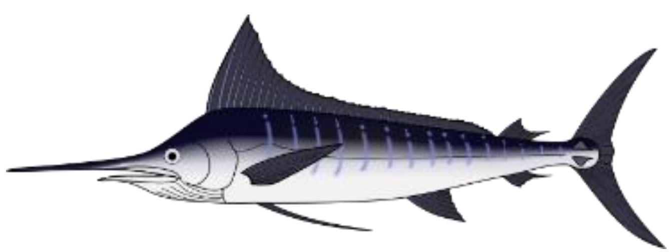
ISC/25/BILLWG-01/03

On the Review of the Stock Assessment Results for Striped Marlin in the Western and Central North Pacific: Japan's Response and a Proposal to the Billfish Working Group[1]

Mikihiko Kai[2] and Marko Jusup[2]

[2]Fisheries Resources Institute, Highly Migratory Resources Division,
Japan Fishery Research and Education Agency
2-12-4 Fukuura, Kanazawa, Yokohama, Kanagawa 236-8648, JAPAN
Email: kai_mikihiko61@fra.go.jp



[1]Working document submitted to the ISC Billfish Working Group Workshop, 13-18 Jan 2025, Honolulu, USA.

Abstract

This working paper summarizes the Japan's response to the review of the stock assessment for striped marlin in the western and central North Pacific conducted in 2024. It also provides a proposal to the billfish working group to develop a detailed plan of tasks for each member country. This plan will be submitted to ISC Plenary 25 to address the short-term and long-term issues identified by the reviewers.

Introduction

A benchmark stock assessment of striped marlin (*Kajikia audax*) in the western and central North Pacific (WCNPO MLS) was conducted using Stock Synthesis (SS3) with fishery data from 1977-2020 by the ISC BILLFISH Working Group (BILLWG or WG) in 2023 (ISC 2023a). The results indicated that the stock was overfished ($SSB_{2018\text{-}2020}/20\%SSB_{(F=0)} = 0.37$) and was likely subject to overfishing ($F_{2018\text{-}2020}/F_{20\%SSB(F=0)} = 1.09$) relative to the dynamic 20-year $20\%SSB_{(F=0)}$-based reference points, where $SSB_{(F=0)}$ is the average of the dynamic B_0 over the last 20 years (2001-2020). However, the WG recognized substantial uncertainties, encompassing—among others—the stock structure, the catch and CPUE data, the life history parameters such as growth and maturity, and the initial equilibrium conditions. To improve confidence in future WCNPO MLS stock assessments, the ISC 23 Plenary had decided to seek the opinions of experts through an external peer review (ISC 2023b), and the WG conducted the in-person review meeting in 2024 (ISC 2024a). The review panel submitted a single consensus report to the ISC 24 Plenary. The ISC 24 Plenary determined that the WG should draft a formal response to the peer review recommendations to be reviewed at the ISC 25 Plenary (ISC 2024b; WCPFC 2024). Additionally, the ISC 24 Plenary requested that this response includes a prioritization of the tasks arising from the recommendations.

The main purpose of this working paper is to summarize Japan's response to the review of the WCNPO MLS stock assessment, including the prioritization of short-term and long-term tasks. It also aims to propose that the WG develops a detailed plan of tasks for each member country and submits it to the ISC 25 Plenary to address the short-term and long-term issues identified by the reviewers.

Japan's response to the review

The italicized text below represents the reviewers' recommendations, while the regular text in red represents Japan's responses, and the regular text in blue represents general comments or requests from Japan to the WG or each member country. Japan's response focuses solely on recommendations related to the Japanese data, rather than addressing all recommendations.

TOR 1 - Review the information available on Pacific MLS stock structure and conceptual model and provide any recommendations for changing WCNPO MLS stock boundaries or to the fleet structure.

1.1. Short-term

1.1.1. No short-term recommendation related to stock structure and boundaries.

1.1.2. Simplify the Japanese fleet structure. The panel recommends not using the CPUE component in the current finite mixture model as a primary determinant for fleet structure. More focus should be given to operation level information to define the fleet structure as the primary objective should be to ensure a consistent fleet-specific selectivity. The key determinant of the fleet structure should be the size structure of the catch. Any fleet categorization should not simply be determined by the model output.

Res: Yes. Japan will stop using finite-mixture modelling to determine fleet structure, and will attempt to simplify the Japanese fleet structure based on the MLS size data caught by the Japanese longliners.

