INTERNATIONAL PACIFIC HALIBUT COMMISSION

IPHC-2024-SRB025-06

Development of the 2024 Pacific halibut (*Hippoglossus stenolepis*) stock assessment

PREPARED BY: IPHC SECRETARIAT (I. STEWART & A. HICKS; 20 AUGUST 2024)

PURPOSE

To provide the IPHC's Scientific Review Board (SRB) with a response to recommendations and requests made during SRB024 (IPHC-2024-SRB024-R) and to provide the Commission with an update on progress toward the 2024 stock assessment.

INTRODUCTION

The International Pacific Halibut Commission (IPHC) conducts an annual coastwide stock assessment of Pacific halibut (*Hippoglossus stenolepis*). The most recent full assessment was completed in 2022 (IPHC-2023-SA01), following updates in 2020 and 2021. The 2023 stock assessment updated the 2022 analysis and all data sources where new information was available but made no structural changes to the methods. Development and supporting analyses arising from the 2023 assessment were reviewed by the IPHC's SRB in June (SRB022; IPHC-2023-SRB022-08, IPHC-2023-SRB022-R) and September 2023 (SRB023; IPHC-2023-SRB023-06, IPHC-2023-SRB023-R).

A summary of the 2023 stock assessment results (IPHC-2024-AM100-10) as well as stock projections and the harvest decision table for 2024 (IPHC-2024-AM100-12) were provided for the IPHC's 100th Annual Meeting (AM100). In addition, the input data files are archived each year on the stock assessment page of the IPHC's website, along with the full assessment (IPHC-2024-SA-01) and data overview (IPHC-2024-SA-02) documents. All previous stock assessments dating back to 1978 are also available at that location.

In June 2024, the Secretariat produced a summary of stock assessment development to date (IPHC-2024-SRB024-08). That preliminary development included extending the time-series and updating to the newest version of the Stock Synthesis software, neither of which affected the model results. Development also included an improvement on the parameterization of selectivity, allow for uncertainty in the random-walk process to be propagated into forward projects; this had very small effects on model projections.

This document includes a response to requests made during SRB023 and SRB024, including the results of FISS design simulation experiments, and an overview of topics planned for exploration in the 2025 full stock assessment. The final 2024 analysis will be an updated stock assessment, consistent with the schedule for conducting a full assessment and review approximately every three (3) years. Standard data sources and model configurations are expected to remain unchanged.

SRB REQUESTS AND RECOMMENDATIONS

The SRB made the following assessment recommendations and requests during SRB023 and SRB024:

1) SRB023–Rec.19 (para. 59):

   *"The SRB **RECOMMENDED** that the Secretariat continue exploring ways of estimating the impacts of different FISS designs and efficiency decisions on stock assessment outputs and*

fishery performance objectives. The end goal should be to provide a decision support tool that can frame decisions about FISS design in terms of costs and benefits in comparable currencies."

2) SRB023–Req.07 (para. 60):

   *The SRB **REQUESTED** that the Commission NOTE that some longer-term (2025 and beyond) implications of reduced FISS designs are predictable and potentially consequential. For instance, higher FISS CVs will generally result in higher inter-annual variation in TCEY under the current decision-making process. This would occur for two reasons: (1) biomass estimates and projections from the assessment model will have greater uncertainty and therefore greater variability in outputs and (2) ad hoc management adjustments to the interim harvest policy recommendations would be more frequent and/or more variable for greater input uncertainty. The SRB therefore REQUESTED the following analyses for SRB024:*

   a) Assessment of reduced FISS designs (2025-2027) via simulation tests of assessment model outputs (e.g. probability of decline, estimated stock abundance and status, TCEY) under alternative revenue-neutral FISS designs using the existing stock assessment ensemble;

   b) Mitigation options of reduced FISS designs (short-term and long-term) via MSE simulations of management procedures that deliberately aim to reduce inter-annual variability in TCEY via multi-year TCEYs and (possibly) fixed stock distribution schemes;

   c) Components (a,b) above would be integrated since (a) will need to inform simulations in (b)."

