ANNEX 15

21st Meeting of the
International Scientific Committee for Tuna
and Tuna-Like Species in the North Pacific Ocean
Held Virtually
July 12-20, 2021

REPORT OF THE ALBACORE WORKING GROUP WORKSHOP

July 2021

Left Blank for Printing

REPORT OF THE ALBACORE WORKING GROUP WORKSHOP

International Scientific Committee for Tuna and Tuna-like Species
in the North Pacific Ocean

31 Aug – 3 and 8 Sep 2020 (Eastern Pacific)
1 – 4, 9 Sep 2020 (Western Pacific)

1. OPENING AND INTRODUCTION

1.1 Welcome and introduction

An intersessional workshop (WS) of the Albacore Working Group (ALBWG or WG) of the International Science Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) was convened as a webinar for 31 August – 3 and 8 September 2020 (Eastern Pacific time) and 1 – 4 and 9 September 2020 (Western Pacific time). Fifteen participants (**Attachment 1**) attended the WS. The objectives of this workshop were to: (1) review MSE progress after the 1st round of MSE, (2) assess if any additional modifications to the MSE framework or additional analyses were required, and (3) provide feedback on presentation of results.

1.2 Meeting protocol

The ALBWG Chair noted that the efforts of the WG at this meeting would be collegial and follow the scientific method with an emphasis on empirical testing, open debate, documentation and reproducibility, reporting uncertainty, peer review, and constructive feedback to authors and presenters.

1.3 Review and adoption of agenda

The draft agenda was circulated prior to the meeting, reviewed and adopted at the WS (**Attachment 2**).

1.4 Assignment of rapporteurs

Rapporteuring duties were assigned to Steven Teo, Yoshinori Aoki, Yuichi Tsuda and Naoto Matsubara.

1.5 Distribution of presentation file availability

Presentation files were distributed to WG members prior to the WS and author contact details was provided (**Attachment 3**).

2. Recommendations from the 1st round of NPALB MSE (4th NPALB MSE Workshop)

H. Kiyofuji provided a short overview of the recommendations from the 1st round of NPALB MSE. The WG thanked the Chair for the refresher and there were no further discussions.

3. Improvements to meet recommendations from the 1st round of NPALB MSE

D. Tommasi gave a presentation entitled: "4th ISC ALB MSE Workshop Recommendations and How They Addressed in the 2nd Round of MSE".

Summary: The presentation outlined how the 23 recommendations from stakeholders and managers at the 4th ISC NPALB MSE Workshop were addressed in the 2nd round of NPALB MSE. To meet some of the recommendations, the MSE framework code had to be modified to 1) compute the probability of breaching SSB_{limit} using the projection software rather than the maximum likelihood estimate of SSB from the operating model; 2) include a management option where TAE applies to the surface fleets, which are the Japan pole-and-line fleet and Eastern Pacific Ocean (EPO) surface fleets, and a TAC is applied to the remaining fleets; 3) limit levels of fishing intensity set by management to historical levels achieved by NPALB fisheries; 4) have an implementation error that includes both positive and negative values; 5) include non-zero candidates for minimum fishing intensity, when $SSB < SSB_{limit}$, ; 6) test all additional harvest control rules proposed in Table ES4 of the 4th ISC NPALB MSE Workshop Report, and 7) include a robustness scenario where an unmanaged new fishery is removing increasing amounts of catch. The results of an analysis assessing the impact of calculating performance metrics using different number of iterations (35, 45, 55, 70) on MSE results was also presented.

Discussion: The WG thanked the author for the work done to improve the models and address the recommendations and concerns from the 4th MSE workshop. Some members expressed concern that non-scientists may have difficulty understanding the use of the word 'depletion' to represent relative biomass (i.e., SSB/SSB_0). However, the presenter noted that non-scientists in previous MSE workshops have not expressed this specific concern but careful explanations of this and other terms in the presentations and report (e.g., in a glossary) will be a good idea. **The WG agreed with the idea**.

A member of the WG suggested using 'worm' plots to show individual simulation runs for various time series plots, in addition to the median and 5th and 95th quantiles because they appeared to be more intuitive for non-scientist participants in other MSE projects. A WG member also suggested that worm plots could be accompanied by pie charts showing proportions of different outcomes (e.g. proportions of years fishing intensity is above average historical). **The presenter agreed to try these suggestions and showed examples to the WG the next day.**

Other WG members enquired about the definition of 'historical' in the plots of projected fishing intensity under different scenarios. The presenter clarified that the period for 'historical' was based on the 2017 assessment (i.e., 1993-2015). Some WG members suggested that it would also be useful to have 'current' periods from both the 2017 (2012-2014) and 2020 assessments (2015-2017) identified in the plots. **The presenter and WG agreed that it was a good idea to do so and recommended that this be done for the WS in December**.

The WG discussed the recommendation on the ranking and/or weighting of management objectives. WG members considered that this was best done during a WS with scientists, managers, and stakeholders together. **Therefore, it was recommended that this be discussed during the MSE WS in early 2021 for managers and stakeholders**.

The presentation stated that the management objective of 'maximizing economic returns' will not be evaluated for the 2nd round of MSE, in part due to the lack of expertise in the WG. However, a WG member noted that CPUE was being used as a proxy for economic returns in the SPALB MSE and wondered if the same approach could be taken for NPALB. There was substantial discussion on taking this approach and there were pros and cons identified for doing so. One important

advantage is that using CPUE as an economic proxy is intuitive to non-scientists and relatively easy to understand and calculate from data. However, in the current OMs and EMs, there are only standardized CPUE indices from the Japanese longline fleet, and there may be substantial work necessary to project the nominal CPUE for the various fleets. This management objective is also likely to be most important to the albacore-targeting fleets, primarily the surface fleets (e.g., troll, pole-and-line), and the changes in the CPUE of these fleets are likely driven primarily by variability in recruitment, and availability/catchability. Unfortunately, these processes are not well understood and would have to be sampled from a statistical distribution estimated during the conditioning procedure. Therefore, it would be similar to examining the juvenile or total population with substantial noise around it. Overall, no clear decision was made by the WG on whether to include CPUE (at least for the surface fleets) as a performance indicator.

There was a discussion on the changes in the projection software used in the 2017 and 2020 assessments and how that would affect the results of the MSE. The projection software is used in the EM to estimate the probability of exceeding the LRP in the next 10 years for each assessment period and initiate any needed management action. Some members of the WG wondered if the changes in the projection software would change the results. **The WG recommended that a small subset of runs be performed to compare the MSE results between the different versions of the projection software, as well as the use of the maximum likelihood estimate of the terminal year SSB and associated uncertainty, as was done in the 1st round of MSE. The decision of which subset of runs to perform was left to the presenter.**

A member of the WG asked for clarification on whether the projected fishing intensity in the 2nd round of MSE is based on resampling historical fishing intensity (as in the recommendations) or is it forced to the target levels (as in the first MSE round). The presenter clarified that in the 2nd round of MSE, the fishing intensity is based on resampling historical fishing intensity (as in the recommendations). Another WG member enquired if fishing intensity was based on some measure of effort like number of vessels or number of fishing days. The presenter clarified that the fishing intensity is currently not modeled with measures of effort in the OMs. Based on a previous WG paper (ISC/20/ALBWG-01/05), there is a correlation between effort and fishing intensity for the main surface fleets but not so for the longline fleets that do not target albacore. It may therefore be possible to use this correlation to translate between fishing intensity and effort in the future. **The WG considered that such a translation between fishing intensity and effort would be useful at the future MSE workshops with managers and stakeholders and recommended preparing one for that.** This was followed by a discussion on the implementation errors assumed for TAEs and TACs. It is currently assumed that the implementation error for TAEs and TACs are the same. However, it may not be the case, and this is worth investigating.

