IPHC Secretariat MSE Program of Work (2022–2023) and an update on progress

PREPARED BY: IPHC SECRETARIAT (A. HICKS & I. STEWART; 18 AUGUST 2022)

PURPOSE

To provide the Scientific Review Board (SRB) with an update of progress on the Management Strategy Evaluation (MSE) program of work for 2022–2023 and a look at preliminary results.

1 INTRODUCTION

The current interim management procedure (MP) at the International Pacific Halibut Commission (IPHC) is shown in Figure 1.

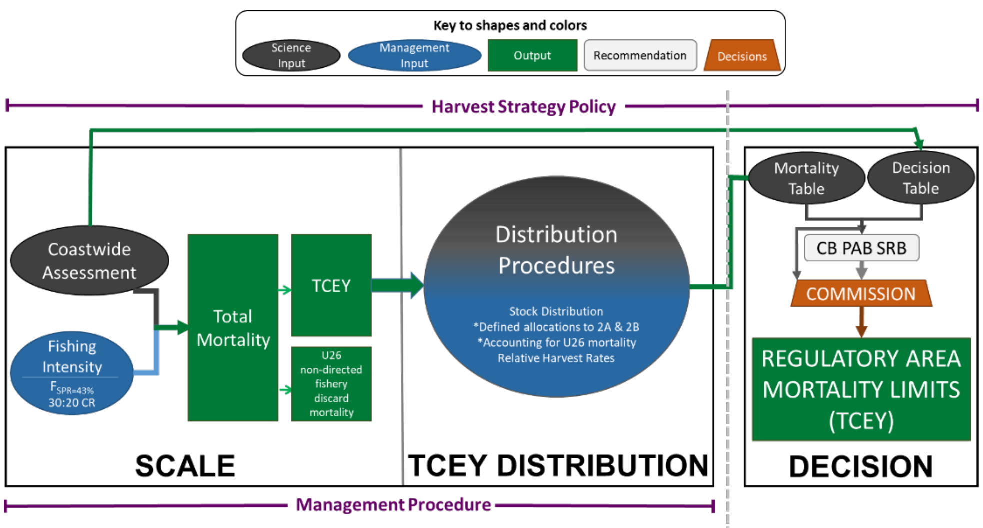

Figure 1. Illustration of the Commission interim IPHC harvest strategy policy (reflecting paragraph ID002 in IPHC-2020-CR-007) showing the coastwide scale and TCEY distribution components that comprise the management procedure. Items with an asterisk are interim agreements in place through 2022. The decision component is the Commission decision-making procedure, which considers inputs from many sources.

The Management Strategy Evaluation (MSE) at the IPHC completed an evaluation in 2021 of management procedures (MPs) relative to the coastwide scale and distribution of the Total Constant Exploitation Yield (TCEY) to IPHC Regulatory Areas for the Pacific halibut fishery using a recently developed closed-loop simulation framework. The development of this closed-loop simulation framework supports the evaluation of the trade-offs between fisheries management scenarios. Descriptions of the MPs evaluated and simulation results are presented in Hicks et al. (2021). Additional tasks were identified at the 11th Special Session of the IPHC (IPHC-2021-SS011-R) to supplement and extend this analysis for future evaluation (Table 1). Document IPHC-2021-MSE-02 contains details of the current MSE Program of Work.

Table 1. Tasks recommended by the Commission at SS011 (IPHC-2021-SS011-R para 7) for inclusion in the IPHC Secretariat MSE Program of Work for 2021–2023.

ID	Category	Task	Deliverable
F.1	Framework	Develop migration scenarios	Develop OMs with alternative migration scenarios
F.2	Framework	Implementation variability	Incorporate additional sources of implementation variability in the framework
F.3	Framework	Develop more realistic simulations of estimation error	Improve the estimation model to more adequately mimic the ensemble stock assessment
F.5	Framework	Develop alternative OMs	Code alternative OMs in addition to the one already under evaluation.
M.1	MPs	Size limits	Identification, evaluation of size limits
M.3	MPs	Multi-year assessments	Evaluation of multi-year assessments
E.3	Evaluation	Presentation of results	Develop methods and outputs that are useful for presenting outcomes to stakeholders and Commissioners

This document provides updates on the progress for the framework related tasks and the MP related tasks. Potential improvements to the evaluation and presentation of results are provided in this document and work will continue in 2022 with input from the SRB and MSAB.

2 Closed-Loop Simulation Framework

The closed-loop framework (Figure 2) with a multi-area operating model (OM) and three options for examining estimation error was initially described in Hicks et al. (2020b). Technical details are updated as needed in IPHC-2022-MSE-01 on the IPHC MSE webpage. Improvements to the framework have been made in accordance with this program of work and a new OM has been developed.

2.1 Development of a new Operating Model

The IPHC stock assessment (Stewart & Hicks 2022) consists of four stock synthesis models integrated into an ensemble to provide probabilistic management advice accounting for observation, process, and structural uncertainty. A similar approach was taken when developing the models for the closed-loop simulation framework along with some other specifications to improve the efficiency when conditioning models and running simulations.

2.1.1 General specifications of the OM

The emerging understanding of Pacific halibut diversity across the geographic range of its stock indicates that IPHC Regulatory Areas should be only considered as management units and do not represent relevant sub-populations (Seitz et al. 2017). Therefore, four Biological Regions (Figure 3) were defined with boundaries that matched some of the IPHC Regulatory Area boundaries (see Hicks et al 2020b for more description). The OM is a multi-regional model with population dynamics modelled within and between each Biological Region, and fisheries mostly operating at the IPHC Regulatory Area scale. Multiple fisheries within a Biological Region may have different selectivity and retention patterns to mimic differences similar to that of an areas-as-fleets approach. Thirty-three fisheries were defined for five general sectors consistent with the definitions in the recent IPHC stock assessment:

- **directed commercial** representing the O32 mortality from the directed commercial fisheries including O32 discard mortality (from lost gear or regulatory compliance);
- **directed commercial discard** representing the U32 discard mortality from the directed commercial fisheries, comprised of Pacific halibut discarded due to the minimum size limit;
- **non-directed commercial discard** representing the mortality from incidentally caught Pacific halibut in non-directed commercial fisheries;
- **recreational** representing recreational landings (including landings from commercial leasing) and recreational discard mortality; and
- **subsistence** representing non-commercial, customary, and traditional use of Pacific halibut for direct personal, family, or community consumption or sharing as food, or customary trade.

Additionally, there are four modelled surveys, one for each Biological Region.

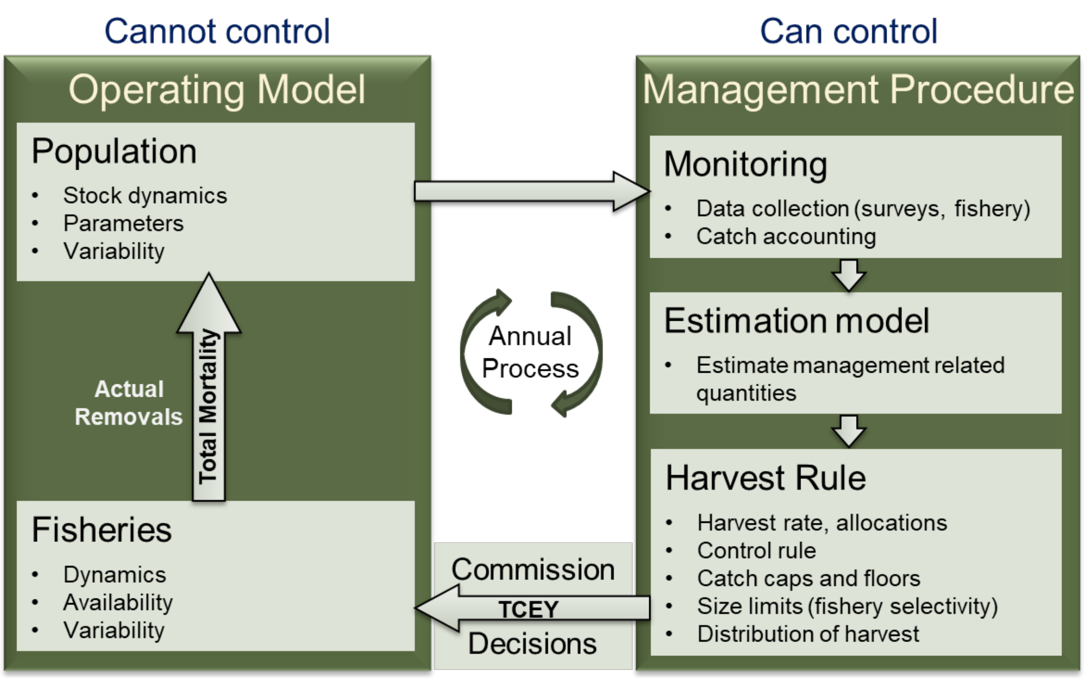

Figure 2. Illustration of the closed-loop simulation framework with the operating model (OM) and the Management Procedure (MP). This is the annual process on a yearly timescale.

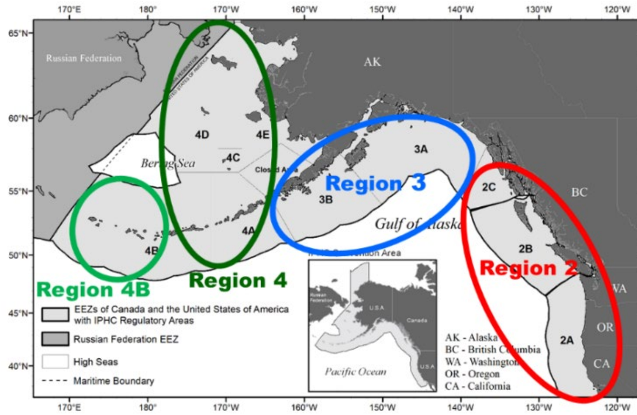

Figure 3. IPHC Regulatory Areas, Biological Regions, and the Pacific halibut geographical range within the territorial waters of Canada and the United States of America.

