specifically defined in the summary biomass (as with exploitable biomass). Overall, this is not a useful metric when there is more than one fishery.

F (instantaneous fishing mortality): a measure of the fishing mortality on the most highly selected age, size, and sex. Catch is a function of F and selectivity, and a change in selectivity results in a change to the meaning of F. The scale of F is also not easily interpretable. Overall, this is a useful measure of fishing mortality for modelling, but is not as useful as a metric.

SPR (spawning potential ratio): a measure of the effect of fishing on the long-term reproductive potential of the stock. More specifically, it is the percentage of long-term, equilibrium spawning output-per-recruit when fishing at a constant fishing intensity (F_{SPR}), divided by the long-term, equilibrium spawning output-per-recruit without fishing. Spawning output for Pacific halibut is measured by spawning biomass. The higher the fishing intensity (F_{SPR}), the lower the SPR (Figure 26). For example, SPR=100% is, by definition, no fishing; and SPR=40% is a fishing level that reduces the equilibrium spawners-per-recruit (i.e., spawning potential) to 40% of the unfished level. The general equation for SPR is

$$SPR = \frac{\widetilde{SB}_F \Big/ R_F}{\widetilde{SB}_{noF} \Big/ R_{noF}} \quad (1)$$

where $\widetilde{SB}$ is the spawning biomass simulated forward to equilibrium with fishing (*F*) or without fishing (*noF*), and *R* is recruitment.

SPR, in general, is slightly different than simply dividing equilibrium spawning biomass when fishing by unfished equilibrium spawning biomass because SPR is on a per-recruit basis, thus eliminating the density-dependent effects of the spawner-recruit curve and simply measuring equilibrium spawning potential (see a comparison in Table 5**Error! Reference source not found.**). In other words, SPR is the r elative spawning potential of a recruit when faced with natural and fishing mortalities. SPR-based harvest policies are commonly used in the management of many fisheries around the world, including fisheries under U.S. fishery management council jurisdiction. An $F_{SPR=46\%}$ policy is currently the interim harvest policy at IPHC. Clark (1993) recommended that a $F_{SPR=40\%}$ for groundfish fisheries would maintain a high average yield.

To calculate SPR, the biology of the species (e.g., natural mortality, maturity, etc.), the selectivity for each fishery, and an overall fishing intensity (or fishing intensities for each fishery) are needed. The calculation of SPR always uses the biology and selectivities in the year of interest, thus accounts for changes in these parameters. However, an appropriate SPR for management should be robust to these changes.

This calculation of SPR is called static %SPR by Mace et al. (1996), and we will simply refer to it as SPR. Mace et al (1996) also presented the concept of “transitional SPR”, which looks at the impact of fishing on existing cohorts in the stock (those that were present back in time) and thus is more of a retrospective measure, rather than quantifying current or future impacts. We do not consider transitional SPR metrics because those metrics are better suited to determine the level at which a stock has been fished, rather than providing a metric of how the stock is to be fished. The static %SPR (from now on simply called SPR) provides a measure of SPR given the current biological regime, fishery patterns, and a fishing intensity (F_{SPR}). See Mace et al. (1996) for further discussion of the difficulties calculating transitional SPR.

The metrics SPR and F_{SPR} has been reported in previous Pacific halibut assessments and are commonly calculated in many stock assessments around the world. It is a useful metric because it accounts for complex and temporally changing population dynamics and selectivities. It can be thought of as a measure of the spawning potential given fishing under the current conditions.

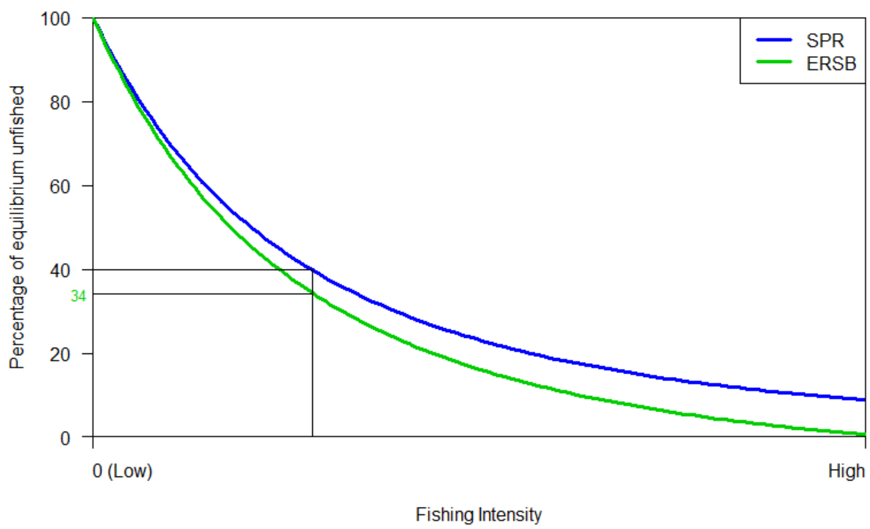

Figure 26: SPR and ERSB plotted against fishing intensity for a generic equilibrium model with constant recruitment (unweighted SPR) and time-invariant biology and selectivity.