3) SRB024 (para. 42):

   *The SRB **RECOMMENDED** that the Secretariat investigate:*

   a) Fitting a power function to the AI/CNN vs manual age determination to show how bias increases with age;

   b) Training the model with more otoliths from older age classes;

   c) Alternative objective functions that put more weight on correctly estimating ages of older individuals;

   d) The importance of different aspects of aging accuracy/bias on the stock assessment.

Recommendations 1 &2 – Simulation testing FISS designs

Results of a stock assessment simulation 'self test' along with a proposal for FISS design simulation experiments were presented during SRB024. Following that basic test of the stock assessment ensemble performance, simulation experiments were developed to compare the effects of three potential FISS designs implemented over the period 2025-2027 on the stock assessment and management results:

1) A 'base block design' including good spatial coverage (at least one charter region in all Biological Regions and all IPHC Regulatory Areas each year), low CVs and very low potential for multi-year bias due to sampling all survey stations on a frequent basis.
2) A 'core design' including sampling in those areas with the highest biomass at a reduced sampling cost. This design will produce larger CVs than the block design and will have a

high likelihood of biased trends and age compositions due to low abundance and/or high-cost areas going unsampled for multiple consecutive years.

3) A ‘reduced core design’ that provides sampling only in areas that are close to or above revenue positive thresholds. This design will produce larger CVs than the core design and will have a very high likelihood of introducing biased trends and age compositions due to the extremely restricted geographic coverage.

For each of these designs, the annual index variance was calculated for 2025 through 2027 (see IPHC-2024-SRB024-06). Projections using the space-time model naturally propagate the variance associated with reduced FISS designs; however, because the reduced designs do not represent a random draw from all 1,890 survey stations there is the potential for bias in addition to reduced precision. The degree of potential bias is unknown and will depend on how the design interacts with localized trends and patterns in cohort structure, movement rates, and other factors known to vary interannually. Based a summary of previous changes in different areas of the stock, the Secretariat used +/- 15% bias in the FISS index over 3 years as a basis for investigating short-term stock assessment performance.

The current stock assessment can be used to simulate new data, given an assumed trend and precision for all data sources. This is achieved via the internal semi-parametric bootstrap used in the ‘self-test’ presented at SRB023. This same approach was applied for the FISS design simulations:

1) Using the 2024 bridging stock assessment models, extend the time-series to 2028 assuming constant harvest levels at the projected 2024 mortality for each fishery sector.
2) Fit ‘true’ models to FISS projections that include no trend, a linear 15% positive trend over the next three years (i.e. the FISS index at the end of the period is 15% larger than that observed in 2024; as the actual estimate for 2024 is not yet available it was assumed that the 2024 index was identical to 2023), and a linear 15% negative trend over the next three years using the CVs projected for the base block design. Assume all other data sources (fishery CPUE and age composition information) are sampled at the observed rates from 2023.
3) Using the ‘true’ models, bootstrap all of the data (FISS and fishery) in 2025-2028, to create 100 replicate ‘true’ data sets for each of the three trends.

When evaluating alternative or restricted survey designs it is common to consider only the index of abundance (e.g., Anderson et al. 2024); however, the age composition information is also critically important to estimating year-class strengths which can lead to very different management outcomes for the same or similar index trends. The bootstrapping approach described above naturally produces age composition information along with trend information, that can be either biased or unbiased depending on how it is used.

Based on the simulated data sets from the ‘true’ states, three experiments were conducted (Table 1). Each experiment compared a stock assessment ensemble (all four models) using unbiased trend information (the base block design) to a stock assessment using data representing the two reduced designs. This analysis therefore produced 9 ensembles, crossing the three designs with three trends. These three experiments were: compare core and reduced core designs with no trend (unbiased) to the base block design (also unbiased) to explore the effects of increased CVs and compare biased core and reduced core designs to an unbiased base block design given true FISS trends of +15% and -15%. For models fitting to data based on the restricted designs (core and reduced core), the sample sizes for the age composition data were reduced in proportion to the geographic extent of the sampling (e.g., a reduced core design will include smaller sample sizes than the other two designs and the areas-as-fleets models will have missing data from some biological regions). Fishery CPUE and age composition data were simulated as unbiased in all cases.