Given that the next MSE WS for managers and stakeholders will likely occur before the ISC Plenary, it may be a good idea to distribute the preliminary report to the WS participants prior to the WS even though the ISC Plenary has not reviewed it. **The WG thought it was a good idea and recommended doing so as long as the ISC Plenary agrees. The WG Chair agreed to ask the ISC Chair about this matter in the near future.**

The presenter demonstrated several options for ranking the performance indicators of candidate HCRs, which was one of the recommendations from managers and stakeholders. Some members

noted that using fine scale ranking to differentiate between similar performing HCRs may be misleading. **The WG recommended using broader risk classes based on the Table 4 from the 2nd ISC NPALB MSE workshop (attachment 5 in http://isc.fra.go.jp/pdf/ISC16/ISC16_Annex_08_Report_of_the_ALBWG(Apr2016).pdf) to group performance metrics based on probabilities.** For metrics not based on probabilities, it is suggested that the metric be split in classes prior to ranking. The presenter agreed with these comments and will follow this recommendation in the later presentations. **The WG also agreed with the presenter that the 70 iterations for the 2nd round of MSE was adequate.** If certain iterations of the runs did not converge, the same set of converged iterations should be used to compare the candidate HCRs, noting which HCRs failed to complete the 70 iterations and why.

A member of the WG, when looking at the results comparing the impact of iteration number on performance metrics, noted the similarity in performance metrics results across HCRs and asked how the uncertainties in the OMs and EMs affected the results. The presenter explained that in this particular example, the similarities in the performance of the HCRs were mainly due to the HCRs only differing in the threshold reference point or the management action when the limit reference point is breached, behaving similarly. This will be explained in greater detail over the next few days of the meeting as all preliminary MSE results are presented.

4. Overview of TAC results

D. Tommasi gave a presentation entitled: "ISC ALB MSE Round 2 Preliminary TAC Results".

Summary: The preliminary results of all HCRs with TAC control and where fishing intensity equals the F_{target} when SSB is above $SSB_{threshold}$ were presented for all management objectives and performance metrics. Results were presented across reference scenarios (OMs) 1, 3, and 4 and separately for scenario 6, the low productivity scenario. Results of a robustness scenario, the ghost fleet scenario, where an unmanaged fishery is removing increasing amounts of catch were also presented. The largest differences in performance under the ghost scenario occurred for Performance Metric (PM) 4, the odds of catch being greater than historical, calculated for the managed fleets. Management for all HCRs tested was able to reduce fishing intensity over time ensuring low odds of the limit reference point being breached, but this was at the cost of reduced catches for the managed fleets.

Discussion: The WG discussed the preliminary results of the 2nd round of the MSE, as they were presented. A member of the WG noted that plot describing the management actions associated with each HCRs are based on fishing intensity, which might be confusing to stakeholders and managers as the results being presented refer to TAC control. The presenter agreed to improve the figures explaining the HCRs so as to make it easier to understand.

A WG member noted discrepancies between the OM scenarios listed here and those reported in the report of the 1st round of MSE. The presenter was not aware of this but will check and get back to the WG as soon as possible. Subsequently, the presenter reported that apparent discrepancy was due to different OM scenario IDs being used here and in the previous report. The OM numbers listed in Table 8 of the 1st round of MSE Report are different than the scenario numbers, which here have been taken as OM numbers. However, the scenario numbers in Table 8 are the same and can be used to link the OM number from the 1st round of MSE to the OMs described here. In addition, another WG member enquired if the estimation error for each scenario was

from an assumed distribution. The presenter clarified that the estimation error is the actual error between the EM, where a full assessment model is run, and the OM.

The presenter showed examples of 'worm' plots and pie charts of fishing intensity that were suggested in Section 3. The WG could see the individual trajectories of fishing intensity and identify periods when certain HCRs were triggered as RPs were breached. **The WG thought that the 'worm' plots and pie charts were quite useful and recommended that 'worm plots' and pie charts be used to help illustrate the time series results in addition to the other plots presented (e.g., median, quantiles).** One WG member asked if the same random seed was used for the corresponding runs between candidate HCRs (i.e., recruitment deviates would be the same for runs with the same run ID). The presenter clarified that this was indeed the case. For example, Run 1 with HCR1 would have the same recruitment deviates as Run 1 from HCR10. The WG member followed up by suggesting using example trajectories to help explain what the differences between HCRs (i.e., using a storytelling approach to help explain differences). The presenter agreed with it and stated that such an approach was used in the 1st round of MSE and was quite successful. **The WG subsequently discussed this and recommended that such a 'storytelling' approach be used again for the 2nd round.** In addition, one member of the WG also suggested using violin plots instead of the median and quantile plots because violin plots are better able to show the distribution of the results. **The WG agreed with this suggestion and recommended trying violin plots instead of median and quantile plots.**

A member of the WG commented that since the WCPFC had already adopted a LRP of $SSB_{20\%,F=0}$ for NPALB, it may be confusing for managers that there are three candidate LRPs ($SSB_{20\%,F=0}$, $SSB_{14\%,F=0}$, and $SSB_{7.7\%,F=0}$) in the MSE. The presenter responded that these candidate LRPs were requested by the managers and stakeholders in the previous MSE WS. In addition, another WG member responded that the current LRP for NPALB in the WCPFC is considered interim and can be changed by the managers if better candidates are found during the MSE process. Subsequently, there was a discussion of the scientific basis for the 3 candidate LRPs and why there are three candidate LRPs. Some members of the WG responded that the $SSB_{20\%,F=0}$ is the current interim LRP adopted by the WCPFC in the WCPO, the $SSB_{7.7\%,F=0}$ LRP is based on the process used by the IATTC to develop LRPs in the EPO, and the $SSB_{14\%,F=0}$ is approximately MSY. Therefore, all 3 candidate LRPs have scientific validity and are used by various RFMOs as LRPs. In addition, it was noted that this stock spans the WCPO and EPO, and is managed by both WCPFC and IATTC, and the managers requested to explore candidates from both RFMOs as well as MSY for the LRP. **After much discussion, the WG agreed to keep all 3 candidate LRPs in the 2nd round of MSE**.