ERSB (Equilibrium Relative Spawning Biomass): the long-term equilibrium relative spawning biomass given a level of fishing. Relative spawning biomass is the percentage of equilibrium spawning biomass with fishing ($F_{XX\%}$) relative to that without fishing. This was called ESD, or Equilibrium Stock Depletion, by Cordue (2012), but the term relative spawning biomass is used at the IPHC instead of depletion. The calculation is simply the equilibrium spawning biomass when fishing divided by unfished equilibrium spawning biomass. The calculation uses constant recruitment, and accounts for density-dependence of the stock-recruit relationship. In other words, this is the effect of fishing on the deterministic spawning potential of the stock, which reflects the decline in recruitment as the spawning biomass declines.

$$ERSB = \frac{\widetilde{SB}_F}{\widetilde{SB}_{noF}}$$

where $\widetilde{SB}$ is the spawning biomass simulated forward to equilibrium with fishing (*F*) or without fishing (*noF*). The only difference from SPR is the division by the number of recruits, and ERSB can be easily calculated from SPR using the following equation (with a Beverton-Holt stock-recruit relationship).

$$ERSB = \frac{4hSPR + h - 1}{5h - 1} \qquad (2)$$

where *h* is steepness in the Beverton-Holt stock-recruit relationship. Notice that when steepness is equal to one (constant recruitment at all spawning stock sizes), ERSB is equal to SPR.

As with SPR, when temporal trends are present, the biology and selectivity used when calculating ERSB can affect the outcome. It is proposed to use the current conditions and project forward to determine the equilibrium spawning biomass with and without fishing. This keeps ERSB consistent with SPR and maintains the relationship in Equation (2). However, SPR and ERSB are similar metrics that can be calculated from one another, thus only one should be used for setting fishing intensity. RSB is currently used in the 30:20 control rule of the harvest policy, which may be a useful place in the harvest policy to use ERSB as a translation of the SPR value to a target RSB. However, RSB is slightly different than ERSB because the denominator in RSB is consistently B_0, which does not consider current biological conditions (but defined equilibrium conditions) when calculating. We'll discuss this more in the Control Rule section below.

FR (Fishing Ratio): the ratio of the biomass of fish that die due to fishing to the biomass of fish that die due to natural causes. This is not an equilibrium metric, but provides an insight into the current effect of fishing on the stock. This metric may be useful to gauge recent or current impacts due to fishing, but is not as useful for long-term management and strategic thinking. It could be used, for example, to set a maximum annual impact on the stock.

SER (Spawning Exploitation Rate): a measure of the reduction in spawning biomass due to fishing at a certain level, and was also termed "Annual Foregone Reproduction" by Mace et al. (1996). This metric was suggested by the SRB (at SRB09) and is calculated as $1\text{-}(SB_{\text{fishing},y}/SB_{\text{noFishing},y})$, where *SB* indicates spawning biomass and *y* is the year. This metric ranges from 0 to 1, with higher values indicating higher fishing intensity. It is not an equilibrium metric and does not specifically account for the mortality of smaller, immature fish. A target value will take into account the selectivity patterns of the fisheries, but it may be sensitive to shifts in selectivity. Overall, this metric is similar to SPR except that it is based on the immediate term rather than long-term equilibrium calculations. It may be an interesting metric to report annually, regardless of the fishing metric used to define the scale.

RFY (Relative Foregone Yield): this is the equilibrium yield calculated using current conditions and fishing intensity divided by the maximum equilibrium yield under current conditions. It provides a measure of the percentage of MSY, which can be related to "Pretty Good Yield" (Hilborn 2010). More thought needs to be given to this metric, but it is likely not a useful metric to determine fishing intensity mainly because there are two sides to the yield curve, and being at a percentage of MSY could mean that the stock is larger than expected at MSY or smaller than expected at MSY. It may be a useful metric to report and monitor, or to use in evaluations of different management procedures.

CONTROL RULE

The control rule is an additional part of the harvest policy that affects the fishing intensity or FCEY. The premise of a control rule is that if the stock declines below a threshold reference point (typically measured in relative spawning biomass) the fishing intensity is reduced, and if the stock declines below a limit reference point there is no harvest. This is used to avoid low stock sizes by acting in a precautionary manner when the stock size begins to approach the limit. The current IPHC control rule is called a 30:20 rule because the threshold is 30% RSB and the limit is 20% RSB (Figure 27).