Two of the four models in the IPHC stock assessment (Stewart & Hicks 2022) consider a long time-series of observations beginning in 1888. One model specifies coastwide fisheries (called the coastwide (CW) long model) and the other model specifies four regions in an areas-as-fleets approach (called the areas-as-fleets (AAF) long model). The previous MSE OM also started in 1888 and simulated the entire time-series up to recent years before starting the forward simulations. However, the early portion of the time-series is challenging to model due to relatively little data, some significant catches in Biological Region 2, and the potential for unknown differences in population dynamics (e.g. movement between Biological Regions) compared to recent periods. To reduce the technical complexity and focus on information contained in the richer data set in the later period, the 2022 OM models were started in 1958. In order to allow for flexible starting conditions, 30 years of initial recruitment and an average fishing mortality were estimated for each fleet. This initialized the models after a bottleneck of potentially high fishing mortality in the 1930's that is confounded with the estimation of movement, while allowing for a sufficient period of time to burn-in the population such that projections began at an appropriate population size and age composition. The period from 1958 to the present includes major changes in fishery catches, weight-at-age in the population, and population size.

To account for structural uncertainty, as with the ensemble stock assessment, four individual models are integrated into a single OM. The first model was parameterised from and conditioned to results from the long AAF stock assessment model. The second model was parameterised from and conditioned using results from the long CW stock assessment model. Because these two OM models started in 1958, they are called the medium AAF (medAAF) and medium CW (medCW) models. The two remaining models also started in 1958 and were conditioned to the same observations, but parameterised with lower values of natural mortality, as in the 2021 'short' assessment models. These two models are noted as medAAF_lowM and medCW_lowM. All four models are regional models with movement between the four biological regions.

The "lowM" models were added after SRB020 because the medAAF and medCW models alone seemed to be overly optimistic relative to short-term projections of fishing mortality compared to the ensemble stock assessment (Figure 4). The inclusion of the "lowM" models produced short-term projections from the OM that were reasonably similar to the short-term projections from the ensemble stock assessment (Figure 5).

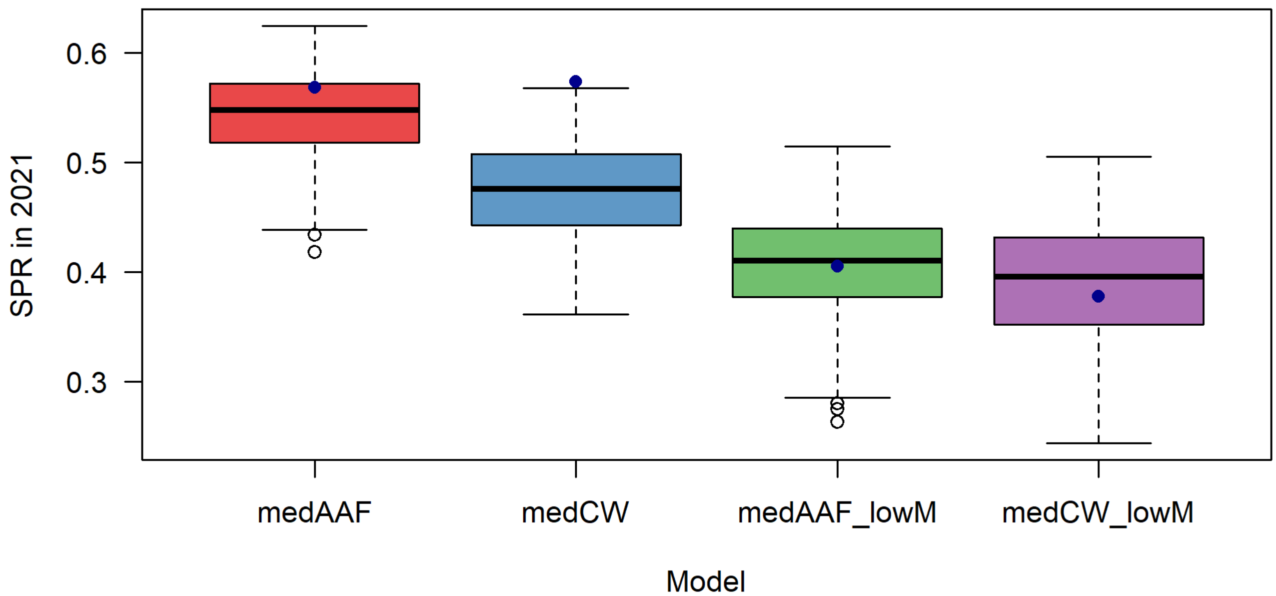

Figure 4. Estimated SPR in 2021 for each OM model (colored boxplot) compared to the estimated SPR in 2021 for each comparable individual 2021 stock assessment model (medAAF=long AAF, medCW=long CW, medAAF_lowM=short AAF, medCW_lowM=short CW).

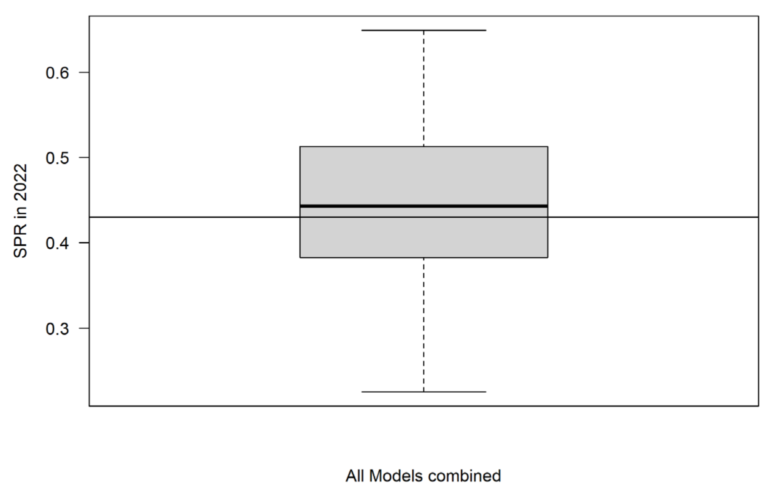

Figure 5. SPR in 2022 given fixed catches and distribution set by the Commission at the 98th IPHC Annual Meeting (IPHC-2022-AM098-R). The gray horizontal line is an SPR of 43%, corresponding to the coastwide mortality limit.

Many parameters used in the OM models were drawn from the corresponding stock assessment model. Natural mortality was fixed in each model, separately for males and females. Maturity, mean weight-at-age, recruitment deviations, the relationship between R_0 and the Pacific Decadal Oscillation (PDO), selectivity, and fishing mortality were fixed at the values from the stock assessment.

Parameters estimated during conditioning included

- R_0: initial average recruitment for the low PDO period;
- multinomial logit parameters for recruitment distribution among Biological Regions: there are 6 parameters, 3 defining the proportion among Biological Regions and 3 adjusting those parameters in high PDO years to change the distribution of age-0 recruits;
- a multiplier on initial fishing mortality: increased or decreased the initial fishing mortality input to initialize the population;
- movement from Biological Region 4 to Biological Region 3 (5 parameters) and movement from Biological Region 3 to Biological Region 2 (5 parameters), which were estimated for low PDO and high PDO periods (thus 20 total parameters).

There is considerable confounding between the recruitment distribution and movement parameters (which was evident during the conditioning process), thus some parameters for movement between Biological Regions were fixed at values estimated from previous analyses (see Figure 3 in Hicks et al 2020). The previous OM estimated considerably higher movement rates-at-age from region 2 back to region 3, which was unexpected. Fixing movement from Biological Region 2 to Biological Region 3 at values estimated directly from data resulted in more stable estimation with similar outputs.

The models were conditioned to five general sources of information:

- Historical spawning biomass estimated from the corresponding stock assessment. For example, the medAAF model was conditioned to the spawning biomass estimates from 1958 to 1992 from the 2021 long AAF stock assessment model.
- Recent ensemble spawning biomass from the corresponding spatial structure of the stock assessment. For example, the medCW model was conditioned to the spawning biomass estimates from 1993 to 2021 from the integration of the 2021 long CW and short CW stock assessment models.
- Fishery Independent Setline Survey (FISS) indices of abundance for each Biological Region.
- FISS estimates of proportions-at-age for each Biological Region. This component was downweighted compared to other components.
- Proportion of all-sizes weight-per-unit-effort (WPUE) in each Biological Region from the space-time model analysis of FISS observations. This is also called stock distribution and was given the highest weight as this is an important component for the OMs to mimic.

The conditioning was heavily weighted to the stock distribution and spawning biomass components. The goal was to have models adequately predicting stock distribution and spawning biomass in recent years, with some variability.

Even though many parameters were fixed when conditioning the models, variability was propagated from the estimated as well as some fixed parameters, accounting for correlations between parameters. Bounds were enforced on some parameters and randomly drawn parameter sets that resulted in unrealistically low population sizes or extremely poor fits to stock distribution or spawning biomass were rejected. Multiple trajectories from 1958 through 2021 were produced for each model.

2.1.2 OM results and outputs

The four individual OM models showed important structural differences in terms of movement rates-at-age, recruitment distribution, and historical spawning biomass trends. The long AAF and long CW stock assessment models, which are the basis for conditioning each OM model, estimate significantly different historical spawning biomass trajectories before the early 2000s and subtle differences in recent trajectories (Figure 6). These differences are attributable to the very different assumptions about how the stock was distributed and connected via movement in relation to historical fishing mortality, and it is important to capture these differences through movement in the OM.

The four OM models generally captured these trends in spawning biomass with the medCW models fitting the lower spawning biomass trend of the long CW assessment model and the medAAF model fitting the higher spawning biomass trend of the long AAF assessment model (Figure 7). The lowM models showed a higher probability that the spawning biomass is declining in recent years. The uncertainty in the OM also spanned the 2021 ensemble stock assessment uncertainty, except for the low spawning biomass in the 1970’s (Figure 8).