1.1.3. For all fleets, a better understanding of the spatial and temporal structure should be explored to more fully understand the patterns of CPUE, size, targeting, fleet structure and vessel turnover.

Res: Yes. Japan will aim to show the spatial and temporal changes in CPUE, size, targeting, fleet structure, and vessel turnover in the future work.

1.2. Long-term

1.2.1. Continue to pursue genetic research to more fully understand mixing between the genetically distinct population. This research should encompass the whole Pacific. There is an opportunity to utilize the genetic samples from the current IBBS program. This may require some modifications to the current sampling programs to ensure that genetic samples meet the requirements for the questions being pursued.

Res: Japan agreed to pursuing further understanding of stock mixing and structure in the whole Pacific using genetic samples from the current IBBS program and the other Tuna RFMOs in the future work. If Taiwan and the United States agreed, it would be acceptable for Japan to take the lead in this research.

1.2.2. The WG should explore the use of an index fishery approach which would link to an exploration of a unified CPUE analysis (Maunder et al 2020, Xu et al 2023).

Res: We will attempt to apply this method after reevaluating the CPUE of main fleets.

TOR 2 - Model inputs, commenting on the adequacy and appropriateness of data sources and data inputs to the stock assessment.

2.1. Growth: review the approach to estimation of growth parameters and consider the implications of potential regional variations in growth

2.1.1. Short-term

2.1.1.1. Understanding the interaction between growth and selectivity is perhaps the most critical aspect of improving the assessment. The review panel recognized that there is a deficiency of samples at older ages and this has consequences in the estimation of a growth model.

Res: Japan is presently focusing on collecting marlin under 100 cm eye fork length, whereas for the next year, the aim will be to collect the samples of larger individuals over 200 cm as well.

2.1.1.2. The WG should explore the possibility of fitting the growth model within the assessment, together with the inclusion of conditional length-at-age data. There is the option to use a fleet currently in the assessment for selectivity or to create another fleet for this purpose.

Res: Japan strongly recommends that the WG explores the possibility of fitting the growth model within the stock-assessment model, together with the inclusion of conditional length-at-age data.

2.1.1.3. The panel notes that the current IBSS project is a significant step in resolving growth *uncertainty and this project should continue to be well supported. The spatial evaluation of growth is supported by the panel and careful consideration to how spatial growth applies to fleet structure and the interpretation of length data is needed.*

Res: Japan agreed to continue the IBBS project. First, it is necessary to collect the required number of samples from small to large individuals. Then, growth equations should be estimated from otoliths and spines, while simultaneously examining regional differences in growth. This will likely advance the understanding of fleet structure and length data. Japan is currently preparing to do this in time for the 2027 stock assessment.

2.1.1.4. Should there be delays in the production of a new growth curve then effort needs to be put into resolving the potential age bias of using spines to produce the growth curve.

Res: In Dr. Kanaiwa's lab, the daily age assessment using otoliths for age 0 and 1 is almost complete. This could mitigate the suggested bias during young ages (from 0 to 2). Further examination should be pursued to highlight the relationship between length, age, possible ageing bias, and asymptotic length.

2.1.2. Long-term

2.1.2.1. There needs to be continuing effort to resolve growth curves across space, sex, time, and genetic origin.

2.2. Catch: review the treatment of the catch data, especially with regard to catch prior to 1993, when the driftnet catch level is highly uncertain due to unspecified species attribution and spatial extent.

2.2.1. Short-term

2.2.1.1. Efforts should be made to continue to improve the full historical time series of catch and associated uncertainty, even if the full catch time series is not used in the model. This would include an evaluation of the reported catch by other nations and whether or not reported catches make sense given fleet effort and area specific catch rates of reference nations. As an example, given the fishing effort of one nation in specific locations, if the CPUE of another nation was used would you estimate the same scale of catch.

Res: Japan plans to review the catch data for longline and high seas driftnet fisheries and, if necessary, revise the catch amounts before the 2027 stock assessment. Japan requests member countries that have conducted high seas driftnet fishing, such as Taiwan and South Korea, to review their catch amounts.

2.2.1.2. Efforts to improve the characterization (spatial location) and associated uncertainty in the driftnet data would increase confidence in these data and should be pursued.

Res: See 2.2.1.1.

2.2.1.3. Efforts should be taken to more fully understand discarding in all fleets. A potential starting point is the comparison of discarding between observed / training vessels and logbooks.

Res: Japan plans to compare the MLS catch numbers between observed / training vessels and logbooks before the 2027 stock assessment.