The multiplier can act on the fishing level (i.e., fishing intensity) or the catch (i.e., FCEY), and it would be somewhat straightforward for the fishing intensity to be adjusted. For example, if $F_{SPR=46\%}$was the fishing intensity, the F could be adjusted, or the SPR could be adjusted. The relationship between SPR and FI is nonlinear (Figure 26) thus a linear adjustment to one would result in a nonlinear adjustment to the other. It would be most straightforward for the SPR to be adjusted.

Table 5: A comparison of the different fishing intensity metrics.

Metric	Name	Multiple fisheries and areas	Equilibrium	Easy to calculate	Easy to interpret	Range	Account for all fishing mortality on all sizes	Current conditons or regime	Use?
U	Exploitation Rate	No	No	Yes	Yes	0–100%	No	Possibly	Leave to modelling
F	Instantaneous Fishing Mortality	No	No	Yes	No	0–∞	No	Yes	
SPR	Spawning Potential Ratio	Yes	Yes	Yes	Yes	0–100%	Yes	Yes	For Fishing Intensity
ERSB	Equilibrium Relative Spawning Biomass	Yes	Yes	Yes	Yes	0–>100%	Yes	Possibly (dynamic?)	For Control Rule
FR	Fishing Ratio	Yes	No	Yes	Yes	0–∞	Yes	Yes	Report or use as a Performance Metric
SER	Spawning Exploitation Rate	Yes	No	Yes	Yes	0–100%	Yes	Yes	
RFY	Relative Foregone Yield	Yes	Yes	Yes	Ues	0–100%	Yes	Yes	

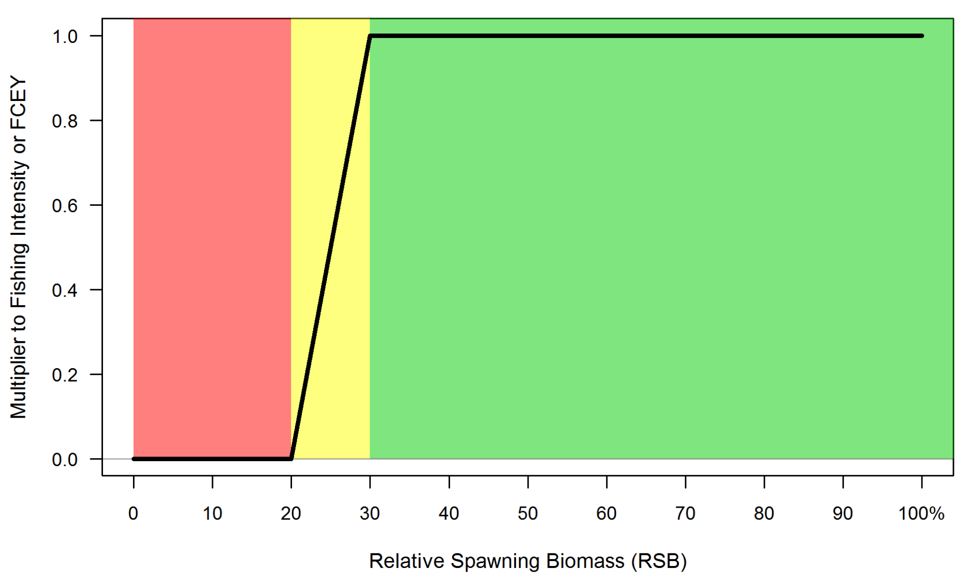

Figure 27: Control rule for the IPHC harvest policy. It is commonly called a 30:20 control rule because the downward adjustment begins at a relative spawning biomass (RSB) of 30% (threshold reference point) and no harvest occurs when the RSB is below 20% (limit reference point). The adjustment has been made to the harvest rates in the past, but the adjustment may be made to the fishing intensity or to the FCEY.

Adjusting the catch may be more difficult because there are portions of the catch that are not directly controlled by the IPHC. It would be possible to adjust the FCEY, but the other components of the TCEY as well as the U26 mortality would not be adjusted. This also brings up an important point about adjusting the FI. The FI defines the total mortality, some of which is not controlled by the IPHC, therefore the FI would not decline to zero when below the limit threshold unless cross-agency management measures were agreed upon.

The current IPHC assessment used RSB to determine stock status with a static unfished equilibrium biomass (B_0), calculated assuming good size-at-age and poor recruitment, as the reference. This static definition has many potential problems. First, it is not necessarily reflective of current conditions. Second, if fishing were to stop and current conditions remained constant, the RSB would not go to one, but could be less than or greater than one. Lastly, a change in conditions could potentially result in a RSB below the threshold even without fishing. In some cases a specific static reference point may be the desired target, but not accounting for current conditions may be misleading when managing a dynamic stock subject to changing conditions.