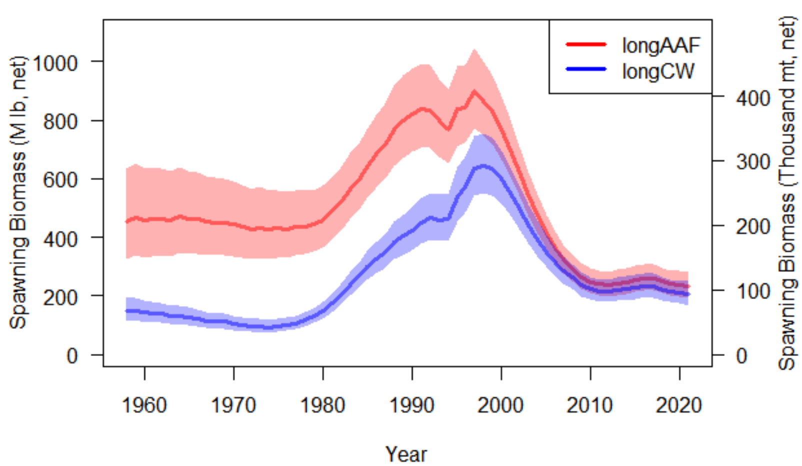

Figure 6. Estimated spawning biomass trajectories from 1958 to 2021 from the 2021 long AAF and long CW stock assessment models (Stewart & Hicks 2022).

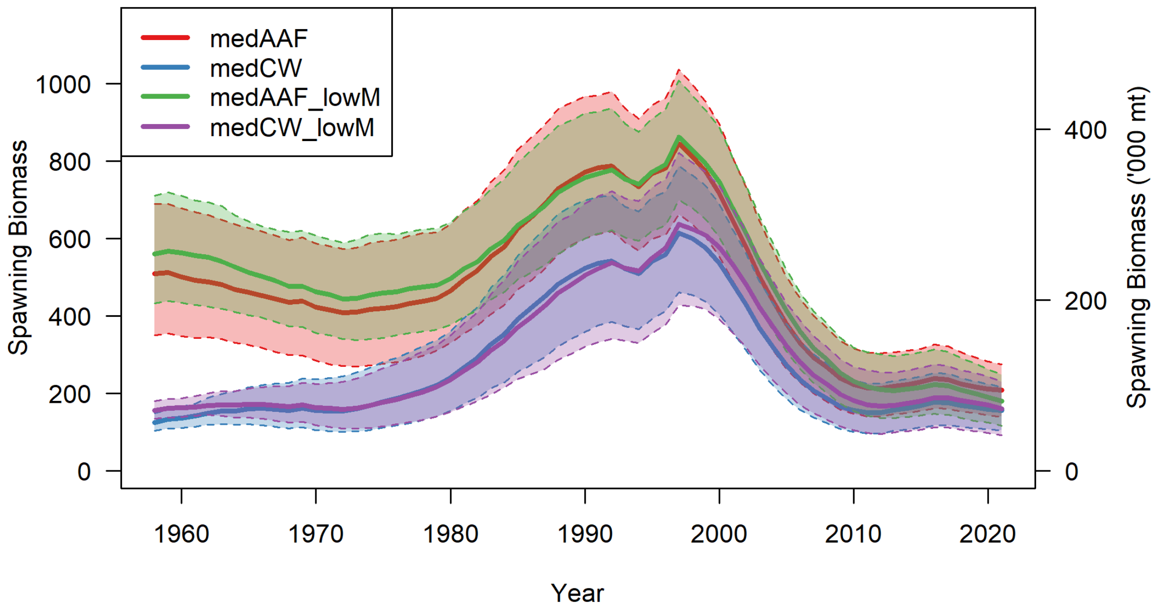

Figure 7. Median, 5th, and 95th quantiles for the four OM models.

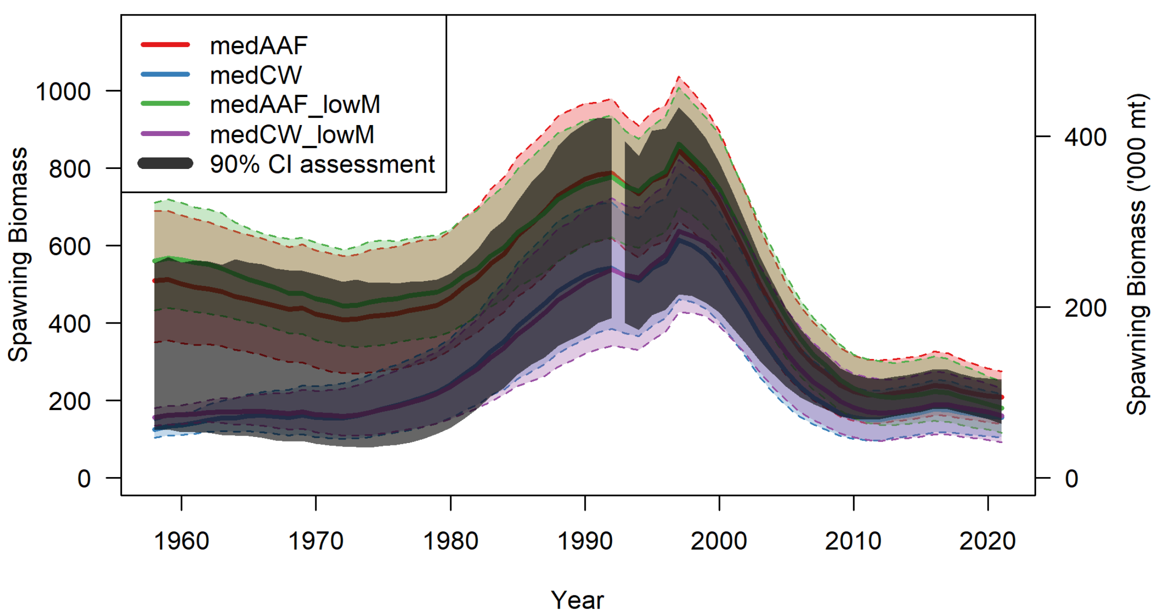

Figure 8. Median, 5th, and 95th quantiles for the four OM models with the ensemble stock assessment range between the 5th and 95th quantiles shown in grey.

Stock distribution was fit well by both OM models (Figure 9) and showed similar patterns of lack of fit across all models. Specifically, the earliest years in Biological Region 4 were overfit by the OM, and recent years overfit in Biological Region 3 corresponding with a slight underfitting in region 4. All OM models matched closely with the proportion of biomass observed in 2021.

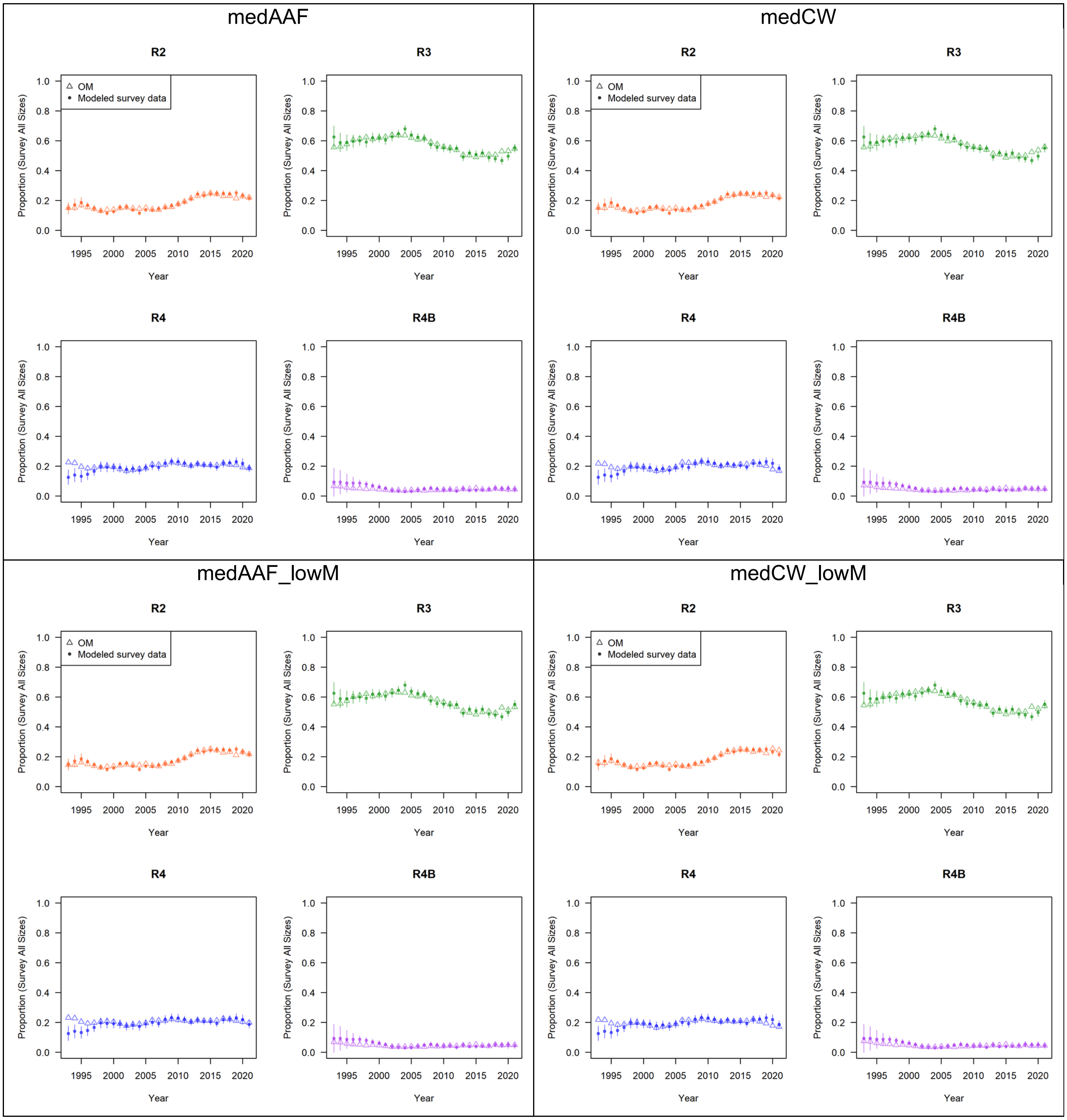

Figure 9. Fits to stock distribution across Biological Regions for each OM model.