2.2.1.4. The reporting of MLS may not have been consistent across vessels and fleets. Efforts should be made to more fully understand if there are reporting biases in all fleets. A potential starting point is the comparison of MLS reporting rates between observed / training vessels and logbooks, at a reasonable spatial and temporal resolution.

Res: Japan plans to compare the MLS reporting rates between observed / training vessels and logbooks, considering spatial and temporal resolution, before the 2027 stock assessment.

2.2.2. Long-term

2.2.2.1. None

2.3. Size composition: review the approach for pre-treatment of size composition data (i.e., reweighting), how size composition is weighted for the likelihood function, and how decisions are made to determine which size data are included.

2.3.1. Short-term

2.3.1.1. Length frequency should be explored in detail at the national level for full historical time series. Methodologically this would include the application of a standardization method to understand the factors that cause spatial and temporal variability in the size data. This would allow a more complete understanding of the spatial and temporal nature of the data, how it relates to population size structure and fleet structure, and the standardization of size data should an index fishery be used.

Res: Japan plans to raise the length-frequency data using catch data before the 2027 stock assessment. Japan requests Taiwan and the United States to conduct the same analysis of their length-frequency data.

2.3.1.2. The US shallow set longline length data should be removed from the length composition data to resolve the discontinuity in the selectivity as a result of the fishery changes that occurred in the early 2000s.

Res: Because it is unclear at this time whether the issue will persist, the final decision should be made during the 2027 stock assessment.

2.3.1.3. Weight data should be used for the Taiwanese fleet instead of the length data since the length data may not always be measured, but can be estimated. The conversion from weight to length should be done outside the model using the fleet-specific length weight relationship. The panel acknowledged that the length-weight relationship in the model is the Taiwanese fleet's relationship but still recommended doing this conversion external to the assessment model. Consider inputting these data as generalized compositions to allow for bin sizes that can span potentially rounded weights.

Res: Japan requests Taiwan and the BILLWG to address this issue.

2.3.1.4. Bootstrapping or a model-based approach should be used to establish initial sample sizes external to the model to account for the properties of the underlying data (Thorson et al. 2017, Stewart et al. 2014).

Res: If upon further examination this method is proven to be effective, its future use could be considered by the WG.

2.3.2. Long-term

2.3.2.1. None.

2.4. CPUE: review the standardization methods and spatio-temporal structure of the CPUE data for each fleet, and the decision process for data weighting and exclusion of indices from the model.

2.4.1. Short-term

2.4.1.1. Make every effort to not split CPUE series (see discussion in Hoyle et al. 2024).

Res: Japan plans to standardize CPUE using data gathered over a long time period (1977-2025) with new standardization methods such as an index fishery approach and a multispecies approach.

2.4.1.2. If the time period of the assessment includes the period with high seas driftnet fishing as early longline fishing, effort should be made to standardize Japanese longline CPUE without *splitting its time series since this could be a data source of the population dynamics bridging before and after 1994.*

Res: See 2.4.1.1.

2.4.1.3. Investigate the potential for monofilament branchlines and other technological advances that may have affected Japanese CPUE during the 1980s.

Res: Japan plans to investigate if such information is available in the fishery data from the 1980s.

2.4.1.4. Use the annual variance estimates from the CPUE standardization as a starting point for model inputs. These can be rescaled but should reflect the differences in precision among years within each series. Where necessary, either iterate or estimate additional variance such that the model fit (RMSE - root mean squared error) is consistent with the average input standard error by fleet.

Res: Japan requests the main modeler of the BILLWG to address this issue.

2.4.1.5. For the Taiwanese CPUE series, omit data before 2006 until there is confidence that these data can be used to provide reliable information. Explore the data for the period prior to 2006 to identify factors that may have caused the jump in CPUE 2003-2005.

Res: Japan requests Taiwan and the BILLWG to address this issue.

2.4.1.6. When there are conflicting CPUE series covering the same period, and these conflicts cannot be resolved, they should be included in alternative model scenarios rather than combining them in the same model (Francis 2011, Schnute and Hilborn 1993, Hoyle et al 2024). This assumes that consideration has been given to the nature of the conflict and that the conflict is not due to a model misspecification (e.g., an error in the fleet designation).

Res: The trends of the new/updated CPUE indices should be compared with each other while considering the area, stock migrations, and fleet selectivity. Age structured production model (ASPM) diagnostics might help to understand the (in)consistency among the indices and catch time series. After those considerations, the final decision should be made during the 2027 stock assessment.