Table 1. Design matrix for simulations of FISS design effects on the stock assessment.

‘True’ FISS trend	Estimation models	Inference
No trend	*Unbiased*: No FISS trend, base block design *Unbiased*: No trend, core design *Unbiased* No trend, reduced core design	Effect of increased CV due to reduced designs
+15% over 3 years	*Unbiased*: +15% FISS trend, base block design *Biased*: No trend, core design *Biased*: No trend, reduced core design	Effect of failing to identify an increasing trend
-15% over 3 years	*Unbiased*: -15% FISS trend, base block design *Biased*: No trend, core design *Biased*: No trend, reduced core	Effect of failing to identify a decreasing trend

This approach provides inference on how a reduced FISS might affect the overall results of the stock assessment ensemble. Specifically: How does a reduced but unbiased FISS affect the results? How will management information be affected if we fail to detect an increasing trend? How will management information be affected if we fail to detect a decreasing trend? For each of these questions we compared key management inputs between the ensemble using the base block design and those that are either less precise and/or biased. Because they are central to management decision-making we compared the estimated spawning biomass, the estimated fishing intensity (SPR), and the estimated risk of stock decline at the end of the three-year period.

Overall, there was not a large bias in the estimated spawning biomass for any of the three experiments. The core and reduced core designs, when unbiased, resulted in only a -1% and -2% bias in spawning biomass between 2025 and the beginning of 2028. When the true trend was increasing but the FISS designs were biased (no trend), both ensembles underestimated the true spawning biomass by either -2% (core design) or -3% (reduced core design) at the end of the projection period. When the true trend was decreasing but the FISS designs were biased (no trend), both ensembles overestimated the true spawning biomass by either 3% (core design) or 2% (reduced core design) at the end of the projection period. This relatively small effect size for imprecise and biased FISS indices and age composition data makes sense for several reasons: most of the recruitments that will mature into the spawning biomass over the next few years are already informed by data through 2023, all fishery data was simulated to be unbiased and therefore stabilizes the model results, and reduced FISS designs produce less informative data than a full design, thus influencing the model fit less. This might seem to lead to the counterintuitive conclusion that when conducting a reduced survey that is potentially biased it seems better to have it be *less* informative. However, this is incorrect as the more reduced the survey design becomes the more likely it is that the results are biased.

Estimated fishing intensity (using SPR) in 2027 also did not show a large response when FISS designs were imprecise and/or biased. Fishing intensity remained unbiased for both the core and reduced core designs without bias. When the true FISS tend was increasing (true SPR=47%), but the FISS designs were biased, both the core and reduced core designs underestimated the SPR (overestimated the fishing intensity) by 1%. Conversely, when the true FISS trend was decreasing (true SPR=44%) but the FISS designs were biased, both the core and reduced core designs overestimated the SPR (underestimated the fishing intensity) by 1%. To put this degree of bias in SPR in context, in recent year's decision tables if the Commission wanted to increase the SPR by 1% (at or near the status quo harvest level) a reduction of 1.0-1.5 million pounds of TCEY would have been required. Given an average price of $6 USD per pound in the commercial fishery, this equates to approximately $7.5 million USD that would need to be temporarily forgone to ensure that the management decision was precautionary for a bias of up to 15% in the FISS index.

The third metric that was compared was the probability of spawning biomass decline at the end of the 3-year period from 2027 to the beginning of 2028. As for SPR, there was no bias created in the estimated probability of spawning biomass decline due to more uncertain but unbiased FISS designs. When the true FISS trend was increasing (a 40% chance of stock decline), but we fail to detect this change due to either biased FISS design, we overestimate the probability of stock decline by 6%. When the true FISS trend was decreasing (a 65% chance of stock decline), and we fail to detect this change we underestimate the probability of stock decline by 9%. Comparing to recent decision tables, in order to reduce the probability of stock decline by 9% in the upcoming year, recent management decisions would have required a short-term reduction in the TCEY of approximately 4 million pounds, or $24 million USD given an average commercial fishery price of $6 USD per pound.