There was also a clarification that the LRP adopted for tropical tunas by the IATTC is the 7.7% of the equilibrium SSB_0, while the one in the NPALB MSE is 7.7% of the dynamic SSB_0. These two quantities are likely to be different. The presenter suggested to do more runs with the equilibrium LRP to assess the differences, the WG agreed with that suggestion

The WG discussed possible improvements to the 'spider' plots used to identify tradeoffs between management objectives and compare HCRs. Discussion centered on the use of 'Conservation Risk' as the label for the probability of not breaching the LRP. There was some concern that the label can be misleading because the probability of breaching the LRP depends largely on which LRP was in the HCR and is not necessarily an indicator of conservation risk. **Therefore, the WG recommended not using 'Conservation Risk' as the label for this management objective**

and suggested other options like 'Odds of not breaching LRP'. In addition, some WG members warned that using this performance indicator may lead to choosing a LRP at a lower SSB level while thinking that it has lower conservation risk because it does not get breached as much. However, other WG members reminded the WG that the managers suggested this performance indicator not as an indicator of conservation risk but as an indicator of the frequency of drastic management action (linked to the LRP) being required. In addition, the conservation risk of SSB reaching low levels is already covered by the 'Odds depletion > historical' indicator. Instead of changing this performance indicator, it was instead suggested to change the label (as above) and add a performance indicator for the level of depletion being above low levels (e.g., Odds of depletion > 7.7% and Odds of depletion > 20%). The WG noted that the managers did not request this indicator, but it could be part of a table with other indicators in the report. **The WG agreed with this and recommended the 'Odds of depletion > 7.7% and 20% based on both dynamic and equilibrium SSB_0' indicators be placed in such a table with other indicators**.

The presenter showed the preliminary results of the scenario with a 'ghost' fleet, which is used to evaluate the robustness of the HCRs to such a scenario. A member of the WG enquired about how the TACs were calculated from the EMs and OMs, and how the data for the EM were generated. The presenter explained that the exploitation rates from the EM were multiplied by the biomasses estimated in the EM to calculate the appropriate TACs because the only information available to set the TAC comes from an assessment (i.e., the EM). The presenter further explained that Stock Synthesis is able to generate 'bootstrap' files during the OM part of the run that acts as data input files for the EM. Another member of the WG suggested it may be good to show the impacts of the 'ghost' fleet on recent catches relative to the historical maximum, and to add some context to the results and conclusions from this analysis. Another member of the WG asked how large an error there was in the SSB estimates from the EM vs the OM in the 'ghost' fleet scenario.

A member of the WG also highlighted that there is a need to discuss whether a specific fleet should be identified as representing the 'ghost' fleet for future presentations, and if not, how the specifics of the fleet should be described.

5. Comparison of results with different options for when SSB > threshold reference point

D. Tommasi gave a presentation entitled: "ISC ALB MSE Round 2 Impact of Different Management Options when SSB > $SSB_{threshold}$".

Summary: Results of all HCRs with TAC control, where the fishing intensity is sampled from historical fishing intensity when SSB is above $SSB_{threshold}$ rather than being set equal to the F_{target} were presented. This was done for Scenarios 1 and 3 as setting fishing intensity equal to the F_{target} when SSB is above $SSB_{threshold}$ implies setting a fishing intensity (and a TAC) higher than what has been observed historically. As historical fishing intensity is lower than the F_{target}, particularly for an F_{target} of F40, biomass is maintained at a higher level, reference points are not breached as often, and there is less difference in performance across HCRS as compared to the option where fishing intensity equals the F_{target} when SSB is above $SSB_{threshold}$.

Discussion:

The presenter showed the effect of using the new algorithm for generating exploitation rates and catches for the 2nd round of MSE. A member of the WG enquired if the materials presented earlier were all from the old algorithm. The presenter affirmed that this was the case. **The WG recommended that the 2nd round of the MSE should prioritize the new algorithm to avoid confusion.** Although it may be useful to have a subsection showing the effect of the change.

The WG also came up with several additional suggestions for improving the presentation for managers and stakeholders. It would be useful to have a table of performance indicators with the indicators as numbers and color-coded. It would also be useful to have the additional performance indicators like (P(management action), P(SSB > $SSB_{7.7\%,F=0}$**).**

6. Overview of results for mixed TAC/TAE control and comparison with TAC control

D. Tommasi gave a presentation entitled: "ISC ALB MSE Round 2 Preliminary Overview for the Mixed TAC/TAE Control Strategy".

Summary: The preliminary results of all HCRs with mixed control, TAE control on the surface fleets and TAC control on longline fleets, were presented for all management objectives and performance metrics for Scenarios 3 and 6. While HCRs performance between TAC and mixed control was comparable for Scenario 3, mixed control for Scenario 6 maintained a higher biomass, depletion, and catch stability than TAC control and similar odds of catch being greater than historical as less management intervention was required. Under mixed control, stock biomass does not fluctuate as widely as surface fleets catches respond quickly to changes in available biomass and they are not impacted by assessment errors in biomass estimates. However, under Scenario 6, the assessment overestimates current biomass, and thus catch ratios of TAE controlled fleets (surface fleets) are lower than under a TAC.

Discussion: The presenter started by showing details of the OM because some members of the WG requested a refresher. Some members asked for a clarification on how the actual catch or fishing intensity was calculated. The presenter also clarified that the RPs are all based on dynamic SSB. Some members wondered if the results, especially for $SSB_{7.7\%,F=0}$, would be different if it were instead based on equilibirum SSB_0. The presenter agreed to take a look at the $SSB_{7.7\%,F=0}$ results and report back in December.

One member of the WG noted that the median of the SSB/LRP performance indicator may not show differences that are apparent if you compare the same replicate for different HCRs. The presenter agreed to take a look at the feasibility of doing so. Other members of the WG noted that even though median SSB was similar for several HCRs, the probability of breaching the LRP can be quite different. The presenter explained that this was because the LRPs are different for the different HCRs and the probability of breaching the LRP is a management objective. This is similar to the discussion in Section 4.

Responding to WG queries, the presenter explained in more detail how the TAE and TAC were calculated, especially with the new algorithm sampling historical fishing intensity instead of

forcing the fishing intensity to match the TAE or TAC. A WG member enquired about whether the calculation of the TAC under the TAC or TAE/TAC approach is based on the EM. The presenter explained that the calculations for the TAC were based on EM SSB estimates, while the expected catch for the TAE were based on OM SSB estimates. Other WG members further enquired if differences in the results of the TAC vs TAE/TAC approaches are due to the differences in biomass estimated from the OM (TAE) vs EM (TAC). The presenter agreed that the differences seen in the results between TAC and TAE/TAC may be due to this difference, since the implementation error is currently assumed to be the same for both.

There was substantial discussion of implementation error and OM weighting during this session. The presenter noted that the model fits of the OM during conditioning process and the relationship between effort and exploitation rate will be presented in a later session. The WG agreed to postpone the discussion until then.

Some members of the WG suggested that the performance indicators could be broken up into various periods (e.g., short, medium, and long term). The presenter agreed that it might be useful for some managers and stakeholders and would attempt to do so if there is time.