2.4.1.7. Provide full diagnostics for all CPUE series that may be included in the assessment, including residual plots, effect plots (i.e., the effect of the covariate on the expected CPUE), and influence (i.e., the impact of the covariate on the index over time) plots.

2.4.2. Long-term

2.4.2.1. Develop joint CPUE series across nations and multiple fleets, to address the following issues:

- *Provide indices that cover the majority of the stock across the whole time series. Such an approach would help to limit the effect of contracting spatial coverage in some of the fleets. It is unclear what impact declining spatial coverage has on the estimation of annual random fields even when only a portion of the random fields are used in standardization.*
- *Provide a single series using consistent methods for data cleaning and model fitting, rather than multiple series that may conflict.*
- *Develop a shared understanding among the collaborators.*

Res: Japan requests the BILLWG to develop joint CPUE in the future work.

2.5. Data inputs: identify and provide recommendations on the key areas for improvement in data collection (both fishery data and biological information).

2.5.1. Short-term

2.5.1.1. As previously mentioned, continue to improve the full historical time series of catch, regardless of whether the full catch is used in the model, more fully understand and characterize discarding in all fleets, more fully understand and account for reporting discrepancies in all fleets.

Res: See 2.2.1.1.

2.5.2. Long-term

2.5.2.1. The range contraction of some of the fleets highlights the importance of information sharing across fleets and nations. Consideration is needed to ensure that biological information used to make inferences at the population level has the appropriate spatial coverage.

Res: Japan believes that the IBBS project will prove useful in solving this problem.

2.5.2.2. Age sampling for a fleet could provide better estimates of population scale.

2.6. Other life history parameters: review the other life history parameters used weight-length, maturity, natural mortality, stock-recruitment, etc.) for internal consistency and appropriateness for the WCNPO stock.

2.6.1. Short-term

2.6.1.1. There have been developments in how to estimate life-history parameters. New life history data has been and is being collected. Revisiting the life history values used in the model in light of this new information and approaches should be considered, with the goal of internal consistency across the development of the full suite of parameters used in the model.

Res: Japan requests the BILLWG to revisit the estimation of life history parameters other than the growth and maturity curves, such as natural mortality and steepness.

2.6.1.2. There is evidence to suggest a west to east difference in maturity. Note also that samples from spawning areas may not be representative of the population-level maturity ogive. Consideration should be given to addressing these differences by establishing a CPUE-weighted maturity ogive (Farley et al. 2014).

Res: Japan believes that the IBBS project will prove useful in solving this problem.

2.6.1.3. The IBBS project collecting information on growth, maturity, and genetics is a significant step to helping understand the spatial distribution of the underlying life history characteristics and needs continued support to ensure success.

2.6.1.4. The steepness prior should be updated to take into account changes in the input parameters. It is important to allow for uncertainty in the values of all input parameters and uncertainty about the structure of the stock recruitment relationship, which may result in a flatter prior.

Res: Japan requests the BILLWG to update the steepness prior in SS3.

2.6.1.5. New information on growth and maturity will require natural mortality to be updated. Given the potential for spatially varying life history values, spatial consideration will need to be given to how to weight such information appropriately in the estimation of these values.

Res: Japan requests the BILLWG to consider the spatial differences in the life history parameters.

2.6.2. Long-term

2.6.2.1. None.

TOR 3 - Model configuration, assumptions and settings

3.1. Fleet structure: review fleet definitions and spatio-temporal structure of catch, CPUE, and size composition inputs.

3.1.1. Short-term/long-term

3.1.1.1. See recommendation in previous sections

3.2. Selectivity: review selectivity assumptions and settings.

3.2.1. Short-term

3.2.1.1. The population level size bins should be reduced from 5 cm (to 1 or 2 cm) as this has the potential to provide a smoother likelihood surface where selectivity is highly dome-shaped. This does not mean that the bins for the fleet specific length frequency data need to be reduced.

Res: Japan could agree with the reviewers on this point, but it should be recalled that a run with the finer population length bin will make the model run slower. The BILLWG should consider a reasonable width of population size bins in the 2027 stock assessment.

3.2.1.2. Consider setting the parameter defining the width of the top of domed selectivity curves to span at least 2 population size bins.

Res: The BILLWG confirmed that a model run, which estimated all parameters defining the width of the top of dome-shaped selectivity, did not exhibit parameter boundary issues. The WG may want to start from a model that just estimates the top of dome shaped selectivity, and could further explore the minimum width for those parameters.