When all the model and data assumptions are met perfectly (as in this simulation) the effects of a reduced FISS, even when biased by up to 15%, are relatively small in the short-term. However, it is our experience that the most challenging situations in stock assessment do not arise from expected outcomes, but from either rare events that cannot be included in simulations or from cases where multiple deviations from expectations occur simultaneously. Therefore, we caution that the results of these FISS design simulations should be considered 'best case' outcomes, and that actual stock assessment ensemble results and management performance may be worse under real conditions.

This simulation experiment does not quantify the value of stakeholder perception and confidence in the FISS. Across years including a range of FISS designs, from very large (e.g., 1,558 stations in 2019 and 1,489 stations in 2018) to very small (951 stations in 2020 and 544 stations in 2023), it has become very clear that the entire decision-making process relies heavily on the perception of whether the FISS was comprehensive and sufficient to capture coastwide and regional trends. Even large survey designs have often required repeated comparisons with commercial fishery catch rates and age composition information as well as the specific experiences of harvesters in each of the IPHC Regulatory Areas before a reasonable level of confidence was achieved. Where entire IPHC Regulatory Areas, or entire Biological Regions have gone unsampled, the lack of direct information has affected management allocation decisions and led to stakeholder proposals to freeze mortality limits at or below the previous year's level (Appendix II in IPHC-AM100-INF01-Rev_5). We recognize that stakeholder perception cannot be easily quantified without a specific social science analysis; however, it is nonetheless critically important to the Pacific halibut management process. We suggest that the long-term goal should be to create a sustainable survey design that meets quantitative objectives (both in the annual process and the full MSE), but also satisfies stakeholder needs and represents a point of stability in the management process rather than a point of concern.

Recommendations 3 – Ageing accuracy and precision

Age reading for Pacific halibut otoliths has used two methods over the history of the Commission: counting the rings on whole otoliths (surface reading) and counting the rings along the edge of an otolith that has been broken in half and baked to enhance the contrast in color between the dark and light bands on the structure (break-and-bake reading). Until 2002, all ages were estimated based on surface reading; in that year the primary method transitioned to break-and-bake ageing. During both periods an extensive quality control program (~5-10% per year) resulted in multiple reads (either by the same individual but blind to the first read, or by different individuals), and also comparisons between surface and break-and-bake age estimates of the same otoliths. In addition, most of the 1998 FISS ages were read a second time using break-and-bake ageing.

Break-and-bake aging has been shown to be unbiased using bomb radiocarbon validation (Piner and Wischnioski 2004) and also found to be very precise relative to the ageing of many groundfish species (Clark 2004). Re-aging of samples from each decade from the 1920s to the 1990s has shown that surface aging, although biased for older ages (Figure 1), has remained quite consistent over the full 100-year time series (Clark and Hare 2006; Forsberg and Stewart 2015). The imprecision in break-and-bake ages and the relative bias in surface ages were simultaneously estimated using software that accounts for the joint probability of two (or more) ageing methods based on double- and triple-reads of the same otoliths conducted as part of the IPHC's quality control protocols (Punt et al. 2008). The stock assessment treats ageing error by first calculating the underlying numbers at age in the modelled population, then multiplying these numbers by the age-imprecision key (Figure 2) before comparing the 'expected' numbers at age for each ageing method to the observed data (Methot and Wetzel 2013). The current model treats surface ages and break-and-bake ages separately, using only break-and-bake ages for those years in which that type of data are available. Further, to reduce the potential impact of estimates of the bias in surface ages, the stock assessment uses a 'plus group' (accumulating all ages at or above that age) of age-20 for this source of age composition data. For break-and-bake age compositions a plus group of age 25 is used. These plus groups describe only the aggregation of the data and expected values; the population dynamics in the assessment models include all ages to a plus group of 30 years.