The distribution of age-0 recruits showed a high proportion going to Biological Region 4 in both low and high PDO regimes. The medCW showed a higher proportion of recruits going to Biological Region 4 in high PDO years, but the medAAF model showed a slightly smaller proportion.

Movement rates between Biological Regions 3 and 2, and Biological Regions 4 and 3 were different between the four OM models (Figure 10). Both models generally showed high movement rates around ages 4 and 5 and slight differences between low and high PDO periods. Movement of fish younger than age 4 was very small from Biological Region 4 to 3 for both models and regimes, but there are few observations of fish younger than age 6 and a number of different movement rates of very young fish in combination with ages 4–6 could achieve similar results.

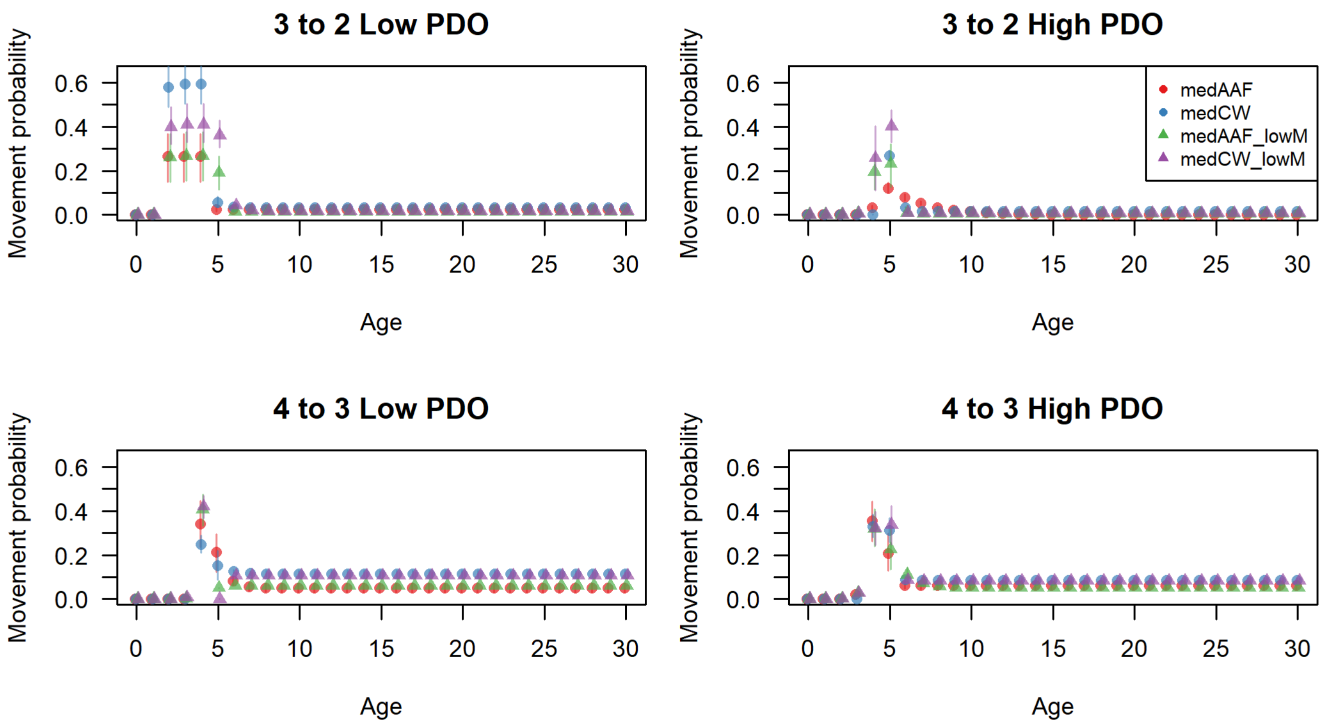

Figure 10. Probability of movement-at-age from Biological Region 3 to region 2 (top) and region 4 to region 3 (bottom) in low PDO (left) and high PDO (right) regimes for the four OM models.

2.2 Projections

The conditioned OM with multiple trajectories is the base of setting up the replicate projections of population and fishery processes. After which, they are left untouched as the closed-loop simulation projects forward in time using various management procedures (MPs) and assumptions. The simulated projection of weight-at-age, selectivity/retention deviations, and the environmental regime do not depend on the population dynamics and can be created ahead of time to save time in the simulations, although any of these processes could be dependent on the size of the population, or a certain demographic, and included in the simulation process.

Other processes, such as implementation variability, are simulated during the closed-loop simulations.

2.2.1 Implementation variability and uncertainty

Implementation variability is defined as the deviation of the fishing mortality from the mortality limit determined from an MP. It can be thought of as what actually (or is believed to have) happened compared to the limits that were set. Decision-making variability is the difference between the MP mortality limits and the adopted mortality limits set by the Commission.

Decision-making uncertainty can be applied to the mortality limit specified by the MP ($TCEY_t$) as a multiplier.

$$\widetilde{TCEY}_t = TCEY_t \varepsilon_I$$

where $\widetilde{TCEY}_t$ is the adopted mortality and ε_I is the multiplier. Using observations from 2014 to 2021 of the MP mortality limit determined from the interim management procedure and the adopted mortality limits set by the Commission for that year and IPHC Regulatory Area, the multipliers are shown in Figure 11. These years were chosen because they used a relatively consistent management procedure, although explicit use of SPR was added in 2017, additional agreements were added in 2019 and 2020, and the reference SPR changed from 46% to 43% in 2021. Decision-making uncertainty is likely different depending on the management procedure and the presence of any agreements. Additionally, in 2021 and 2022, the adopted coastwide TCEY was equal to the coastwide TCEY specified by the interim management procedure, thus distribution was the only decision-making variability.

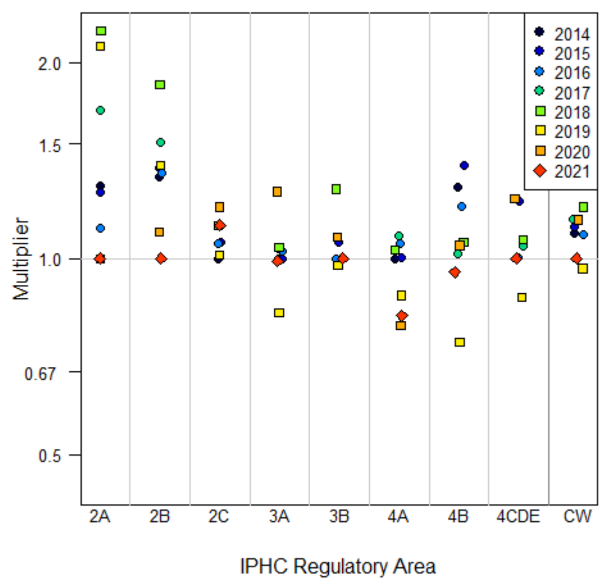

Figure 11. Multipliers for the difference between MP mortality limits and adopted mortality limits from 2014 to 2021. “CW” refers to coastwide.

2.2.1.1 Method to simulate decision-making uncertainty

The multiplier to simulate decision-making uncertainty is drawn from a lognormal distribution with correlation between multipliers for each IPHC Regulatory Area. The mean ($\boldsymbol{\mu_\varepsilon}$) and standard deviation ($\boldsymbol{\sigma_\varepsilon}$) of that distribution are modified as follows depending on the TCEY from the MP.

$$\boldsymbol{\mu_\varepsilon}\ or\ \boldsymbol{\sigma_\varepsilon} = \begin{cases} \boldsymbol{\bar{x}}\ or\ \boldsymbol{s} & TCEY < TCEY_{low} \\ \boldsymbol{a} + \boldsymbol{b} * TCEY & TCEY_{low} \leq TCEY \leq TCEY_{high} \\ \mathbf{1.0}\ or\ \boldsymbol{s}/\mathbf{2} & TCEY > TCEY_{high} \end{cases}$$

Using IPHC Regulatory Area 2A as an example (without a TCEY agreement in place), with a coastwide $TCEY_{low}$ of 30 Mlbs and a coastwide $TCEY_{high}$ equal to 60 Mlbs, the distribution of simulated multipliers gets closer to 1 as the TCEY increases (Figure 12).

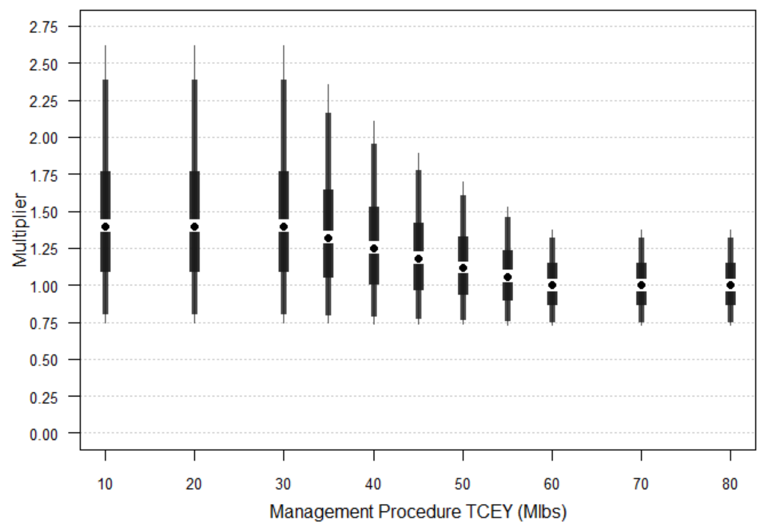

Figure 12. Simulated multipliers for IPHC Regulatory 2A at different values of the coastwide TCEY (without the recent agreement on the 2A TCEY). The thickest portion of the vertical bar represents the 25^{th} and 75^{th} percentiles, followed by the 5^{th} and 95^{th} percentiles, and then the 2.5^{th} and 97.5^{th} percentiles.