3.2.1.3. Given the structure of the area-implicit assessment model, a fleet assumed as an asymptotic selectivity should be chosen based on the observed data (i.e. empirical selectivity method). The panel recommends fleet 18, the TWN DWLL, as this fleet has the largest observed fish, though this may be revisited after the data source has been changed from length frequency to weight frequency observations.

Res: The current base case model includes multiple fleets for which asymptotic selectivity was assumed. The reviewers recommended changing the selectivity assumptions for the Japanese drift net fleets (Fleets 13 and 14) from asymptotic to dome-shaped. The current assessment results exhibited some ill fits to the corresponding composition data, in particular right tails of the distributions. When the BILLWG tested a model run with the dome-shaped selectivity for Fleets 13 and 14, the test run showed much better fit to those composition data (see Figure below) without disturbing the model fits to the other data sources. This change did not disturb the model convergence either. Japan agrees with the reviewers on this point.

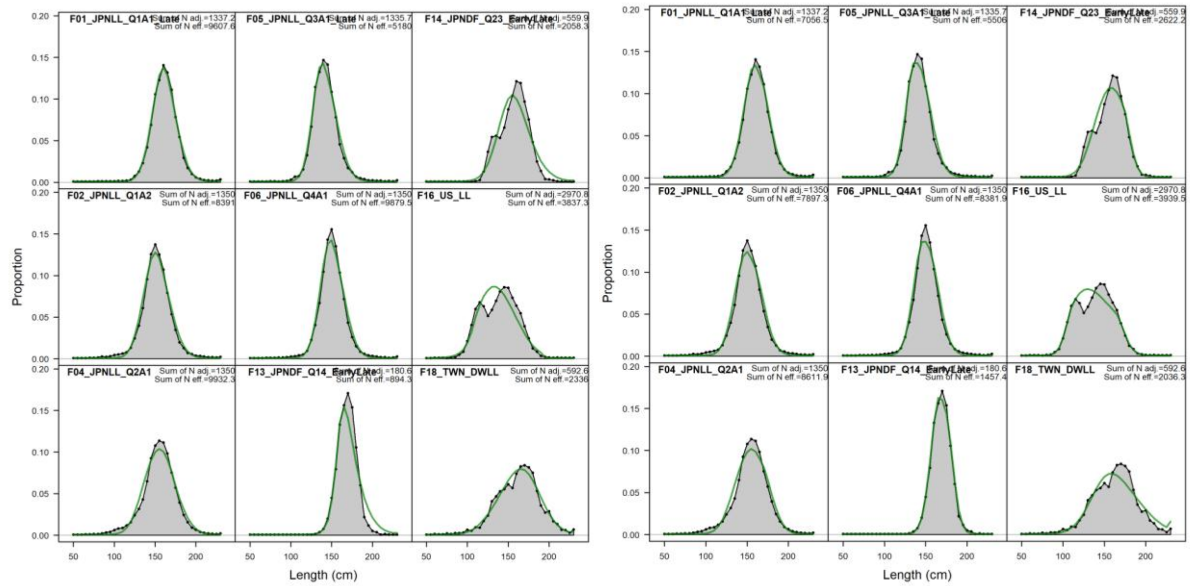


Figure 1.Comparison of observed (gray shaded area and black dots) and model predicted (green solid line) aggregated size compositions for the 2023 ISC WCNPO MLS assessment base case model (left panel) and a test model assuming dome-shaped selectivity for all fleets except Taiwanese longline fleet (fleet 18) (right panel).

3.2.1.4. Aim to remove time blocks from the selectivity parameterization of the Japanese and US fleets.

Res: Because of the possible change in the fishing technology as well as the management measure for MLS or other target species, fishing practice could be changed over time. This could lead to a change in size selectivity. Removing time blocks, as suggested by the reviewers, is a reasonable measure to make the model simpler, but the WG still needs to address the residuals in the size composition data due to the temporal change in observed size distributions. The BILLWG should discuss this point for the 2027 stock assessment.

3.2.1.5. Explore more flexible selectivity for the US longline fleet in order to better fit the bimodal size-composition information. There is also the potential to use an age-based selectivity of this fleet given the apparent length-based modal progression seen in the data.

Res: Japan requests the BILLWG and the US to consider these points in the 2027 stock assessment.

3.2.1.6. Review parameter and asymptotic variance estimates from all selectivity curves and reparametrize where there is no apparent information (e.g. Size_DblN_descend_se_F16_US_LL (16) in the 2023 base case model).