For future use of Artificial Intelligence (AI) based ages (IPHC-2024-SRB025-10) it would be possible to estimate both bias and imprecision through comparisons with break-and-bake ages. Ideally, all ages from a particular year and source (e.g., FISS, commercial fishery, recreational fishery, …) would be aged using the same method, such that it would be possible to model a single age composition. Duplicate age compositions for the same year and source are also possible but would reduce the effective sample size of each and require careful partitioning such that both compositions were random with regard to the overall sampling frame. Specifically, it would not be possible to include a single age composition where otoliths were read by multiple methods without defining that approach as a new 'method; itself and creating an associated bias and imprecision estimate through comparison with break-and-bake ages. There is no limit on the number of age reading methods that can be included in the stock assessment; however re-training the AI algorithm and estimating a unique imprecision matrix for each year would add technical overhead to the already compressed stock assessment process.

There is no threshold for imprecision that would make AI-based ages usable vs. unusable in the stock assessment. Instead, the degree of precision and the choice of which age to use as the plus group (especially if there is high imprecision and/or bias in older ages from the AI method) will dictate the information content of the data. Using a lower plus group will tend to reduce the information on mortality rates but aggregating the right-hand side of the catch curve. Greater imprecision will make it more difficult to detect and track strong cohorts moving through the population. When a suitable data set has been developed, it may be helpful fit stock assessment models with age compositions from the same sources and year but different methods to directly evaluate how the models respond.

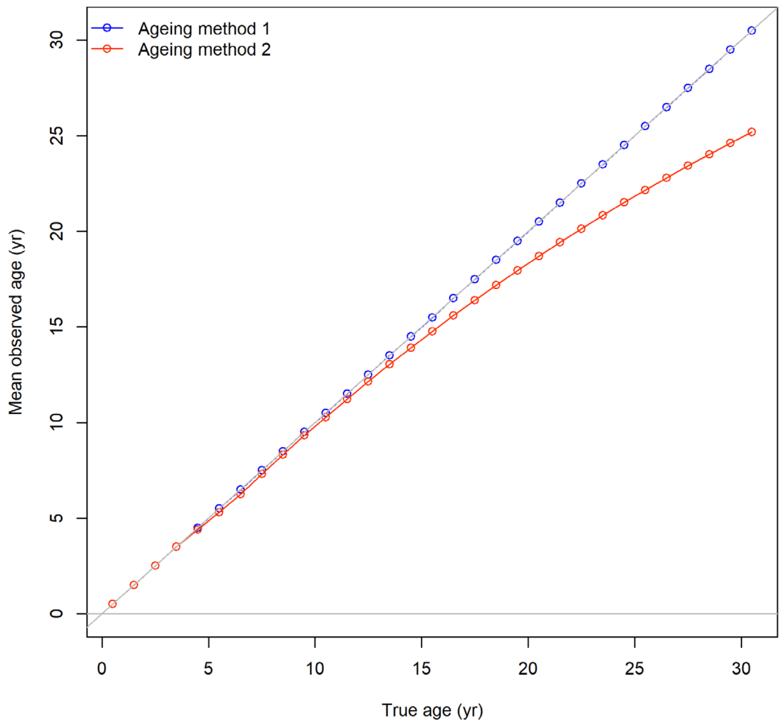

Figure 1. Relative bias estimated for break-and-bake ageing (method 1) and surface ageing (method 2) used in the stock assessment.