Each IPHC Regulatory Area will have a specific parameterisation to simulate decision-making variability, which will be dependent on the specific management procedure. For example, an agreement of a specific TCEY for an IPHC Regulatory Area will not have decision-making variability for that area, but other IPHC Regulatory Areas may have increased decision-making variability as a result. Furthermore, two general concepts will be used for decision-making variability:

1. The coastwide TCEY is equal to the coastwide TCEY from the MP, but distribution contains decision-making variability.
2. The coastwide TCEY may deviate from the MP, along with distribution, due to decision-making variability.

Actual decision-making variability is likely more complex than these simple methods. In fact, some IPHC Regulatory Areas show a consistent adopted TCEY over a range of MP TCEYs (e.g., 4B in Figure 13). However, the goal of including decision-making uncertainty in the MSE simulations isn't to exactly simulate what the pattern is, but to identify the effect of decision-making uncertainty and identify MPs that are robust to a plausible amount of uncertainty. Therefore, simulations will be done with and without decision-making uncertainty to identify MPs that are robust to this uncertainty and/or illustrate the benefits of reducing decision-making uncertainty. Various modifications may be made to decision-making uncertainty to explore sensitivity to various hypotheses. For example, different offsets depending on the trend in the population or TCEY, as suggested by the SRB (SRB019–Rec.06, para. 35).

2.2.1.2 Methods to simulate realized and perceived implementation uncertainty

Realized uncertainty is currently implemented in the OM by simulating a range of actual non-directed discard mortality, recreational mortality, and subsistence mortality. These are likely the largest sources of realized variability in the Pacific halibut fisheries, which is relatively small compared to many fisheries.

Perceived uncertainty is currently not simulated in the OM but will be considered as work progresses. Perceived uncertainty may include uncertainty related to sampling of catch or prohibited discarding (e.g. high-grading) that is not observed. Inclusion of perceived uncertainty in the MSE framework will likely not occur before the 99th Annual Meeting.

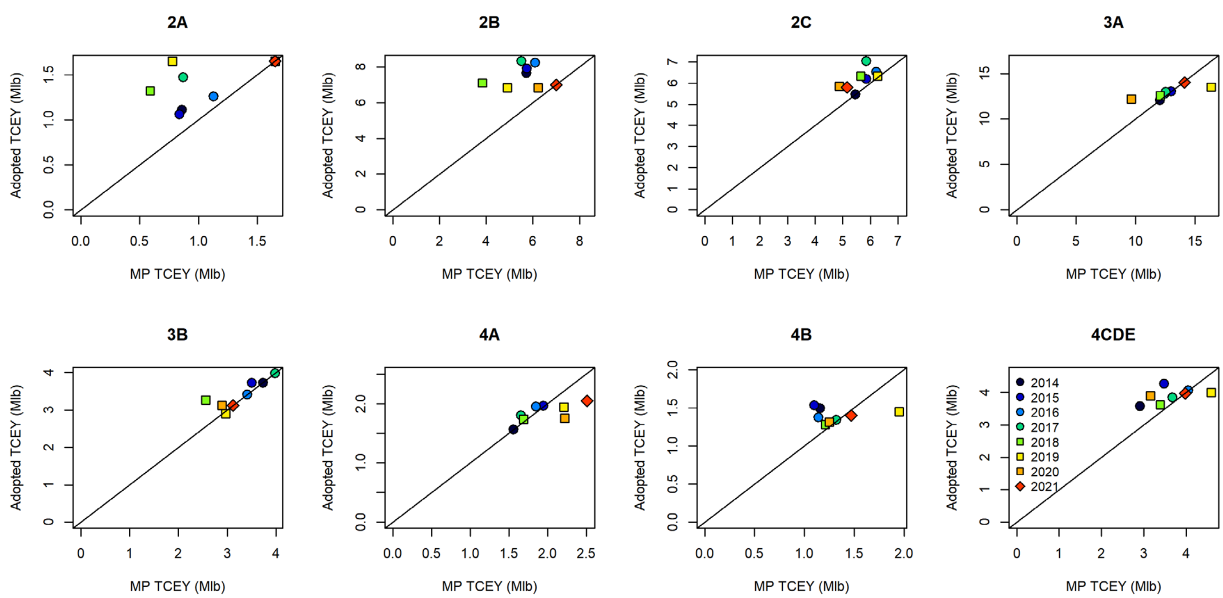

Figure 13. Adopted TCEYs plotted against MP TCEYs for each IPHC Regulatory Area and years 2014 to 2021.

2.2.2 Estimation error

Estimation error is the uncertainty in parameters that are estimated for use in a management procedure. For example, relative spawning biomass is used in the 30:20 control rule and is an estimate from the stock assessment. The total mortality given a fixed SPR is also subject to estimation error.

There are three options for examining the effect of estimation error. The first is No Estimation Error, which is useful to understand the intrinsic qualities of a management procedure. The second is Simulated Estimation Error, which simulates the correlated uncertainty in estimated relative spawning biomass and estimated total mortality. This mimics the variability that may arise from a stock assessment, but not may not capture some of the nuances of the estimates from a stock assessment, such as bias. The third is to run a stock assessment as part of the closed-loop simulation process (Simulated Stock Assessment). This can be time-consuming, especially with a complex ensemble assessment, thus simplifications are often made. Currently, a single simplified model from the Pacific halibut ensemble assessment is implemented in the MSE framework, and is useful for comparison to the simulated estimation error, but is not complete for decision-making purposes. Improvements to the simulated stock assessment method will be made in 2022 if time allows.

2.3 Runs and Scenarios

The primary closed-loop simulations consist of integrating the four OM models with equal weight by simulating an equal number of trajectories/projections from each model. Results from the full set of projections are used to calculate the performance metrics. Additional scenarios may be evaluated that include different types of implementation error or alternative scenarios of fishery selectivity (e.g. targeting or avoiding small Pacific halibut).

3 MANAGEMENT PROCEDURES

Two categories of MPs were prioritised in the MSE Program of Work for 2021–2023. One was the investigation of size limits (M.1) and the other was to investigate multi-year stock assessments (i.e. not conducting the stock assessment annually; M.3). Due to improvements in the MSE framework and changes in the OM, select MP elements investigated previously, such as SPR, may need to be re-evaluated.

3.1 Size limits

Since 1973, IPHC has restricted the directed commercial fishery for Pacific halibut with a 32 inch (81.3 cm) minimum size limit, although other forms of size limits have been in place since 1940 (Myhre 1973). Many investigations of size limits have been completed since then including IPHC (1960), Clark & Parma (1995), Parma (1999), Valero & Hare (2012), Martell et al. (2015a), Martell et al. (2015b), Stewart & Hicks (2018), and Stewart et al (2021). Most of these analyses have focused on short-term effects or effects on reference points. The novelty of this analysis using the MSE framework will be to examine long-term effects of different size limits in relation to defined conservation and fishery objectives. Additionally, long-term changes to the stock and fishery distribution as well as changes in productivity will be examined.

The Commission requested that three size limits be investigated: 32 inches, 26 inches, and no size limit.

> **IPHC-2022-AM098-R, para. 61**: *The Commission RECALLED SS011-Rec.01 and REQUESTED that the current size limit (32 inches), a 26 inch size limit, and no size limit be investigated. to understand the long-term effects of a change in the size limit.*

It is uncertain how selectivity of the directed commercial fisheries may change with the implementation of a different size limit than the current 32 inches. Fisheries may choose to target smaller fish to increase efficiency, they may maintain current practices, or they may target larger fish if that provides improved economic gains. Some sensitivities to changes in selectivity (e.g. alternative scenarios) may be investigated.

An important concept to bring into the evaluation of size limits is market considerations. Stewart et al. (2021) used the ratio between the U32 price and O32 price for Pacific halibut to determine what ratio is necessary for the fishery to break even economically. It is unknown what prices will be for U32 Pacific halibut if a size limit was removed, but the FISS has recently begun selling U32 fish, which may be an indicator for future market conditions of small fish. Regardless, a performance metric related to economics will be important to consider in this evaluation.

3.2 Multi-year assessments

Management procedures with multi-year assessments incorporate a process where the stock assessment occurs at intervals longer than annually. The mortality limits in a year with the stock assessment can be determined as in previously defined MPs, but in years without a stock assessment, the mortality limits would need an alternative approach. This may be as simple as maintaining the same mortality limits for each IPHC Regulatory Area in years with no stock assessment, or as complicated as invoking an alternative MP that does not require a stock assessment (such as an empirical-based MP relying only on data/observations).

The Commission requested that the Secretariat investigate biennial assessments and potentially longer intervals as time allows.

> **IPHC-2022-AM098-R, para 64**: *The Commission REQUESTED that multi-year management procedures include the following concepts:*
>
> *a) The stock assessment occurs biennially (and possibly triennial if time in 2022 allows) and no changes would occur to the FISS (i.e. remains annual);*
>
> *b) The TCEY within IPHC Regulatory Areas for non-assessment years:*
>
> > *i. remains the same as defined in the previous assessment year, or*
> >
> > *ii. changes within IPHC Regulatory Areas using simple empirical rules, to be developed by the IPHC Secretariat, that incorporate FISS data.*

There are many different empirical rules that could be applied to determine the TCEY in non-assessment years. We identified three empirical rules for determining IPHC Regulatory Area specific TCEYs in non-assessment years, which either use no observations or FISS observations .

a. The same TCEY from the previous year for each IPHC Regulatory Area.
b. Updating the coastwide TCEY proportionally to the change in the coastwide FISS O32 WPUE and updating the distribution of the TCEY using FISS results and the applied distribution procedure.
c. Maintaining the same coastwide TCEY as the previous year but updating the distribution of the TCEY using FISS results and the applied distribution procedure.

Empirical rule (a) does not update the TCEY in Regulatory Areas, which may deviate from distributions agreements related to a percentage of the coastwide TCEY, if present, due to changes in the distribution of biomass. Empirical rules (b) and (c) both adjust the distribution of the coastwide TCEY and would maintain any agreements related to distribution.

