SCIENTIFIC COMMITTEE

NINETEENTH REGULAR SESSION

Koror, Palau

16-24 August 2023

Report from the SPC Pre-assessment Workshop - April 2023

WCPFC-SC19-2023/SA-IP-01

20 July 2023

Report compiled by Paul Hamer

Oceanic Fisheries Programme (OFP),
Pacific Community (SPC), Noumea, New Caledonia

Table of Contents

Pre-assessment Workshop Overview 4

DAY 1 – 2023 WCPO yellowfin tuna stock assessment 5

Stock assessment overview presentation and peer review summary 5

Conceptual model development to inform spatial stratification 7

Data inputs 9

Biology 10

DAY 2 – External presentations, 2023 WCPO bigeye assessment, CPUE analysis and model developments (yellowfin and bigeye) 11

External and WCPFC project presentations 11

2023 WCPO bigeye assessment 12

Stock assessment overview presentation 12

Conceptual model development to inform spatial stratification 12

Data inputs 14

The tagging data summary 15

CPUE analyses 16

Biology 17

Model development - yellowfin 17

DAY 3 – MFCL, MSE/harvest strategies, 2022 skipjack assessment follow-up work, TARP – tuna assessment research plan 18

Model development – bigeye 18

Multifan -CL 19

Next generation tuna assessment model 19

Management Strategy Evaluation 22

Follow-up work on the 2022 skipjack assessment 24

Tuna Assessment Research Plan 26

DAY 4 – SEAPODYM, Silky shark assessment, Project 113 – review of assessment uncertainty characterisation, discussion of assessment models/spatial structure/uncertainties and diagnostics 27

SEAPODYM 27

Silky shark assessment 30

Project 113 – review of assessment uncertainty characterisation 32

Diagnostics 35

Discussion of assessment models/spatial structure/uncertainties 37
Post-meeting follow-up 41
References 41
Appendices 42
APPENDIX 1: Agenda 42
APPENDIX 2: List of participants 48
APPENDIX 3: Terms of Reference 50

Pre-assessment Workshop Overview

To help guide stock assessment and related modelling work and analyses for the Western and Central Pacific Fisheries Commission (WCPFC), the Oceanic Fisheries Programme (OFP) of the Pacific Community (SPC) has sought input from regional stock assessment scientists, consultants and representatives from regional fisheries organisations that are part of the WCPFC, through the SPC pre-assessment workshop (PAW) process. The fifteenth PAW was held from the 25th – 28th April 2023. The meeting was held under a hybrid format, with 37 people present in Noumea, including 19 from SPC and 18 scientists who travelled from external organisations, and approximately 33 people from external organisations attending online, for all or parts of the workshop. Fifteen organisations were represented, from across at least 10 countries,

Paul Hamer (OFP, SPC) chaired the meeting. The meeting agenda focused primarily on:

- approaches for the 2023 stock assessments of WCPO yellowfin and bigeye tuna, including an overview of the 2020 yellowfin assessment peer review,
- follow-up work being done on the 2022 WCPO skipjack tuna assessment,
- developments to the MULTIFAN-CL modelling framework and future tuna assessment model options,
- technical developments in WCPO Management Strategy Evaluation and related activities to progress the WCPFC harvest strategy work plan,
- the SEAPODYM modelling framework,
- development of the WCPO silky shark assessment,
- the Tuna Assessment Research Plan, and;
- the WCPFC project to review approaches for ensemble model development and characterising uncertainty in WCPFC stock assessments,
- other topics including, bomb radiocarbon age validation work, length-weight conversion work, caging experiments to study seasonal growth dynamics of yellowfin tuna, diagnostics for stock assessments models.

The full planned agenda is in Appendix 1, and list of attendees is in Appendix 2.

Presentations were invited from all participants, with the majority made by SPC staff or consultants working with SPC. Two external presentations were provided. The meeting operated under the terms of reference provided in Appendix 3.