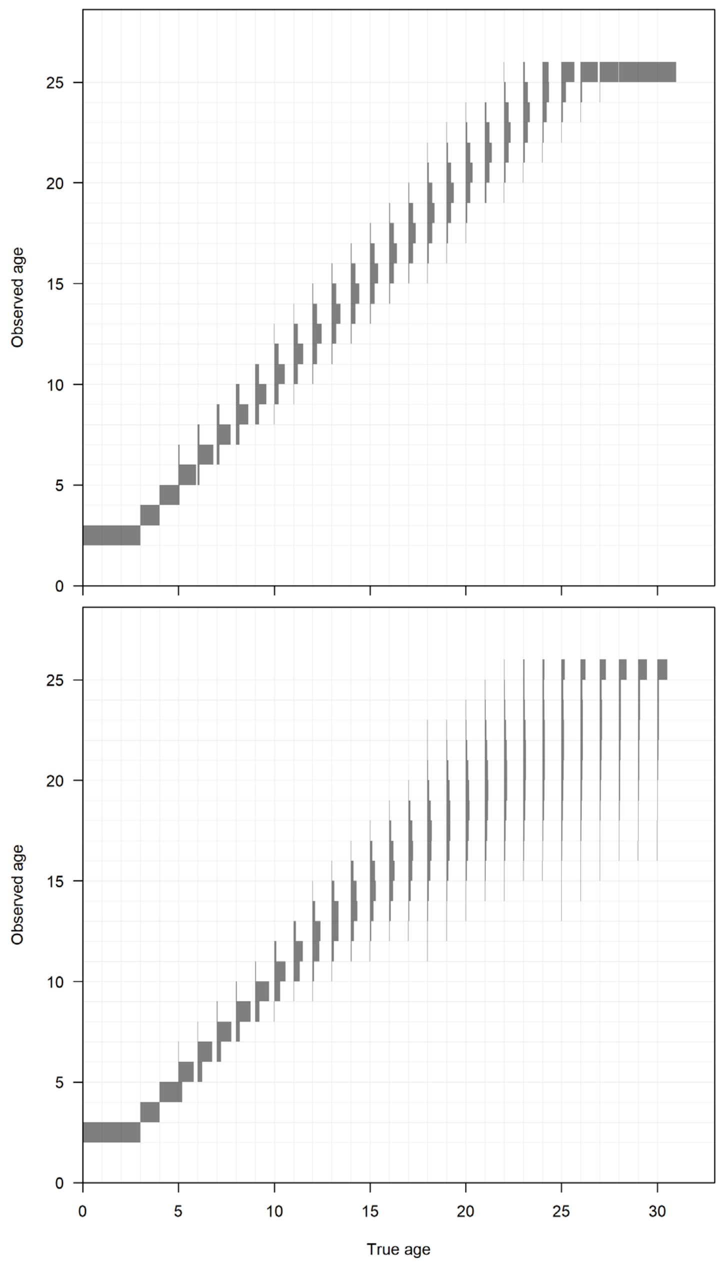

Figure 2. Relative bias and imprecision estimated for break-and-bake ageing (upper panel) and surface ageing (lower panel) used in the stock assessment.

PLANNED DEVELOPMENT IN 2025

The 2025 stock assessment is planned as a full assessment, where all aspects of data processing, model structure and ensemble construction may be revisited. Each recent full stock assessment has included new approaches to data processing and modelling methods, with major changes represented by the addition of commercial sex-ratio information in the 2019 stock assessment (IPHC-2020-SA01) and the estimation of natural mortality in the 2022 stock assessment (IPHC-2023-SA01). For 2025 several development avenues are currently planned:

Maturity and skip-spawning: Histological maturity estimates from 2022-2024 should be available for the 2025 stock assessment. Decisions will need to be made about how to include the multiple years of data (e.g., average them or treat as a time series) and whether to revise the historical maturity time-series or replace recent values with the new results. This ongoing research has the potential to have a large impact on stock assessment results, particularly if evidence of frequent skip-spawning, age/size dependent fecundity, or trends in reproductive output that depend on the environmental conditions are identified. All of these relationships affect estimates of spawning biomass or total reproductive output as well as reference points, thus specifically affecting the potential fishery yield at low stock sizes.

Treatment of the PDO: There is a revised and extended time series for the Pacific Decadal Oscillation (PDO; Mantua et al. 1997) available that includes data from 1854, where the currently used time-series is much shorter. However, the two series are standardized anomalies from the average value over different periods, thus leading slightly different regimes. A comparison of these environmental series, and how they affect stock assessment recruitment estimates and consideration of whether the PDO is still likely to be a useful covariate to recruitment given change in the underlying relationships between environmental variables (e.g., Litzow et al. 2020) is planned for 2025.