The coastwide TCEY set in the assessment year also can be calculated using different methods. The coastwide TCEY may simply be determined from the one-year projection of the stock assessment without any consideration of the projections beyond one year. This is the method assumed in the above empirical rules. An alternative method would be to take an average of the coastwide TCEYs, given a defined fishing intensity, projected for all years before the next assessment. This would account for potential changes in the population and may maintain the stock closer to target biomass levels and the fishing intensity closer to reference SPR levels. Alternative methods of averaging projected TCEYs were not considered.

An alternative approach that would not require a stock assessment for setting mortality limits in any year would be to adopt an empirical-based MP as the method for setting annual mortality limits. The stock assessment would be used at a defined interval to verify that management is effective and to potentially tune the MSE OM and existing MP (Cox and Kronlund 2008).

The Commission has realized that there are some benefits to multi-year assessments, including stability and transparency in mortality limits for multiple rather than single years, additional time during the Interim/Annual meeting process to focus on topics other than setting mortality limits, time for development/improvement of the stock assessment, and the potential for increased collaborative research across branches within the IPHC Secretariat. However, there may be some costs associated with multi-year assessments. For example, performance in meeting conservation and fishery objectives may be reduced depending on the interval for multi-year assessments and the specifics of the selected management procedure.

The Commission has asked the SRB to assist the Secretariat in identifying potential costs and benefits of not conducting an annual stock assessment.

> **IPHC-2022-AM098-R, para 63:** *The Commission REQUESTED that the IPHC Secretariat work with the SRB and others as necessary to identify potential costs and benefits of not conducting an annual stock assessment. This will include a prioritized list of work items that could be accomplished in its place.*

The SRB provided some insight at SRB020 and the Secretariat will continue to work with the SRB in identifying costs and benefits.

> IPHC-2022-SRB020-R, para 27. *The SRB NOTED that assessment research activities (e.g. paras. 23-26) are examples of work that could be done more extensively in non-assessment years within a multi-year assessment schedule. Other work could include investigating optimal sub-sampling designs for ages, sex-ratio, annual assessment methods to use within the MPs, and well as any of the several topics listed under Stock Assessment Research. The quantifiable costs of multi-year assessments could be estimated within the MSE, for example, of potentially lower average yield for longer assessment cycles to achieve the same levels of risk associated with annual assessments.*

It may be premature to begin identifying detailed costs and benefits of multi-year assessments until an evaluation has been done to determine whether multi-year assessments may meet the Commission objectives already defined. An evaluation of multi-year assessments using Commission conservation and fishery objectives will be presented at the 99th IPHC Annual Meeting, after which a discussion of detailed costs and benefits would be informative.

3.3 Modelling distribution

The fisheries in the OM are specified by IPHC Regulatory Area because many of the Commission objectives used to evaluate MPs are specific to IPHC Regulatory Areas and the OM is spatially structured by Biological Region. This makes it necessary to distribute the TCEY across the fisheries to appropriately remove biomass from each Biological Region and allow for the calculation of necessary performance metrics. Distribution procedures have been evaluated (Hicks et al. 2021), but a specific MP has not been implemented. Even though distribution procedures are not currently being evaluated and there is no specific agreement on a single distribution procedure, they are part of the MP and need to be included in the simulations. Therefore, the Commission has recommended five different distribution procedures representing a practicable range to provide a robust analysis of size limits and multi-year assessments.

> **IPHC-2022-SS012-R, para 11**: *The Commission RECOMMENDED the following five distribution procedures to be used in the management strategy evaluation of size limits and multi-year assessments, noting that these distribution procedures are for analytical purposes only and are not endorsed by both parties, thus would be reviewed in the future if the Commission wishes to evaluate them for implementation.*
>
> > *a) Baseline based on recent year O32 FISS results, relative harvest rates of 1.0 for IPHC Regulatory Areas 2-3A, relative harvest rates of 0.75 for IPHC Regulatory Areas 3B-4, and no application of the current interim agreements for 2A and 2B;*
> >
> > *b) Baseline based on recent year O32 FISS results, relative harvest rates of 1.0 for IPHC Regulatory Areas 2-3A, relative harvest rates of 0.75 for IPHC Regulatory Areas 3B-4, and current interim agreements for 2A and 2B;*

c) Baseline based on recent year O32 FISS results with 1.65 Mlbs to 2A and 20% of the coastwide TCEY to 2B;

d) Baseline based on recent year O32 FISS results, relative harvest rates of 1.0 for IPHC Regulatory Areas 2-3, 4A, and 4CDE, a relative harvest rate of 0.75 for IPHC Regulatory Area 4B, and no agreements for 2A and 2B;

e) Baseline based on recent year O32 FISS results, relative harvest rates of 1.0 for IPHC Regulatory Areas 2-3, 4A, and 4CDE, a relative harvest rate of 0.75 for IPHC Regulatory Area 4B, and current interim agreements for IPHC Regulatory Areas 2A and 2B

Three of the five distribution procedures contain agreements for IPHC Regulatory Areas 2A and 2B (b, c, and e). Implementation variability for these two areas is set to zero when agreements are in place.

3.4 MP combinations

It is easy in any MSE to specify a large set of runs due to the combination of many MP elements. Given that the simulation time for a single MP may be days, it is useful to identify a small set of runs that will provide insight into the performance of each element of the MP of interest. The components of size limits and multi-year assessments presented above have multiple elements that are combined as shown in Table 2. For each MP, an SPR of 43% was used, with some specific combinations using SPR values of 40% and 46%.

Table 2. Primary MPs to be evaluated. The multi-year assessment specifies the frequency of the stock assessment and the procedure for years without a stock assessment (see Section 3.2).

MP ID	Multi-year assessment	Size Limit (inches)
MP-A32	Annual	32
MP-Ba32	Biennial, constant TCEY	32
MP-Bb32	Biennial, empirical rule	32
MP-Bc32	Biennial, update distribution	32
MP-A26	Annual	26
MP-A0	Annual	0

A secondary set of MPs will be developed based on the performance of the primary set. This may include crossing size limits with biennial assessments, tuning SPR values to best meet objectives, examining different levels of estimation error, and incorporating various forms of implementation variability. This secondary set will not be a full factorial, but instead a specific investigation of relevant factors, and to refine the best performing MPs relative to stock and fishery objectives.

Furthermore, a set of sensitivities will be done using alternative scenarios as described above. These will be performed on a small set of the best performing MPs.

4 RESULTS AND EVALUATION

The MPs were integrated across the distribution procedures, resulting in the six MPs in Table 2 as distribution is considered an uncertainty in this evaluation. However, any interesting differences between distribution procedures may be reported.

Improvement of the methods to evaluate simulation results and present those for decision-making are ongoing. Current tasks specifically include updates to the MSE Explorer tool, improving the ranking procedure to identify best performing management procedures, determining new methods to identify best performing management procedures, and providing new types of plots and tables that effectively communicate the results. This task will benefit from interactions with stakeholders and management agencies, which may include MSAB meetings.

4.1 Projections

The improvements to the MSE framework, including the updated OM, resulted in some different outcomes, although general conclusions were consistent with previous analyses. The additional years at the end of the historical time-series in the OM resulted in immediate optimistic trends in the spawning biomass (Figure 14) due to a possibly large 2012 year class, a positive PDO regime, and increasing trends in weight-at-age.

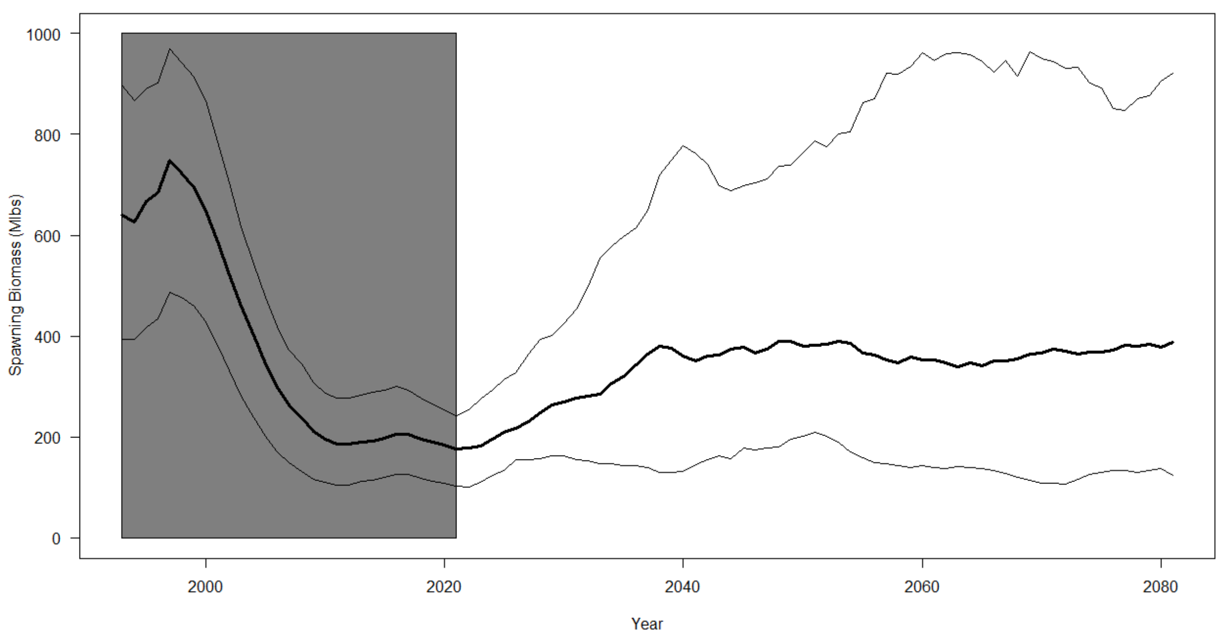

Figure 14. Projected spawning biomass with MP-A32, an SPR of 43%, and no estimation error. The shaded area is the historical region with fixed data and fishing mortality. The thick line is the median and the thin lines are the 5th and 95th percentiles.