The WG members also discussed the potential of a 'Shiny' app to help managers and stakeholders understand the MSE results. Some RFMOs have developed apps for this or similar purposes, and code may be available to be shared to help development. The presenter suggested two possible development paths for such an app: 1) an app to visualize and interact with the MSE results that have already been completed; and 2) a simpler and faster MSE model (primarily simplifying the EM to a random sample of a statistical error distribution instead of a full assessment model) so as to respond to manager and stakeholder requests efficiently during the upcoming MSE workshop. This simpler MSE model could have a Shiny app built around it. **The WG agreed with the ideas and recommended that the presenter consider one or both above mentioned tools to help effectively present the results of the MSE.**

7. Overview of Implementation error, OM weighting, and Assessment error

D. Tommasi gave a presentation entitled: "Preliminary Overview of Assessment Error".

Summary: The relative estimation model errors, calculated as $(Value_{OM}-Value_{EM})/Value_{OM}$, in the inputs from the estimation model (EM) that inform the HCRs, $SSB_{current}$, $F_{current}$, F_{target}, and dynamic SSB_0, were presented. For $SSB_{current}$ and dynamic SSB_0, the error was calculated on the log-transformed values. Estimation error differed across scenarios, being largest for Scenario 6, and smallest for Scenario 1. While errors in both SSB_0 and Ftarget were unbiased for Scenario 1 and, for SSB_0, for Scenario 3, other quantities were over or underestimated relative to the "true" OM value and the direction of the bias differed by scenario. For instance, $SSB_{current}$ was overestimated in the EM under Scenarios 1, 4, and 6, but underestimated in Scenario 3. By contrast, $F_{current}$ was overestimated under Scenario 3, but overestimated under Scenarios 1,4, and 6. While errors in SSB_0 and F_{target} were consistent across HCRs for each scenario, errors in $SSB_{current}$ and $F_{current}$ varied between HCRs. It was proposed that these errors could be used to develop a faster MSE with a simplified EM to enable quicker comparison of potential HCRs at workshops with stakeholders.

Discussion: A WG member noted that the LRP of $SSB_{7.7\%,F=0}$ in the MSE is based on dynamic SSB while the IATTC LRP is based on 7.7% of equilibrium SSB0, and asked if that would result in

different behaviors for the HCRs. The presenter responded that this has not been explored yet but will do some runs to examine this issue.

The presenter showed the results of a simple model examining the relationship between effort in fishing days and estimated exploitation rate for the EPO surface fishery. The estimated errors in this model could be used to parametrize the implementation error for the TAE instead of assuming a fixed error identical to the TAC approach. One WG member suggested that instead of this approach, a more direct approach could be taken to examine the relationship between observed catch and effort, and use that relationship and uncertainty in the simpler EM. Other WGs members thought the approach was plausible but it would need to be corrected for total population size and selectivity. The presenter agreed to examine this approach. Another WG member thought that the relationship used may be inverted because in reality, the relationship between effort and exploitation rate is first used to translate the appropriate TAE from exploitation rate to effort, and used to set the appropriate TAE. However, the resulting effort (TAE + implementation error) results in an exploitation rate with substantial uncertainty that is used in the OM. The presenter agreed to examine this approach to see if it is feasible.

There was a substantial discussion on the weighting of the OM scenarios. In the previous round of MSE, Scenarios 1 (base), 3, 4, and 6 were all included, with equal weights, in the 'reference' set of OMs used to calculate the overall performance metrics of the candidate HCRs. However, some members of the WG also noted that, based on what we know of albacore biology and model fits, some OMs are more plausible than others, and weighting the models may result in a better representation of the performance of the HCRs. In particular, Scenario 6 (a low productivity scenario) was considered to be the least plausible and the WG discussed the possibility of using Scenario 6 as a robustness OM instead of in the 'reference' set. There were various pros and cons to including or excluding Scenario 6. Model fit during the conditioning process was considered to be inadequate as the sole criteria by which to weight these OMs. Previous discussions of the biological plausibility of OM scenarios suggested the following weights: Scenario 1 (high), 3 and 4 (medium), and 6 (low). However, using this ad hoc 'expert' opinion of biological plausibility and model fits will result in Scenario 1 being very heavily weighted. Weighting a single scenario so heavily might result in a lack of robustness for the 'best performing' HCR. **After much deliberation, the WG decided to include Scenarios 1 (base), 3, 4, and 6 with equal weights, in the 'reference' set of OMs used to calculate the overall performance metrics of the candidate HCRs. However, the WG also recommended that the biological plausibility and model fits of the OM scenarios be included in a table when describing the OM scenarios. In addition, the WG recommended that in addition to the overall performance metrics, the performance metrics of each OM scenario be included as an appendix to the upcoming MSE report.**
And some results from specific OMs (e.g. Scenario 6) could be highlighted to point to particular behaviors. The WG decided that the stock assessment model in the 2020 assessment will not be used as an OM for the MSE.

Lastly, the presenter showed an analysis of EM error (i.e., the differences in the estimates between the OM and the EM) by OM scenario and HCR. The presenter noted that the main purpose of this analysis is to examine the plausibility of using a simpler EM in the simplified MSE model used to respond quickly to requests during MSE workshops (e.g., what about F45 as a TRP?). Based on this, the **WG recommended that the work continue on this approach but would need to**

consider and work with the projection software developer to provide estimates of the probability of breaching the LRP.

The discussion on EM error also included a discussion on how to deal with changes in the assessment model (i.e., monitoring system) in the real world, and how to decide when the HCRs may no longer perform as expected from the MSE (i.e., meta-rules for exceptional circumstances). There was substantial discussion on this topic but there were no clear a priori answers to this problem. **The WG recommended one possible approach: when HCR-based management action does not result in the expected response by the stock, including the expected uncertainty and time frame, a new MSE should be considered. The WG also recommended that the need for meta-rules be communicated to managers and stakeholders. The WG noted that it is important that the monitoring system be decoupled from the management system.**

7. Administrative Matters

7.1 Time and place of next meeting

The WG developed a work plan for the 2nd round of MSE (**Attachment 4**). The next meeting of the WG is expected to be a webinar during November 30 -December 3, and 8, 2020 (Eastern Pacific) and December 1 – 4, and 9 (Western Pacific) to review the 2nd round of the MSE.

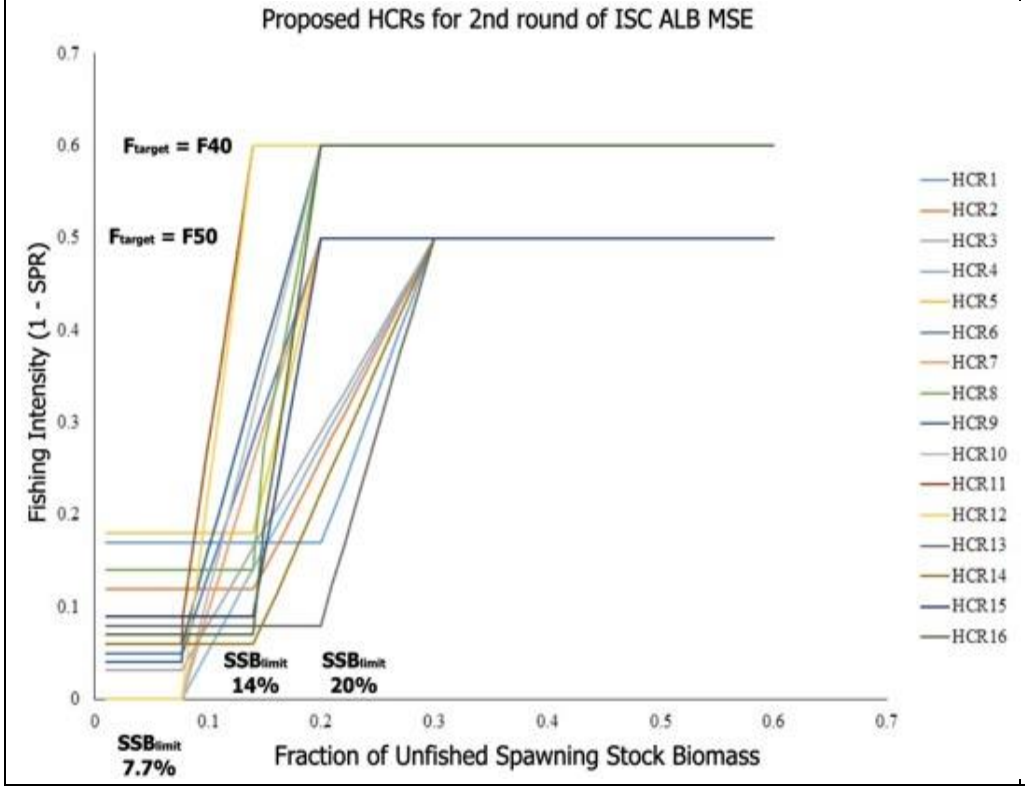
8. Clearing of the report

The WG Chair prepared a draft of the report, which was reviewed by the WG prior to adjournment of the workshop. After the workshop, the WG Chair evaluated and incorporated suggested revisions, made final decisions on content and style and distributed a second draft via email for approval by WG members.