Data weighting: As part of the 2022 stock assessment a bootstrapping procedure (Stewart and Hamel 2014) was included in all age data processing to provide an objective starting point for the weighting of the compositional data (Francis 2011). A recent publication has suggested and extension to that method which includes age-reading imprecision in the calculation of effective sample size (Hulson and Williams 2024). This approach will be added to the existing bootstrapping procedure for 2025. Other developments in weighting of compositional data include: a new formulation of the multinomial-Dirichlet distribution that has linear scaling which more closely resembles the multinomial, and an improved calculation of residuals (One-Step-Ahead or OSA residuals) for diagnosing the model fit that do not rely on standard Pearson residuals, which are statistically invalid for composition data that inherently includes a correlation among bins and are therefore not independent and identically distributed (Thygesen et al. 2017; Trijoulet et al. 2023). Exploration of these methods is planned to be included in the overall evaluation of model fit and data weighting.

Natural mortality estimation: In the 2022 stock assessment three of the four individual models in the ensemble estimated natural mortality for both female and male Pacific halibut. The short coastwide model estimated the value for males, but relied on a fixed value for females as there was no clear minima in the likelihood surface for that parameter over a reasonable range of values. Further investigation of natural mortality in that model, including potential confounding with commercial fishery sex ratio, data weighting, and other structural choices is planned.

Other analyses: There may be other improvements to data processing or model configurations that arise during the full assessment and for which the change in model results will be evaluated and documented.

The 2025 full stock assessment will be initially reviewed during SRB026 providing the opportunity for the SRB to make recommendations and for those recommendations to be explored prior to final review at SRB027. The development of the stock assessment is closely tied to Commission decisions leading to a formal Management Procedure (MP). Importantly, the assessment must be targeted to the specific needs of the Commission – a stronger reliance on the MSE output might allow for development of more complex stock assessment approaches and a change in the stock assessment schedule (e.g., from annual to biennial or triennial) would also have major impacts on assessment development goals. For this reason, final decisions about the specific topics and potential degree of change in the 2025 stock assessment will be made after the IPHC's AM101 in January 2025.

OTHER TOPICS

Assessment development during 2024 is occurring in parallel with the ongoing histological maturity study (IPHC-2024-SRB025-08). Although not yet available at the time this document was produced, a sensitivity analysis including the updated maturity schedule for Pacific halibut in the stock assessment models may be available for SRB025. It is anticipated that any major revisions to the stock assessment or to the management results inferred from it will be included in the full assessment planned for 2025. Any preliminary updates on 2024 data will also be provided if available in time for SRB025.

ADDITIONAL STOCK ASSESSMENT DEVELOPMENT FOR 2024

Per standard procedures for an update stock assessment, the secretariat will include routine minor updates and improvements to each of the models and data sets as needed. Standard data sources that will be included in the final 2024 stock assessment include:

1) New modelled trend information from the 2024 FISS for all IPHC Regulatory Areas. Increased variance and the potential for bias is a concern for IPHC Regulatory Areas 2A, 4A, and 4B due to the reduced design. Further, low spatial coverage in 4CDE, 3B and 3A also has the potential to create bias for those and for stock distribution estimates for all IPHC Regulatory Areas.
2) Age, length, individual weight, and average weight-at-age estimates from the 2024 FISS. These data may also contain bias due to the low spatial coverage in the 2024 FISS design.
3) Directed commercial fishery logbook trend information from 2024 (and any earlier logs that were not available for the 2023 assessment) for all IPHC Regulatory Areas.
4) Directed commercial fishery biological sampling from 2024 (age, length, individual weight, and average weight-at-age) from all IPHC Regulatory Areas.
5) Biological information (lengths and/or ages) from non-directed discards (all IPHC Regulatory Areas) and the recreational fishery (IPHC Regulatory Area 3A only) from 2023. The availability of these data routinely lags one year.
6) Updated mortality estimates from all sources for 2023 (where preliminary values were used) and estimates for all sources in 2024.

RECOMMENDATION/S

That the SRB:

a) **NOTE** paper IPHC-2024-SRB025-06 which provides a response to requests from SRB023 and SRB024, and an update on model development for 2024.

b) **REQUEST** any modifications to the 2024 stock assessment.

c) **REQUEST** any analyses to be provided at SRB026 as part of the development of the full 2025 stock assessment.

References

Anderson, S.C., English, P.A., Gale, K.S.P., Haggarty, D.R., Robb, C.K., Rubidge, E.M., Thompson, P.L., and Kotwicki, S. 2024. Impacts on population indices if scientific surveys are excluded from marine protected areas. ICES Journal of Marine Science. doi:10.1093/icesjms/fsae009.