4.2 Size limits

Applying the three size limits resulted in little change to the biological sustainability performance metrics with or without simulated estimation error (Table 3). Simulated estimation error resulted in a lower average fishing intensity (i.e. higher SPR) but a slightly lower average relative spawning biomass. The lower portion of the distribution of average relative spawning biomass was more compact than without estimation error as shown by the lower probability of being less than 36%. The upper portion of the distribution of average RSB was wider with estimation error (Figure 15).

Table 3. Performance metrics related to primary objectives for size limit MPs with no decision-making variability. The same MPs are simulated with no estimation error or simulated estimation error. Biological sustainability metrics are long-term and fishery sustainability are short-term (4–13 years).

MP name	**MP-A0**	**MP-A26**	**MP-A32**	**MP-A0**	**MP-A26**	**MP-A32**
Decision-making variability	**None**	**None**	**None**	**None**	**None**	**None**
Estimation Error	**None**	**None**	**None**	**Sim**	**Sim**	**Sim**
Assessment Frequency	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**
Size Limit	**0**	**26**	**32**	**0**	**26**	**32**
SPR	**0.43**	**0.43**	**0.43**	**0.43**	**0.43**	**0.43**
Median average SPR	43.00%	43.00%	43.00%	43.90%	43.90%	44.00%
Biological Sustainability						
Median average RSB	39.3%	39.3%	39.3%	39.0%	39.0%	39.0%
P(any RSB_y<20%)	0	0	0	0	0	0
P(all RSB<36%)	0.17	0.17	0.18	0.14	0.14	0.14
Fishery Sustainability						
Median average TCEY	62.26	62.08	58.92	60.18	59.69	58.09
P(any3 change TCEY > 15%)	0.058	0.058	0.072	0.934	0.946	0.966
Median AAV TCEY	5.2%	5.3%	5.7%	18.2%	18.3%	18.7%

Short-term fishery sustainability performance metrics showed some improvements when lowering the size limit (Table 3). The TCEY, on average, was 5.4% higher with a 26-inch size limit and 5.7% higher with no size limit. With simulated estimation error the average TCEY was less, and increases to the TCEY with a 26-inch size limit and no size limit were 2.8% and 3.6%, respectively. The percentage gain in the TCEY is variable across years and is higher in the short-term given starting conditions of the projections (Figure 16), and there is a very small probability that the TCEY is less without a size limit. The high percent gain in recent projected years is due to starting conditions, which declines as recruitment, weight-at-age, and environmental regimes become more integrated across the range of possible values. Annual variability in the TCEY was slightly reduced with lower size limits but above 15% with estimation error (Table 3).

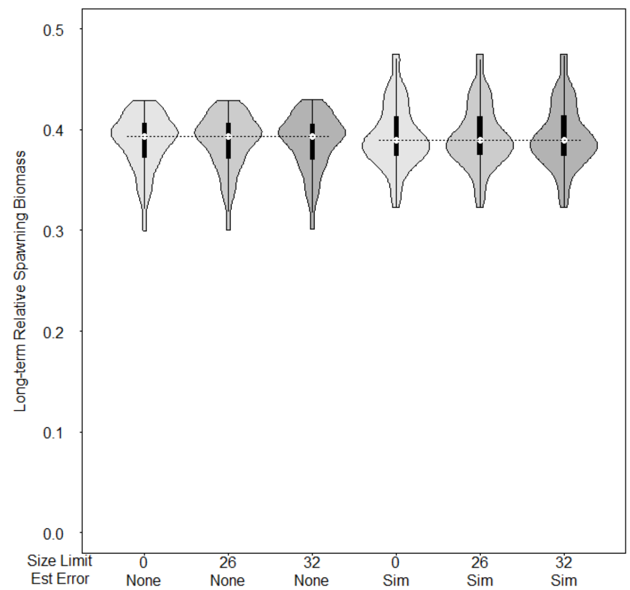

Figure 15. Violin plots of long-term relative spawning biomass for the three size limits (different shades of grey) and no estimation error (left) or simulated estimation error (right). A dashed line is drawn at the median for the 32 inch size limit of each estimation error type.

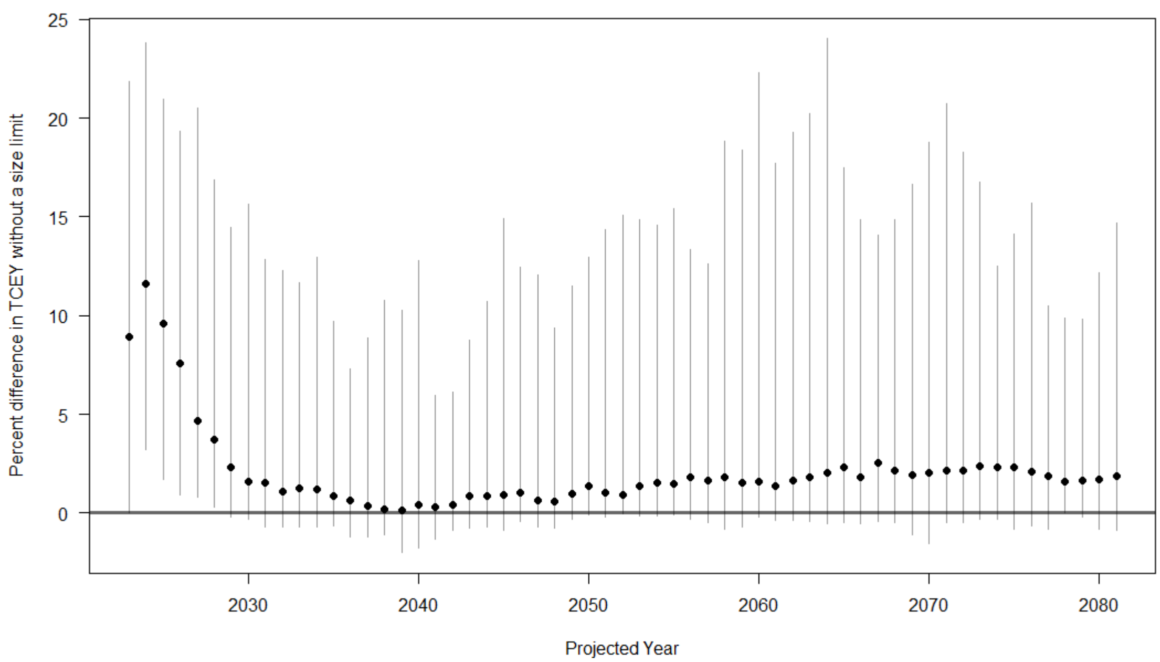

Figure 16. Percent difference in the TCEY without a size limit compared to a 32-inch size limit for each projected year when simulating estimation error. The points are the median and the vertical lines connect the 5th and 95th percentiles.

The coastwide TCEY differed between the short-term and the long-term (Figure 17). The median coastwide TCEY was higher and differences between size limits were less pronounced in the long-term (as also shown in Figure 16). Estimation error also had a greater effect on the range of TCEY in the long-term.

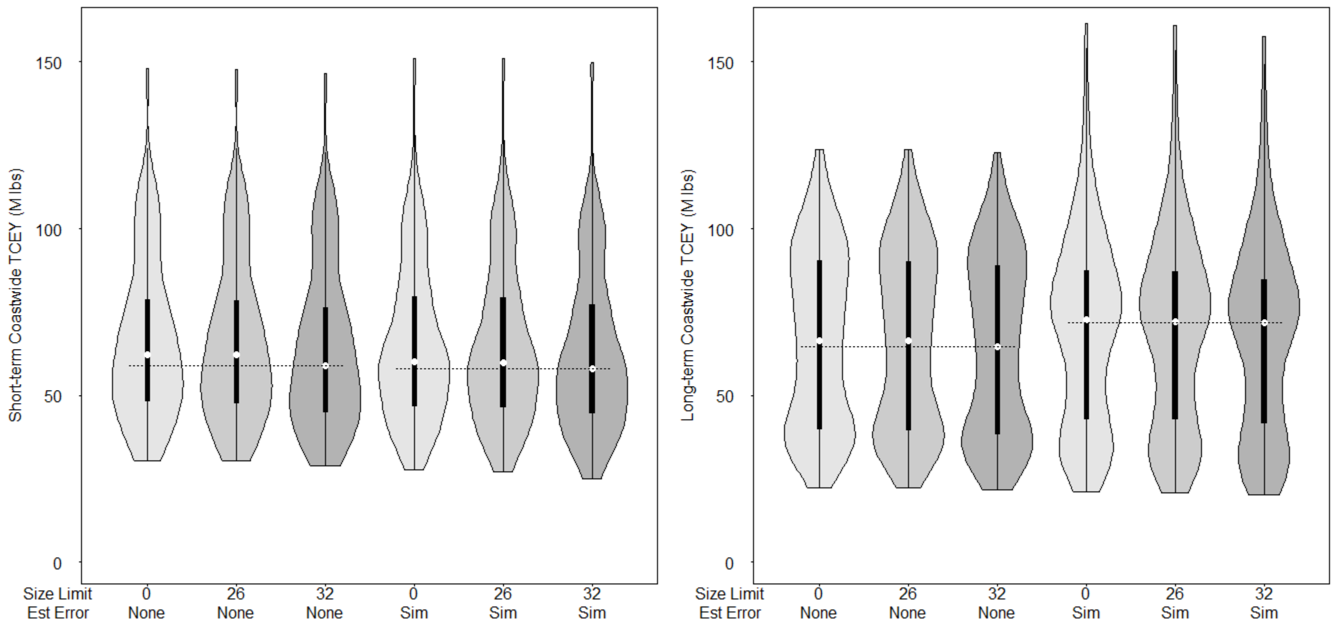

Figure 17. Short-term coastwide TCEY (left) and long-term coastwide TCEY (right) for the three size limits and no or simulated estimation error.