9. Adjournment

The ALBWG meeting was adjourned on 8 and 9 September 2020 (Eastern and Western Pacific Date, respectively). The WG Chair thanked the scientists participating in the workshop for their attendance and contributions on north Pacific albacore MSE.

Table 1. HCRs proposed for the second round of MSE during the 4th ISC ALB MSE workshop, showing the different combinations of TRPs (F50 and F40), threshold reference points (SSB30%, SSB20% and SSB14%), limit reference points (SSB20%, SSB14% and SSB7.7%), and minimum levels of fishing intensity when spawning stock biomass is below the limit reference point.



Biomass corresponding to a 50%Ro Under conservative steepness of 0.75 Maunder and Deriso 2014

B_{msy}

HCR	Target reference point (F_{target})	Threshold reference point ($SSB_{threshold}$)	Limit reference point (SSB_{limit})	Prob SSB > SSB_{limit}	TAC_{min}	Action if SSB > $SSB_{threshold}$
1	F50	30%	20%	0.8	0.25	SSB_{limit} from WCPFC
2	F50	30%	14%	0.9	0.25	
3	F50	30%	7.7%	0.9	0	
4	F50	20%	14%	0.9	0.25	F = TRP
5	F50	20%	7.7%	0.9	0	or
6	F40	20%	14%	0.9	0.25	No harvest control (F sampled from historical distribution)
7	F40	20%	7.7%	0.9	0	
8	F40	14%	7.7%	0.9	0	
9	F40	30%	20%	0.8	0.5	
10	F50	30%	14%	0.9	0.5	
11	F50	30%	7.7%	0.9	0.25	
12	F50	20%	14%	0.9	0.5	
13	F50	20%	7.7%	0.9	0.25	
14	F40	20%	14%	0.9	0.5	
15	F40	20%	7.7%	0.9	0.25	
16	F40	14%	7.7%	0.9	0.25	

Table 2. Progress on improvements to meet recommendations from the 1st round of ISC NPALB MSE. Note that colors in left column is categorized as presentation of results (blue), Management objective (orange), candidate harvest strategies, reference points and harvest control rules (green), workplan (purple) and others (yellow), respectively.

	Recommendation	**Progress**
1	The ALBWG should be more explicit in the labelling of performance indicators and specify if an indicator is based on a probability. For example, for Management Objective #2, the performance indicator labelled "Relative total biomass" was actually the probability of the depletion of total biomass being over the minimum historical depletion and could instead be labelled "probability of total biomass > minimum historical".	Performance Indicators for Management Objectives #2 was replaced to "Odds depletion > historical" and for #4 was replaced to "Odds catch > historical", respectively.
2	Performance indicators using relative total or spawning biomass are likely to be better understood than indicators using probabilities. Separate plots of the mean or median of the relative biomasses coupled with plots of the variability of those relative biomasses may be preferable to a single plot of probabilities. Comparison with historical levels could be done by including indications of the historical levels to be compared.	To use 'worm' plots to show individual simulation runs for various time series plots, in addition to the median and 95% confidence intervals with pie charts to show proportions of different outcomes. (**See Figure 2**). The WG also recommended to try violin plots instead of median and quantile plots.
3	The ALBWG should provide guidance on how to interpret fishing intensity in terms of implications to fleet management. For example, it would be useful for managers to be shown the changes in fishing intensity relative to current fishing intensity.	Additional analyses were conducted and presented by the MSE specialist. (**See http://isc.fra.go.jp/pdf/ALB/ISC20_ALB_1/ISC20-ALBWG01-05.pdf**).
4	Managers and stakeholders should prioritize, rank, or weight the management objectives to assist decision making and help resolve trade-offs in management objectives.	It was recommended that this be discussed during the MSE WS in early 2021 for managers and stakeholders.
5	Management Objective #6 was considered of relatively low priority by managers and stakeholders in evaluating candidate reference points and harvest control rules.	Results relevant to management objective 6 will still be presented in the 2nd round of MSE report.

	Recommendation	Progress
6	The ALBWG should try to obtain the necessary expertise to evaluate the Management Objective of "Maximizing the economic returns of existing fisheries". However, this would be a longer-term goal beyond the 2nd round of MSE.	The WG noted that CPUE could be as an economic proxy, however, in the current OMs and EMs, there are only standardized CPUE and there may be some work necessary to project the nominal CPUE for various fleets. No clear decision was made by the WG on whether to include CPUE to evaluate this objective as a performance indicator at this stage.
7	As the MSE process continues, it should be emphasized that the overarching objective running through all the management objectives of the MSE is to maintain the viability and sustainability of the current NPALB stock and fisheries.	To be emphasized in the final report for the 2nd rounds of MSE.
8	The 2nd round of MSE should focus on Harvest Strategy 3 using the specific reference points and harvest control rules listed in Table 4 (in summary report: http://isc.fra.go.jp/pdf/ISC19/ISC19_ANNEX12_Report_First_North_Pacific_Albacore_MSE.pdf).	• These recommendations were reflected in the development of the 2nd round of MSE. • MSE specialist has been working on HS3 and TRPs of F40 and F50 with different combinations of LRPs and threshold reference points (**See Table 1**). • Three LRPs ($SSB_{20\%,F=0}$, $SSB_{14\%,F=0}$, and $SSB_{7.7\%,F=0}$) requested by the managers and stakeholders were also evaluated for further consideration of LRPS.
9	Harvest Strategy 1 should be removed from further consideration because it performed poorer in terms of Management Objective #1 relative to Harvest Strategy 3, and it was considered undesirable to have a discontinuity in fishing intensity once the limit reference point was breached. In addition, participants of the 3rd MSE Workshop intended to evaluate Harvest Strategy 3 rather than Harvest Strategy 1.	
10	Harvest Strategy 2 should be removed from further consideration because the absence of a threshold reference point required a large drop in fishing intensity once the limit reference point was breached and it performed poorer than Harvest Strategy 3 with F50 or F40 in terms of Management Objective #2.	
11	The candidate target reference point of F30 should be removed from further consideration because it was the worst performing in terms of	