Res: This should be considered with a new candidate model which has reconsidered fleet structure, new size composition data, and a finer population length bin. The BILLWG should review parameter and asymptotic variance estimates from all selectivity curves and reparametrize some of them if there is no apparent information in the 2027 stock assessment.

3.2.1.7. Consider reducing the range of the lower and higher bounds for selectivity parameters (especially those that are logistic transformed) and adjusting the phasing to achieve more reliable convergence. Ideally this would reduce/remove the need for a .par file and assist profile and jitter analysis convergence.

Res: The BILLWG confirmed that a model with a reduced range of the lower and higher bounds (e.g. from -1 to 9 for all ascending and descending rim parameters) did not show any boundary issues. The BILLWG should consider reducing the range of the lower and higher bounds for selectivity parameters and adjusting the phasing in the 2027 stock assessment.

3.2.2. Long-term

3.2.2.1. None.

3.3. Initial equilibrium conditions: review the estimation of initial equilibrium catch and fishing mortality, recommend if the BILLWG should be estimating the equilibrium conditions (as in the 2023 model) or fixing them and running sensitivity runs to evaluate the sensitivity of these conditions (as in the 2019 model).

3.3.1. Short-term

3.3.1.1. As in the 2023 assessment, the initial conditions should be estimated. This includes recruitment deviations for the initial age structure and initial F.

3.3.2. Long-term

3.3.2.1. None.

3.4. Uncertainty: review the approach used to represent uncertainty in model-derived management quantities, considering structural, model and input data uncertainty.

3.4.1. Short-term

3.4.1.1. The review panel felt that the workgroup did a good job of presenting the sensitivity of the results.

3.4.1.2. The panel recommends the adoption of an ensemble model approach. The ensemble should consider: growth, assessment start year, steepness, catch uncertainty, and conflicting time series. Growth should be ensembled based on observed spatial differences in growth if they are identified to exist. Model start date should be ensembled if there is no possibility of linking the early and late parts of the CPUE time series. (A 1977 + and a 1994+ model). Steepness could be ensembled by selecting three steepness values that represent a plausible range of steepness for the species. Uncertainty in catch should be incorporated by including ensembles with high/low, and best estimates. Should conflicting (e.g., JPN vs. TWN) CPUE time series not be resolved the assessment should be ensembled over these conflicting time series.

Res: Japan finds the idea of an ensemble model approach reasonable, but its implementation should be considered once the data update is complete in 2026.

3.4.1.3. The full time series (a 1952+ model) of catch should be included at least as a sensitivity run.

Res: Japan requests the BILLWG to consider using the full time series of catch at least as a sensitivity run.

3.4.2. Long-term

3.4.2.1. Some simulation work would be required to understand the details of how, if it is present, spatial differences in life history characteristics should be accounted for within an assessment model. Best practices have not been determined.

Res: Japan requests the BILLWG to conduct such a simulation study in the future work.

3.5. Start year: review the suitability of the current start year (1977) and suggest potential alternatives, such as 1994 (the start of the high seas driftnet moratorium).

3.5.1. Short-term

3.5.1.1. The panel would prefer the model to begin in 1977, combined with improved long-term catch time series, and with a CPUE series the same length as the assessment period. If discrepancies are not resolved between early and late assessment periods then a model such as the one starting in 1994 was recommended for inclusion in an ensemble approach.

Res: As mentioned in 2.4.1.1., Japan plans to provide the long-term CPUE covering the 1977-2025 period. Japan requests the BILLWG to consider starting the stock assessment in 1977 and in 1994 as an ensemble model approach, if the former scenario has important issues.

3.5.2. Long-term

3.5.2.1. None.

3.6. Alternative models: review the use of SS3 as the modeling software and determine if it is an adequate tool for the assessment.

3.6.1. Short-term

3.6.1.1. The review panel supports the use of an age-structured production model to provide a good diagnostic tool for the current assessment framework.

Res: Japan requests the BILLWG to use an age-structured production model as one of the model diagnostics in the 2027 stock assessment.

3.6.2. Long-term

3.6.2.1. There is a notable reality that the reliance of the assessment on length composition data and the nature of the growth pattern results in challenges in determining the scale of the population. Consideration should be given to the use of close-kin approaches to estimate population scale.

Res: After completing the stock assessment in 2027, Japan requests the BILLWG to start a feasibility study of the close-kin mark recapture (CKMR) approach.

TOR 4 - Model diagnostics

4.1. Review the suitability of the diagnostics used and reported for the assessment.

4.1.1. Short-term

4.1.1.1. Check convergence of all models, to avoid incorrect inferences from models that have not converged. The results of models that have not converged should not be reported, or included as sensitivity runs.