Clark, W.G. 2004. Nonparametric estimates of age misclassification from paired readings. Canadian Journal of Fisheries and Aquatic Sciences **61**: 1881-1889.

Clark, W.G., and Hare, S.R. 2006. Assessment and management of Pacific halibut: data, methods, and policy. International Pacific Halibut Commission Scientific Report No. 83, Seattle, Washington. 104 p.

Forsberg, J.E., and Stewart, I.J. 2015. Re-ageing of archived otoliths from the 1920s to the 1990s. IPHC Report of Assessment and Research Activities 2014. p. 405-428.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences **68**: 1124-1138.

Hulson, P.-J.F., and Williams, B.C. 2024. Inclusion of ageing error and growth variability using a bootstrap estimation of age composition and conditional age-at-length input sample size for fisheries stock assessment models. Fisheries Research **270**. doi:10.1016/j.fishres.2023.106894.

IPHC. 2022. Report of the 21st session of the IPHC Scientific review board (SRB021). IPHC-2022-SRB021-R.

IPHC. 2024. Stakeholder comments on IPHC Fishery Regulations or published regulatory proposals. IPHC-2024-AM100-INF01 Rev_5. 21 p.

Litzow, M.A., Hunsicker, M.E., Bond, N.A., Burke, B.J., Cunningham, C.J., Gosselin, J.L., Norton, E.L., Ward, E.J., and Zador, S.G. 2020. The changing physical and ecological meanings of North Pacific Ocean climate indices. Proceedings of the National Academy of Sciences of the United States of America **117**(14): 7665-7671. doi:10.1073/pnas.1921266117.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.R., and Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society **78**(6): 1069-1079.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research **142**(0): 86-99. doi:http://dx.doi.org/10.1016/j.fishres.2012.10.012.

Piner, K.R., and Wischnioski, S.G. 2004. Pacific halibut chronology of bomb radiocarbon in otoliths from 1944 to 1981 and a validation of ageing methods. Journal of Fish Biology **64**: 1060-1071.

Punt, A.E., Smith, D.C., KrusicGolub, K., and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. Canadian Journal of Fisheries and Aquatic Sciences **65**: 1991-2005.

Stewart, I., and Hicks, A. 2020. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2019. IPHC-2020-SA-01. 32 p.

Stewart, I., and Hicks, A. 2023. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2022. IPHC-2023-SA-01. 37 p.

Stewart, I., and Webster, R. 2023. Overview of data sources for the Pacific halibut stock assessment, harvest policy, and related analyses. IPHC-2023-SA-02. 59 p.

Stewart, I., and Hicks, A. 2024a. Assessment of the Pacific halibut (*Hippoglossus stenolepis*) stock at the end of 2023. IPHC-2024-SA-01. 37 p.

Stewart, I., and Hicks, A. 2024b. Development of the 2024 Pacific halibut (*Hippoglossus stenolepis*) stock assessment. IPHC-2024-SRB024-08. 12 p.

Stewart, I., Hicks, A., Webster, R., and Wilson, D. 2023. Summary of the data, stock assessment, and harvest decision table for Pacific halibut (*Hippoglossus stenolepis*) at the end of 2022. IPHC-2023-AM099-11. 21 p.

Stewart, I.J., and Hamel, O.S. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences **71**(4): 581-588. doi:10.1139/cjfas-2013-0289.

Thygesen, U.H., Albertsen, C.M., Berg, C.W., Kristensen, K., and Nielsen, A. 2017. Validation of ecological state space models using the Laplace approximation. Environmental and Ecological Statistics **24**(2): 317-339. doi:10.1007/s10651-017-0372-4.

Trijoulet, V., Albertsen, C.M., Kristensen, K., Legault, C.M., Miller, T.J., and Nielsen, A. 2023. Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research **257**. doi:10.1016/j.fishres.2022.106487.

Webster, R., Stewart, I., Ualesi, K., and Wilson, D. 2024. 2025-27 FISS design evaluation. IPHC-2024-SRB024-06. 24 p.