	Recommendation	Progress
	Management Objectives #1, 2, and 5, and had a similar performance to F40 for Management Objective #4.	
12	The candidate target reference point of F0204 should be removed from further consideration because the actual fishing intensity of this reference point varied substantially between productivity scenarios. It also performed poorer than TRP40 and TRP50 for Management Objectives #1, 2, and 5.	
13	A stricter risk level of 90% (rather than 50%) should be used when evaluating the risk of breaching the candidate limit reference points of SSB7.7% and SSB14% (i.e., the LRP is breached if the probability of being above the limit reference point drops below 90%). Given that the candidate limit reference point of SSB20% is relatively conservative, a risk level of 80% was considered appropriate for that reference point. This risk level should be calculated in the same way as is currently done in NPALB stock assessments, by using future projection software over a period of 10 years and calculating the probability of breaching the limit reference point.	● New HCRs tested in 2nd round of MSE use a 90% or 80% risk level of breaching candidate SSB_{limit} (See Table 1). ● Code was modified to calculate the probability of breaching the SSB_{limit} using the projection software (2017 SA version) rather than the MLE estimate from EM output as in the 1st round of MSE. ● The projection software is run for 10 years with 1,000 iterations within the MSE loop. The uncertainties in the projection software are derived from recruitment variability and initial N at age based on the CV of SSB.
14	In addition to harvest control rules where all fisheries are managed by total allowable effort (TAE) or total allowable catch (TAC), there should be an evaluation of harvest control rules where surface fisheries (i.e., Japan pole-and-line and EPO surface) are managed by TAE and all other fisheries are managed by TAC.	Code was modified to include a mixed TAC/TAE option as follows. ✓ Compute the overall TAC using the fishing intensity (1-SPR) according the the status of the SSB relative to the reference points (as per TAC rule). ✓ The TAC is split across fleets according to the pre-agreed upon allocation (1999-2015 catch ratios) and is kept constant between assessments for the non–surface fleets. ✓ For the EPO surface fleet and the Japanese pole-and-line fleets the exploitation rate is kept constant between assessments, but the catch varies given the biomass from the OM.

	Recommendation	Progress
15	The levels of fishing intensity should be limited by the historical (1997 – 2015) levels (or distributions of historical fishing intensity levels) achieved by the NPALB fisheries. However, if these levels of fishing intensity are not high enough to compare performance of threshold and limit reference points, low productivity scenario should be used in the operating models to evaluate these reference points, where appropriate.	Code was modified to set F as a random F sampled from historical Fs rather than F
16	A future fishing effort scenario where an unmanaged new fishery is removing an increasing amount of unreported catch should be evaluated to understand how large amounts of unreported catch may affect the performance of the harvest control rules.	Code was developed to include this as a robustness scenario. The new fishery has the characteristics of the Chinese longline fleet operating in area 2 and 4.
17	Implementation error distribution should include both positive and negative errors.	Both positive and negative errors were included as 1.05 + N(0, σ=0.05).
18	The ISC ALBWG should continue working on the MSE process for a 2nd round because the results presented at the 4th ISC ALB MSE Workshop were useful for understanding the trade-offs and potential performance of candidate reference points and harvest control rules. However, some candidate reference points and harvest control rules developed at the 3rd MSE Workshop were not evaluated in time due to computer resource limitations. Therefore, the workshop participants developed a focused list of candidate reference points and harvest control rules to be examined for the 2nd round of MSE.	• Three LRPs ($SSB_{20\%,F=0}$, $SSB_{14\%,F=0}$, and $SSB_{7.7\%,F=0}$) requested by the managers and stakeholders were also evaluated for further consideration of LRPS. • This will be discussed at the 2nd round of MSE WS.
19	Pending approval by the ISC Plenary and resolving potential conflicts with the workload of the ALBWG, results of the 2nd round of MSE should be presented at the 5th ISC ALB MSE Workshop as soon as possible, and no later than late 2020.	It may be a good idea to distribute the preliminary report to the WS participants prior to the WS even though the ISC Plenary has not reviewed it. **The WG thought it was a good idea and recommended doing so as long as the ISC Plenary agrees. The WG Chair agreed to ask the ISC Chair about this matter in the near future.**

	Recommendation	Progress
20	Given the timeline and previous computer resource limitations, it is important that improved computer resources be available for the 2nd round of ISC ALB MSE.	Some additional resources at NOAA were available until early 2020. Results will be completed by late 2020 as planned.
21	The adequacy of 45 replicates per "run" (i.e., each OM-MP combination) should be examined to a) determine if the rank order of each run for each performance indicator was stable as more replicates are added; and b) determine if and how the value of each performance indicator varied with increasing numbers of replicates.	• **The WG recommended using broader risk classes based on the Table 4 from the 2nd ISC NPALB MSE workshop (attachment 5 in http://isc.fra.go.jp/pdf/ISC16/ISC16_Annex_08_Report_of_the_ALBWG(Apr2016).pdf) to group performance metrics based on probabilities.** For metrics not based on probabilities, it is suggested that the metric be split in classes prior to ranking. • **The WG also agreed with the presenter that the 70 iterations for the 2nd round of MSE was adequate.** If certain iterations of the runs did not converge, the same set of converged iterations should be used to compare the candidate HCRs , noting which HCRs failed to complete the 70 iterations.
22	The relationship between how effort is modelled in the MSE operating models (i.e., fishing intensity) and effort in the real world should be examined by the ALBWG and included in the future round of MSE to help managers and stakeholders, if possible.	MSE fishing intensity was compared to real world effort.
23	Economic expertise, even though now is not available for the ALBWG, may be needed for future round of MSE since economic aspects are important incentives for the fishery industry.	This is related to Rec. #6.

Table XX. List of tasks that will be presented at next ISC MSE WS in December, 2020.

No.	Tasks
1	Compare the MSE results between the different versions of the projection software and the MLE estimate for a small subset of runs
2	Prepare information for the translation between fishing intensity and effort. This would include an examination of the relationship between observed catch and effort as well as the model-based effort metric (exploitation rate derived from SPR-based fishing intensity) and real-world effort.
3	Prepare worm plots, pie charts and violin plots to show results of PIs to help illustrate the variability in the MSE results (**See Figure 2**).
4	In plots of simulated fishing intensity over time show 'current' fishing intensity from both the 2017 (2012-2014) and 2020 assessments (2015-2017) in addition to average hisotrical fishing intensity.
5	Prepare explanations for the new algorithm for generating exploitation rates and catches for the 2nd round of MSE to avoid confusion from the 1st round of MSE.
6	Prepare table of PIs with the results for each performance metric both as numbers and color-coded (**See Figure 3**).
7	Prepare additional PIs such as P(management action) and P(SSB > $SSB_{7.7\%,F=0}$ from the OM).
8	Label Performance Metric 1 "odds of not breaching the LRP" rather than "conservation risk".
9	Compare the impact on performance metrics results of using a $SSB_{7.7\%,F=0}$ based on dynamic SSB_0 versus using equilibrium SSB_0.
10	Improve plot describing management actions for HCRs (**Table1**) for explanation of TAC control.
11	Extract example trajectories from set of results to be used in presentation of results using a "storytelling approach".
12	In the table describing OMs include additional columns for biological plausibility and model fit.
13	In 2nd round of MSE report include appendix tables with the performance metrics calculated across the four reference scenarios (1, 3, 4, and 6) as well as separatly for each scenario.

14	Consider how the need for meta-rules should be communicated to mangers and stakeholders, and consider some potential examples.
15	Consider possible usage of "ShinyApp" to show MSE results in more effective manner by the 5th MSE workshop (Presented at early 2021 workshop for managers and stakeholders).