This report briefly describes the various presentations made and issues discussed by participants, and specific suggestions made. The report does not attribute comments to countries or individuals except for those that provided presentations and where the comment related to the agreement to provide data or to undertake particular analyses. The relevant stock assessment scientists will consider the recommendations provided in this report. The extent to which suggestions can be explored and/or incorporated into the stock assessments prior to WCPFC SC19 will be constrained by the available time and requirement to prioritise some aspects over others, which will be at the discretion of the SPC stock assessment scientists. Ultimately the final decisions on model development, data inclusion, and approach

to characterising uncertainty are made by the SPC-OFP assessment team, or the SPC-OFP assessment team in consultation with external contractors involved in the assessments or supporting work.

The outcomes of this meeting will be reflected in the stock assessment related papers submitted to WCPFC-SC19. Copies of presentations prepared by SPC can be provided on request from paulh@spc.int.

2023 SPC Pre-assessment workshop in-person attendees at SPC Noumea headquarters.

Back row: Nick Davies, Moses Mataika, Joe Scutt Phillips, Nicholas Ducharme-Barth, Tim Adams
Middle row: Peter Williams, Raijeli Natadra, Finlay Scott, Leyla Knittweis, Yoshinori Aoki, Ren-Fen Wu, Keisuke Sato, Yuichi Tsuda, Nan Yao, Jemery Day, Inna Senina, Arni Magnusson, Thom Teears, Patrick Lehodey, Phil Neubauer, SungKwon Soh
Front row: Laura Tremblay-Boyer, Eric Chang, Keith Bigelow, Hidetada Kiyofuji, Rob Scott, Rob Campbell, Francisco 'Curro' Abascal, Claudio Castillo-Jordan, Takaaki Hasegawa, Paul Hamer, Graham Pilling, John Hampton. Missing: Jed Macdonald, Sam McKechnie.

DAY 1 – 2023 WCPO yellowfin tuna stock assessment

Stock assessment overview presentation and peer review summary

Day 1 focused on the stock assessment of yellowfin tuna (*Thunnus albacares)* in the Western and Central Pacific Ocean (WCPO) being led by Arni Magnusson (OFP-SPC). The session began with Arni providing an overview of the 2020 assessment that outlined the outcomes of that assessment, noting the large change in stock status between the 2017 and 2020 assessments, and the issues that led to the peer review of that assessment. Following this Paul Hamer (OFP-SPC) provided a summary of the independent peer review of the 2020 yellowfin assessment, noting the process of the review, key recommendations, and the areas on

which the 2023 stock assessment is aiming to focus on improvements. It was noted that the peer review provides direction for improvements for both the yellowfin and bigeye assessments and, potentially, also for other assessment of WCPO stocks. The peer review panel was thorough, with 42 requests for information or analysis made to the SPC assessment team. The panel report can be found at https://meetings.wcpfc.int/node/18561. Many recommendations were made by the panel, and these were summarised in the presentation, noting that not all recommendations can feasibly be explored or implemented for the 2023 assessments and that some would require further detailed work and potential modifications to MFCL.

Aspects from the peer review that were noted for consideration in the 2023 assessments included:

- Developing a conceptual model for yellowfin and use this to consider revising/simplifying the model spatial stratification.
- Improving fits to length composition data, better understand the size composition data and consider triaging poor/unrepresentative data.
- Size data should be weighted using a metric that reflects the likely information content of the data.
- Work on the CPUE analysis, explore various alternatives that result in more plausible relative abundance among northern, southern, and equatorial model regions.
- See if length-based selectivity can be implemented in MFCL (Note: this has been explored prior to the workshop and is not feasible for 2023).
- Specify overdispersion in CPUE as an additive rather than multiplicative factor.
- Implement conditional age at length as the default, with internal growth estimation.
- Explore age specific movement.
- Improve the diagnostics suite, including definition of convergence, Hessian status.
- Avoid nested steps in the stepwise analysis, single steps at a time.
- Clear documentation of model developments' and improved systems for repeatability.
- Include tables of model parameters.

The peer review also noted that the SPC analysts will need to prepare for a transition to an alternative assessment platform given that recommended enhancements to MULTIFAN-CL may be difficult to develop due to limited developer resources (one person) and the retirement of the lead developer, Dave Fournier.