Res: Japan requests the BILLWG to accept this advice.

4.1.1.2. Given that stock rebuilding has been evaluated based on future projection, longer term (i.e., an assessment cycle or generation time) hind-casting could be conducted to determine the prediction skill of the model.

Res: Japan requests the BILLWG to accept this advice.

4.1.1.3. Continue to use a broad suite of metrics to characterize model suitability.

4.1.1.4. Consider the absolute scale of the residuals from diagnostic plots and reweight data sets accordingly (e.g., some of the size data residuals were notably larger than others).

Res: This issue too should be considered with a new candidate model which has reconsidered fleet structure, new size composition data, new initial input sample size, and reconsidered selectivity assumptions. The BILLWG could consider several methods for data weighting such as re-weighting based on the absolute scale of the residuals or model-based estimates of effective sample size in SS3 using the Dirichlet-multinomial distribution.

4.1.2. Long-term

4.1.2.1. None.

4.2. Consider the diagnostics provided for the 2023 WCNPO MLS assessment and provide guidance on follow-up work where the diagnostics suggest issues, i.e., data conflicts.

4.2.1. Short-term

4.2.1.1. The panel felt that the diagnostics provided were helpful.

4.2.1.2. As noted earlier, data sets that are in conflict such as CPUE time series should not be included in the same model but accounted for in an ensemble approach.

4.2.2. Long-term

4.2.2.1. None.

4.3. The driver of the pattern of higher fishing mortality after the high-seas driftnet fishery was banned in 1993.

4.3.1. Short-term/Long-term

4.3.1.1. The flat CPUE through the high catch period is scaled independently of the later period where length frequency data is fit. This appears to have forced the model into a domain where it needs to be highly responsive to recruitment deviations as well as fishery removals to fit the data. The adoption of the review panel's recommendations related to selectivity and continuity of time series should change this pattern. The response of the model to these changes as demonstrated by the additional runs requested suggest these are productive areas of exploration.

Res: Japan requests the BILLWG to consider these key points in the 2027 stock assessment.

4.4. Evaluate the adequacy of the sensitivity analyzes in regard to completeness and incorporation of results. Recommend improvements to the communication of the sensitivity run results (plots, tables, and/or text)

4.4.1. Short-term

4.4.1.1. The panel found the sensitivity analysis presented by the workgroup to be quite comprehensive.

4.4.1.2. *The use of non-zero sum recruitment deviations is a reasonable sensitivity check. However, the use of non-zero sum recruitment deviation can cause a difference between the time-series and the reference point and therefore the reference point would need to be calculated based on the net result of the recruitment deviations.*

Res: Japan requests the BILLWG to consider this point in the 2027 stock assessment.

4.4.2. Long-term

4.4.2.1. None.

TOR 5 - Comment on the proposed reference points and management parameters (e.g., MSY, F_{MSY}, SSB_{MSY}, 20%$SSB_{F=0}$); if possible and feasible, estimate values for alternative reference points or alternative methods of determining the appropriate reference years for the dynamic B0 calculations.

5.1. Short-term

5.1.1. Recommend calculating and reporting both the 20-year moving average as well as the annual dynamic B0 so that the trends can be compared.

Res: Japan requests the BILLWG to accept this suggestion and to show both of 20-year moving average and the annual dynamic B0 in the 2027 stock assessment.

5.1.2. Recommend averaging relative Fs over the last 3-5 years but not including the terminal year for the calculation of $F_{SSB20\%}$ rather than using the terminal year.

Res: Japan requests the BILLWG to use averaging relative Fs over the last 3-5 years but leaving out the terminal year for the calculation of $F_{SSB20\%}$ rather than including it.

5.1.3. The panel suggests continued reporting of additional status metrics such as %SPR or 1-SPR.
Res: Japan requests the BILLWG to continuously use the additional status metrics such as %SPR or 1-SPR.

5.1.4. The panel recommends reviewing the standards outlined by the WCPFC and considering the adoption of the same approach.

Res: Japan requests the BILLWG to review the standards for reference points outlined by the WCPFC.

5.2. Long-term

5.2.1. None.

TOR 6 - Suggest research priorities to improve our understanding of essential population and fishery dynamics, necessary to formulate best management practice, with the identification of priorities to improve future assessments.

6.1. Short-term

6.1.1. The development of an age validated growth curve is essential to improve the reliability of the assessment model.

Res: Japan is responsible for the development of the growth curve in the IBBS project and had already started updating the growth curve using the data collected over the entire North Pacific Ocean.