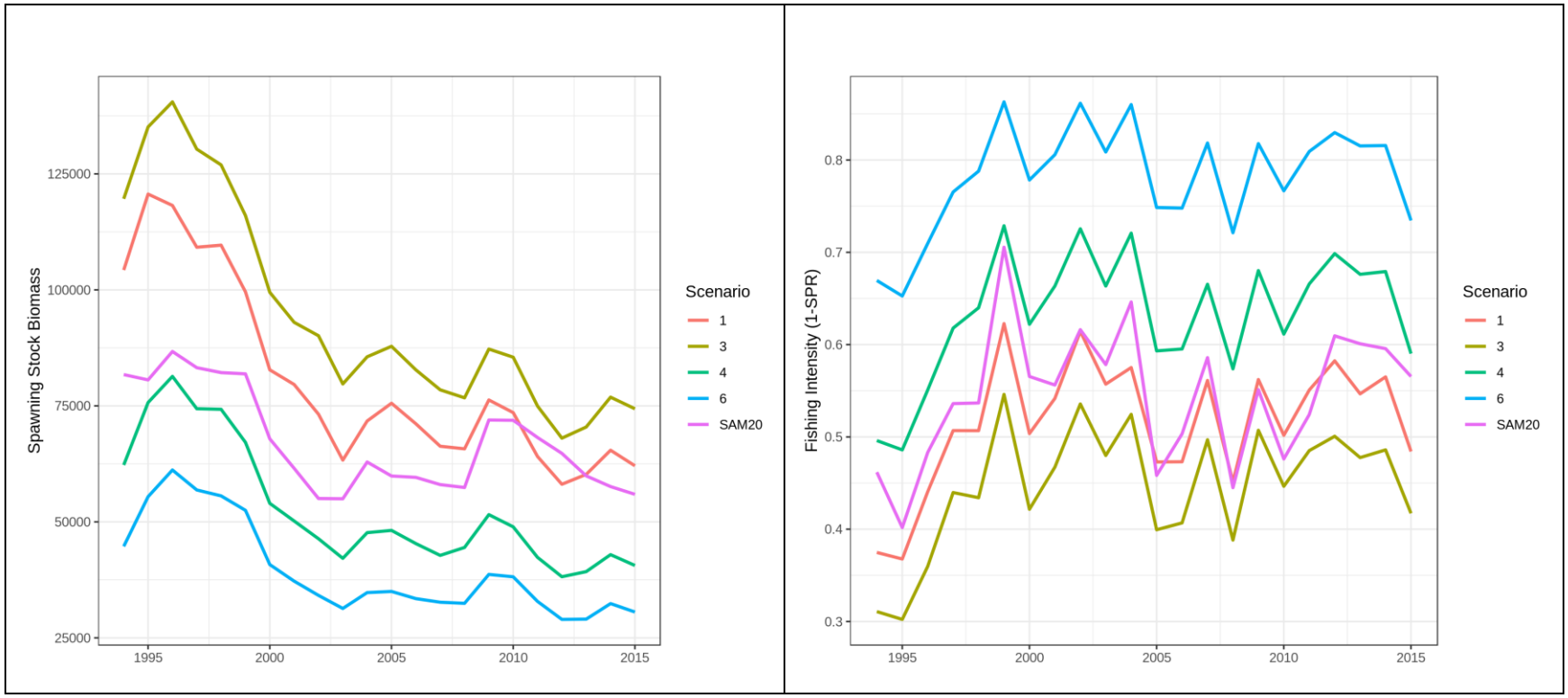


Figure 1. Trends in spawning stock biomass and fishing intensity (1-SPR) for the four reference uncertainty scenarios used for the 2nd round of ISC NPALB MSE and the 2020 stock assessment (pink line labeled as "SAM20").

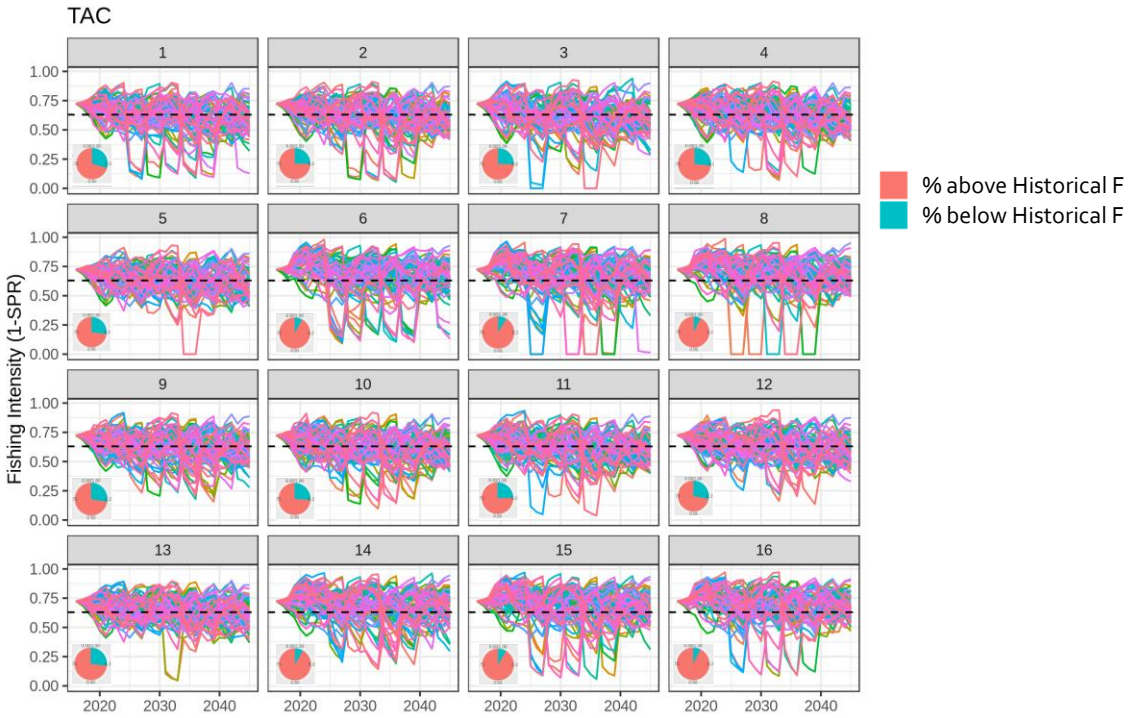


Figure 2. Example of worm plots with pie charts showing the simulated time series of fishing intensity as 1-SPR for HCRs 1to 16 under TAC control for scenario 1. Each panel presents the results for the labeled HCR. Colored trajectories represent separate iterations differing in simulated random recruitment deviates, EPO age-based selectivity deviates, and implementation error. The pie chart shows the % of years across iterations for each HCR that were above or below historical average fishing intensity. Historical average fishing intensity was calculated over the conditioning period of 1993-2015.