SPR is currently used to define fishing intensity, which is an equilibrium concept using current conditions. When a target SPR is defined, a target ERSB is also defined (assuming the Beverton-Holt stock-recruit curve and a value for steepness). With a target related to stock status one may also define a threshold and limit in relation to that target (Figure 28). However, the x-axis of the control (stock status) should also be based on current conditions instead of a static definition.

A dynamic quantity to define current stock status would be consistent with SPR and ERSB and could be used to determine at what stock status fishing intensity is reduced. A desirable property for the current status may be that if fishing had not occurred over the last generation, then the calculation of the status would result in a value of one. In other words, the current status would be a measure of the effect of fishing and not include the effect of changing conditions or recent deviations in recruitment. Dynamic B_0 (McCall et al. 1985) is a dynamic calculation of stock status that uses the conditions and recruitment deviations that the stock has recently experienced. It also corrects for the reduction in average recruitment due to the stock-recruit function. This quantity has also been used in tuna assessments (Harley et al 2015).

Using SPR and translating that to ERSB results in consistent equilibrium quantities for fishing intensity and target stock status. Dynamic B_0 is the consistent link to determine the current fishing effect on stock status, which we call dynamic RSB (dRSB). Using these three quantities provides for a control rule where each component relates to each other in a meaningful way. For example, a stock would be expected to fluctuate around a target ERSB due to natural variability in recruitment. It is likely that dropping below the threshold is not a highly desired state due to a curtailing of fishing effort, and if the threshold was near the target, it would be crossed often due to variability in the population. Setting the threshold less than the target reduces the probability of curtailing fishing effort and builds the stock back to expected levels when the current stock status is lower than desired. However, if the desire is to build back to the target as quickly as possible, a threshold closer to the target may be useful.

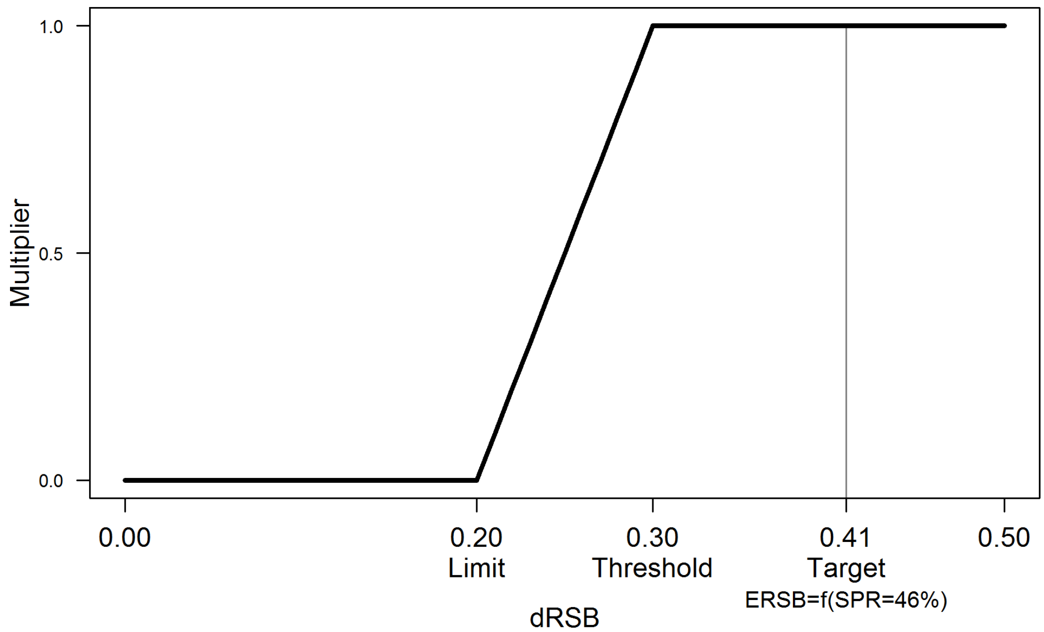

Figure 28: Components of the control rule.

A concern may be that in extreme cases where non-fishing related influences result in a static stock status below a threshold, the dynamic approach would not reduce the fishing intensity appropriately to maintain a minimum spawning biomass. Using SPR to define a fishing intensity helps to alleviate this concern since it determines a relative spawning potential. Even though SPR is based on current conditions, it still maintains a minimum spawning potential.

A consistency between reference points is useful because it helps to relate the different components of the control rule to each other and define meaningful values. The control rule is part of the harvest strategy determining the scale of fishing and the MSAB decided to test 30:20 and 40:20 control rule reference points using dRSB in conjunction with SPR values. Given the advice of the SRB, RSB will also be investigated as well as other thresholds and limits deemed appropriate given initial results.