6.1.2. Consider exploring requirements for CKMR.

Res: See 3.6.2.1.

6.2. Long-term

6.2.1. Continue to develop a more comprehensive understanding of the genetic structure of the entire Pacific as well as the genetic composition of the removals.

Res: See 1.2.1.

6.2.2. Implement CKMR approaches should they prove to be tractable for the population.

Res: See 1.2.1.

6.2.3. Simulation work to understand the best assessment approaches to deal with a complex fishery and life history spatial structure.

Res: See 3.4.2

TOR 7 - Comment on whether the stock assessment methods, results, and assessment decision process are clearly and accurately presented in the detailed report of the stock assessment.

7.1. Short-term

7.1.1. The review panel found the reporting of the process to be well documented, appreciated the extensive supporting material, and was highly appreciative of the effort.

7.1.2. Some of the supporting documentation in the working group papers would benefit from greater detail in the decisions made and well as the diagnostics used. It would be helpful to have this information within these documents. This is important for the development of both CPUE time series and size data. Encourage analysts to follow standard guidelines for documenting these analyses, and development of standards for data areas without them. We also encourage coordination across groups so that they follow similar approaches.

7.2. Long-term

7.2.1. Recommend working with the institutions involved with the assessment and reporting process to ensure that personnel are afforded the time to fully explore data analyses and report comprehensively on the findings.

Discussion

Japan proposes that, in collaboration with the WG members and their organizations, the ISC BILLWG:

1) prepares a response to the ISC 25 Plenary during the meeting period as indicated in this document, and

2) reports the response at the ISC 25 Plenary in June.

References

Farley JH, Hoyle SD, Eveson JP, Williams AJ, Davies CR, Nicol SJ. 2014. 'Maturity ogives for south Pacific albacore tuna (*Thunnus alalunga*) that account for spatial and seasonal variation in the distributions of mature and immature fish'. PLoS One. DOI: 10.1371/journal.pone.0083017.

Francis R.I.C.C. Data weighting in statistical fisheries stock assessment models. 2021. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124-1138

Hoyle SD, Campbell RA, Ducharme-Barth ND, Grüss A, Moore BR, Thorson JT, Tremblay-Boyer L, Winker H, Zhou S, Maunder MN. Catch per unit effort modelling for stock assessment: a summary of good practices. 2024. Fisheries Research. https://doi.org/10.1016/j.fishres.2023.106860

ISC. 2023a. Stock Assessment Report for Striped Marlin (*Kajikia audax*) in the Western and Central North Pacific Ocean through 2020. ISC23 plenary report Annex 14, 12-17 July 2023 Kanazawa, Japan.

ISC. 2023b. Report of the 23rd meeting of the international scientific committee for tuna-like species in the North Pacific Ocean. ISC23 plenary report, 12-17 July 2023 Kanazawa, Japan.

ISC. 2024a. Western and Central North Pacific Ocean Striped Marlin Assessment Consensus Peer Review. ISC24 plenary report Annex 11, 19-24 June 2024 Victoria, Canada.

ISC. 2024b. Report of the 24th meeting of the international scientific committee for tuna-like species in the North Pacific Ocean. ISC24 plenary report, 19-24 June 2024 Victoria, Canada.

Maunder MN, Thorson JT, Xu H, Oliveros-Ramos R, Hoyle SD, Tremblay-Boyer L, Lee HH, Kai M, Chang SK, Kitakado T, Albertsen CM. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fisheries Research. 2020 Sep 1;229:105594.

Schnute, J.T. and Hilborn, R. Analysis of contradictory data sources in fish stock assessment. 1993, Canadian Journal of Fisheries and Aquatic Sciences 50: 1916-1923.

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

Thorson, J.T., Johnson, K.F., Methot, R.D., and Taylor, I.G. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. 2017. Fisheries Research 192: 84-93. doi:10.1016/j.fishres.2016.06.005.

Xu H, Maunder MN, Minte-Vera C, and Lennert-Cody C. Exploratory analysis for the stock assessment of Bigeye tuna in the Eastern Pacific Ocean. The meeting document submitted to the 14th meeting of the Inter-American Tropical Tuna Commission Scientific Advisory Committee, 2023. Document SAC-14-05

WCPFC. 2024. Report of the 24th meeting of the internation scientific committee for tuna-like species in the North Pacific Ocean. WCPFC-SC20-2024/GN-WP-03