	Reference				Low Productivity			
HCR	PM1	PM2	PM4	PM5	PM1	PM2	PM4	PM5
1	0.99	0.68	0.69	0.73	0.66	0.71	0.59	0.58
2	1	0.68	0.71	0.79	0.83	0.71	0.59	0.67
3	1	0.67	0.72	0.82	0.94	0.70	0.61	0.63
4	1	0.68	0.73	0.87	0.83	0.70	0.60	0.66
5	1	0.67	0.74	0.92	0.93	0.69	0.63	0.68
6	0.98	0.62	0.83	0.66	0.74	0.66	0.61	0.47
7	1	0.62	0.84	0.75	0.88	0.64	0.63	0.44
8	1	0.62	0.87	0.82	0.87	0.71	0.63	0.49
9	0.99	0.68	0.71	0.77	0.64	0.70	0.61	0.63
10	1	0.67	0.72	0.82	0.83	0.71	0.59	0.61
11	1	0.68	0.72	0.84	0.95	0.70	0.61	0.69
12	1	0.68	0.73	0.87	0.83	0.71	0.62	0.7
13	1	0.67	0.75	0.91	0.94	0.7	0.62	0.69
14	0.98	0.62	0.84	0.72	0.73	0.64	0.62	0.48
15	1	0.62	0.86	0.78	0.87	0.65	0.62	0.46
16	1	0.62	0.87	0.84	0.88	0.64	0.63	0.5

Figure 3. Example of table with the results for each performance metrics by HCR and scenario as numbers and color-coded according to the risk classes proposed in Table 4 from the 2nd ISC NPALB MSE workshop (attachment 5 in http://isc.fra.go.jp/pdf/ISC16/ISC16_Annex_08_Report_of_the_ALBWG(Apr2016).pdf)

Attachment 1
List of Participants

Canada	
Chris Rooper Pacific Biological Staion 3190 Hammond Bay Road Nanaimo BC V9T 6N7 Canada Chris.Rooper@dfo-mpo.gc.ca	Jennifer Shaw Fisheries and Oceans Canada 200 Kent St. Ottawa, ON, Canada K1A 0E6 Jennifer.Shaw@dfo-mpo.gc.ca
Zane Zhang Pacific Biological Staion 3190 Hammond Bay Road Nanaimo BC V9T 6N7 Canada Zane.Zhang@dfo-mpo.gc.ca	**Chinese-Taipei** Chiee-Young Chen National Kaohsiung University of Science and Technology, Department of Marine Environmental Engineering No. 142, Hai-Chuan Road Kaohsiung, Taiwan chency@nkust.edu.tw
Japan	
Yoshinori Aoki Fisheries Research Institute, Japan Fisheries Research and Education Agency 5-7-1 Orido Shimizu, Shizuoka Shizuoka 424-8633 JAPAN aokiyoshinori@affrc.go.jp	Hirotaka Ijima Fisheries Research Institute, Japan Fisheries Research and Education Agency 2-12-4 Fukuura, Kanazawa, Yokohama Kanagawa 236-8648 JAPAN ijima@affrc.go.jp
Hidetada Kiyofuji (WG Chair) Fisheries Research Institute, Japan Fisheries Research and Education Agency 2-12-4 Fukuura, Kanazawa, Yokohama Kanagawa 236-8648 JAPAN hkiyofuj@affrc.go.jp	Naoto Matsubara Fisheries Research Institute, Japan Fisheries Research and Education Agency 2-12-4 Fukuura, Kanazawa, Yokohama Kanagawa 236-8648 JAPAN matsubaranaoto@affrc.go.jp
Yuichi Tsuda Fisheries Research Institute, Japan Fisheries Research and Education Agency 5-7-1 Orido Shimizu, Shizuoka Shizuoka 424-8633 JAPAN	

u1tsuda@affrc.go.jp	
USA Steven Teo (WG Vice Chair) NOAA/SWFSC 8901 La Jolla Shores Drive La Jolla CA 92037 USA steve.teo@noaa.gov	Desiree Tommasi (MSE Specialist) NOAA/SWFSC 8901 La Jolla Shores Drive La Jolla CA 92037 USA desiree.tommasi@noaa.gov
Inter-American Tropical Tuna Commission Carolina Minte-Vera 8901 La Jolla Shores Drive La Jolla CA 92037 USA cminte@iattc.org	Juan Valero 8901 La Jolla Shores Drive La Jolla CA 92037 USA jvalero@iattc.org
The Pacific Community (SPC) Robert Scott CPS – B.P. D5 | 98848 Noumea New Caledonia robertsc@spc.int	Nan Yao CPS – B.P. D5 | 98848 Noumea New Caledonia nany@spc.int

Attachment 2

ALBACORE WORKING GROUP (ALBWG)

INTERNATIONAL SCIENTIFIC COMMITTEE FOR TUNA AND TUNA-LIKE SPECIES

IN THE NORTH PACIFIC OCEAN

MSE UPDATE WEBINAR

31 Aug – 3 and 8 Sep 2020 (Eastern Pacific)
1 – 4, 9 Sep 2020 (Western Pacific)
Draft Agenda

Time:

JAPAN and KOREA:	09:00 – 13:00 (break: 10:30)
CHINESE TAIPEI:	08:00 – 12:00 (break: 09:30)
NOUMEA:	11:00 – 15:00 (break: 12:30)
CANADA, USA and MEXICO	17:00 – 21:00 (break: 18:30)

August 31 (Mon) and September 1 (Tue)

1. Opening and Welcome
   1.1 Meeting Protocol
   1.2 Review and adoption of Agenda
   1.3 Assignment of Rapporteurs
2. Overview of recommendations from 1st MSE WS for NPALB
3. Main changes from 1st MSE to meet those recommendations

September 1 (Tue) and 2 (Wed)

4. Overview of TAC results
5. Comparison of results with different options for when SSB > threshold reference point

September 2 (Wed) and 3 (Thu)

6. Overview of results for mixed TAC/TAE control and comparison with TAC control
7. Main conclusions from preliminary results presented by MSE specialist

September 3 (Thu) and 4 (Fri)

8. Overview of assessment error and potential for development of faster MSE code for exploratory analyses
9. Others (new assessment as OM)

September 8 (Tue) and 9 (Wed)

8. Clearing of Meeting Report
9. Administrative Matters
   9.1 Review MSE timeline and workplan
10. Adjournment

Attachment 3
List of Presentations

Number	Title and Authors	Availability
Presentation	4th ISC ALB MSE Workshop recommendations and how they were addressed in the 2nd round of MSE D. Tommasi	Contact the author
Presentation	ISC ALB MSE Round 2 Preliminary TAC Results D. Tommasi	Contact the author
Presentation	ISC ALB MSE Round 2 Preliminary Overview of results for the mixed TAC/TAE control strategy D. Tommasi	Contact the author
Presentation	ISC ALB MSE Round 2 Preliminary overview of Assessment Error D. Tommasi	Contact the author

Attachment 4
Workplan

Date	Location	Task/Event
September 21 – 23, 2020	Webinar	IATTC SAC
October 8, 2020	Webinar	NC16
Nov 30 -Dec 3, 8 2020 (EPO) Dec 1 – 4, 9 (WPO)	Webinar	ALBWG: review on MSE progress and finalize report
December	Webinar	WCPFC17
February or March 2021	TBD	5th ISC MSE Workshop: review results from 2nd MSE for NPALB
July 11 – 19, 2021	HI, USA	ISC Plenary: Report results from 5th ISC MSE Workshop
August 11 -19, 2021	Palau	WCPFC SC17
September 2021	TBD	NC17
December 2021	TBD	WCPFC18