The patterns were similar for performance metrics calculated for each IPHC Regulatory Area (Table 4). The median average TCEY (with simulated estimation error) in the IPHC Regulatory Areas increased between 4.5% and 5.7% except for IPHC Regulatory Area 2A (no change since three of the five distribution procedures had a fixed 1.65 Mlbs) and IPHC Regulatory Area 4B (6.9%). Even though the TCEY in IPHC Regulatory Area 3A showed a modest percent increase without a size limit, the absolute increase in the TCEY was over 1 million pounds. Annual variability in the TCEY for each IPHC Regulatory Area decreased when lowering the size limit, but remained above 14.5% when simulating estimation error.

The majority of the gain in median average TCEY and the reduction in annual variability of the TCEY was achieved when lowering the size limit from 32 inches to 26 inches. This is because the directed commercial gear has a low selectivity for Pacific halibut less than 26 inches.

Table 4. Performance metrics related to area-specific primary objectives for size limit MPs with no decision-making variability. Fishery sustainability metrics are short-term (4–13 years).

MP name	**MP-A0**	**MP-A26**	**MP-A32**	**MP-A0**	**MP-A26**	**MP-A32**
Decision-making variability	**None**	**None**	**None**	**None**	**None**	**None**
Estimation Error	**None**	**None**	**None**	**Sim**	**Sim**	**Sim**
Assessment Frequency	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**
Size Limit	**0**	**26**	**32**	**0**	**26**	**32**
SPR	**0.43**	**0.43**	**0.43**	**0.43**	**0.43**	**0.43**
Median average TCEY-2A	1.65	1.65	1.65	1.61	1.61	1.61
Median average TCEY-2B	9.14	9.09	8.72	8.97	8.9	8.58
Median average TCEY-2C	6.82	6.77	6.55	6.7	6.67	6.41
Median average TCEY-3A	24.7	24.59	23.6	24.57	24.36	23.36
Median average TCEY-3B	7.75	7.70	7.46	7.47	7.42	7.09
Median average TCEY-4A	3.72	3.69	3.56	3.74	3.70	3.54
Median average TCEY-4CDE	5.11	5.06	4.89	4.18	4.12	3.99
Median average TCEY-4B	2.47	2.42	2.33	2.93	2.87	2.74
P(any3 change TCEY 2A > 15%)	0.012	0.012	0.012	0.302	0.310	0.336
P(any3 change TCEY 2B > 15%)	0.030	0.030	0.032	0.728	0.738	0.786
P(any3 change TCEY 2C > 15%)	0.040	0.044	0.042	0.762	0.766	0.810
P(any3 change TCEY 3A > 15%)	0.030	0.030	0.036	0.734	0.748	0.790
P(any3 change TCEY 3B > 15%)	0.022	0.020	0.022	0.734	0.746	0.790
P(any3 change TCEY 4A > 15%)	0.034	0.036	0.042	0.818	0.828	0.852
P(any3 change TCEY 4CDE > 15%)	0.006	0.006	0.016	0.580	0.574	0.568
P(any3 change TCEY 4B > 15%)	0.036	0.032	0.03	0.826	0.82	0.848
Median AAV TCEY 2A	0.0%	0.0%	0.0%	3.1%	3.1%	3.4%
Median AAV TCEY 2B	5.7%	5.8%	6.3%	18.3%	18.7%	19.1%
Median AAV TCEY 2C	6.5%	6.6%	7.0%	19.3%	19.5%	20.2%
Median AAV TCEY 3A	5.9%	6.0%	6.4%	19.1%	19.4%	19.6%
Median AAV TCEY 3B	5.9%	6.0%	6.4%	19.0%	19.3%	19.5%
Median AAV TCEY 4A	6.2%	6.2%	6.5%	19.3%	19.7%	20.4%
Median AAV TCEY 4CDE	6.1%	6.2%	6.3%	14.5%	14.6%	14.9%
Median AAV TCEY 4B	6.0%	6.0%	6.4%	19.9%	20.1%	20.6%

4.3 Multi-year assessments

Simulations without estimation error of a biennial assessment frequency using option c (constant coastwide TCEY in non-assessment years but updated distribution) showed very small differences in long-term biological sustainability metrics when compared to an annual assessment frequency (Table 5). Short-term fishery sustainability metrics showed a slightly smaller median TCEY with a biennial assessment frequency. The annual variability of the TCEY was much greater with biennial assessments, even though the coastwide TCEY changed in only 5 of the 10 years used to calculate the metric. This suggests that the TCEY had to make large changes to account for the constant TCEY over two-years. There are no current objectives that would indicate whether a stable 2-year period with a larger biennial change is preferable to annual changes in the TCEY.

Table 5. Performance metrics related to primary objectives for annual and biennial MPs with a size limit of 32 inches and no estimation error or decision-making variability. The biennial MP uses option c. Biological sustainability metrics are long-term and fishery sustainability are short-term (4–13 years).

MP name	**MP-A32**	**MP-Bc32**
Decision-making variability	**None**	**None**
Estimation Error	**None**	**None**
Assessment Frequency	**Annual**	**Biennial**
Size Limit	**32**	**32**
SPR	**0.43**	**0.43**
Median average SPR	43.0%	43.3%
Biological Sustainability		
Median average RSB	39.3%	38.9%
P(any RSB_y<20%)	0.00	0.00
P(all RSB<36%)	0.18	0.17
Fishery Sustainability		
Median average TCEY	58.92	57.53
P(any3 change TCEY > 15%)	0.072	0.784
Median AAV TCEY	5.7%	14.7%

4.4 Additional results anticipated for the 99th IPHC Annual Meeting

Many more results and comparisons will be provided at the 99th IPHC Annual Meeting. Implementation variability and estimation error will be simulated and contrasted to runs without these sources of variability. Additional performance metrics will also be examined, including the age/size composition of landings, the amount of fish discarded and discard mortality in the directed commercial fisheries, and other sector-specific metrics.

RECOMMENDATION/S

That the SRB

a) **NOTE** paper IPHC-2022-SRB021-07 describing improvements to the closed-loop simulation framework, two types of management procedures to simulate and evaluate in 2022, and preliminary results from different size limits.

b) **IDENTIFY** costs and benefits associated with multi-year assessments, including whether multi-year assessments meet the Commission's primary objectives.

c) **RECOMMEND** any changes, additional MPs, or evaluation to be presented at IM098.

d) **RECOMMEND** additional improvements or additional MSE tasks to be done in 2023.

REFERENCES

Clark, W.G. and A.M. Parma. 1995. Re-evaluation of the 32-inch commercial size limit. International Pacific Halibut Commission Technical Report No. 33. 34 p.

Cox, Sean P., and Allen Robert Kronlund. 2008. "Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia, Canada." *Fisheries Research* 94 (3): 224-237. https://doi.org/10.1016/j.fishres.2008.05.006.

Hicks, A, P Carpi, S Berukoff, and I Stewart. 2020a. An update of the IPHC Management Strategy Evaluation process for SRB016. https://iphc.int/uploads/pdf/srb/srb016/iphc-2020-srb016-08.pdf

Hicks, A, P Carpi, S Berukoff, and I Stewart. 2020b. *An update of the IPHC Management Strategy Evaluation process for SRB017.* https://iphc.int/uploads/pdf/srb/srb017/iphc-2020-srb017-09.pdf.

Hicks, A, P Carpi, I Stewart, and S Berukoff. 2021. *IPHC Management Strategy Evaluation for Pacific halibut (Hippoglossus stenolepis).* https://iphc.int/uploads/pdf/am/am097/iphc-2021-am097-11.pdf.

Martell, S., B. Leaman, and I. Stewart. 2015a. Recent developments in the IPHC Management Strategy Evaluation process and size-limit implications. IPHC Report of Assessment and Research Activities 2014. p. 299-312.

Martell, S., I. Stewart, and J. Sullivan. 2015b. Implications of bycatch, discards, and size limits on reference points in the Pacific halibut fishery. In Fisheries bycatch: Global issues and creative solutions. Edited by G.H. Kruse and H.C. An and J. DiCosimo and C.A. Eischens and G. Gislason, S. and D.N. McBride and C.S. Rose and C.E. Siddon. Alaska Sea Grant, University of Alaska Fairbanks.

Myhre, R.J. 1973. Setting the new halibut size limit. Western Fisheries. 85(5): 14IPHC. 1960. Utilization of Pacific halibut stocks: yield per recruitment. IPHC Sci. Rep. No. 28. 52 p.

Parma, A.M. 1999. Effects of imposing a maximum size limit in commercial landings. International Pacific Halibut Commission Annual Meeting Handout.

Stewart, I. and A. Hicks 2018. Evaluation of the IPHC's 32" minimum size limit. IPHC-2018-AM094-14. 1 December 2017. https://www.iphc.int/uploads/pdf/am/2018am/iphc-2018-am094-14.pdf

Stewart, I., and A. Hicks. 2022. Assessment of the Pacific halibut (Hippoglossus stenolepis) stock at the end of 2021. IPHC-2022-SA-01. 16 December 2022. https://www.iphc.int/uploads/pdf/sa/2022/iphc-2022-sa-01.pdf

Stewart, I., A. Hicks, B. Hutniczak. 2021. Evaluation of directed commercial fishery size limits in 2020. IPHC-2021-AM097-09. 15 December 2020. https://www.iphc.int/uploads/pdf/am/am097/iphc-2021-am097-09.pdf

Valero, J.L., and S.R. Hare. 2012. Harvest policy considerations for re-evaluating the minimum size limit in the Pacific halibut commercial fishery. 2012 IPHC annual meeting handout. p. 22-58.

APPENDICES

Appendix A: Supplementary material

APPENDICES

Appendix A: Supplementary material

In addition to this document, an MSE technical document is available electronically. This is document IPHC-2022-MSE-01 and is available on the IPHC MSE page (https://www.iphc.int/management/science-and-research/management-strategy-evaluation).

The MSE Explorer will also be updated with additional results.

(http://shiny.westus.cloudapp.azure.com/shiny/sample-apps/MSE-Explorer/).