Regarding the peer review, the workshop asked whether there was any key issue or change driving the more optimistic view of stock status from the 2020 yellowfin assessment. SPC noted that the peer review felt that each of the individual changes made in the 2020 assessment were considered appropriate, with the cumulative effect leading to the more positive stock status outcome. However, for some steps in the bridging analysis multiple things were changed at the same time making it difficult to disentangle the implications, this was noted for treatment and updating of tag data, which appeared related to some of the larger changes in the stepwise analysis.

The workshop noted the review panel's comments on avoiding 'nested' stepwise changes, where multiple model settings and data changes were made within an individual step, observations on growth assumptions and fits to the size data, and the issue of effort creep in the longline fishery for which CPUE is a key influence on stock status. With regards to effort creep, it was noted this was an area requiring further work that would benefit from being the subject of a specific SC project that involved distant water fishing nation colleagues to identify available information and approaches to estimate plausible effort creep scenarios for long line fisheries in the WCPO. A dedicated project on longline effort creep was recommended by the peer review panel.

The workshop noted that while the peer review panel did not recommend a change in the approach for natural mortality at age (M), it was noted that that other relevant estimates of base/asymptotic values for M of yellowfin tuna are available in the recent review paper by Hoyle et al., (2023). Given the influence of M on assessment outputs, this should be considered for an axis of the uncertainty grid. Other options for parameterising M at age were noted from the recent Tuna Assessment Good Practices CAPAM meeting, including the Lorenzen formulation, that could also be considered.

Conceptual model development to inform spatial stratification

The peer review recognised that that 9 region spatial structure was overly complex for the information in the data. One of the recommendations of the peer review was to consider simplification options, including reviewing the spatial stratification based on a conceptual model of yellowfin tuna population and fishery structure in the Pacific. A series of presentations were provided to explore biological, genetic, otolith, and tagging information (Paul Hamer, SPC), SEAPODYM predictions of distributions of yellowfin tuna life-stages in the Pacific (Inna Senina, SPC), fishery size composition data (Jo Potts, SPC), and fishery CPUE patterns (Jed Macdonald) to provide a conceptual basis for a revised spatial stratification of the yellowfin assessment in the WCPO. These presentations provided support for a simpler spatial stratification than the current 9 region stratification. This was discussed by the workshop, and several options including a 4 and 5 region stratification were suggested to explore further.

The workshop noted that the peer review panel supported the simplification of the yellowfin assessment with the aim of aiding model stability, and that it was not essential to keep yellowfin identical to bigeye for assessment purposes.

With regards the summary of the background information available, much of which is reviewed in Moore et al. (2020), the workshop noted that a high percentage of the catch is also in the equatorial region, and that it could be useful to explore a model that only considers the equatorial region. It was noted that an equatorial region model was performed in 2007 to estimate an alternative growth pattern that was evaluated as a plausible growth sensitivity. SPC noted that the evaluation of an equatorial model was a suggestion by the peer review panel as a potential sensitivity to explore the influences of removing the northern and southern regions and their possible buffering effects, which is different to the 2007 objective.

In relation to simplified spatial structures, issues were raised on tag mixing assumptions, especially for small model regions were lots of tags have been released and recaptured (i.e., regions 8 and 9). There was

some discussion on region 9 and that this region is not supported to be retained for yellowfin at least. A stepwise approach to moving from the current 9 region model to a simpler model was recommended. The potential to remove region 9, along with the influence of the data from the associated tags released within that region, was noted. In turn, the potential to remove the line between regions 1 and 2 and 5 and 6 in the temperate regions was supported by some. The potential to maintain more complex fleet structures through a fleets-as-areas approach was also noted.

Other discussion points included the suitability of the application of SEAPODYM predictions for informing stock assessment assumptions (covered more later in the workshop), different movement rates with age and the possible approach in MFCL to implement such differences, the need to conduct further size composition analysis (i.e. regression tree analysis) on the longline weight frequency data to compliment the length composition analysis for informing fishery stratifications (to be followed up by Jo Potts, OFP-SPC), and various other aspects of the size composition analysis. There was some discussion on the purpose of the length composition analysis, and it was noted that the IATTC use it to define fleets for a fleets as areas approach and not so much for identifying underlying population variation, but the data are influenced by selectivity, underlying population structure and catchability, plus also sampling methodology/variability. Therefore, there needs to be some care in interpreting results, look at the plots of the distributions, ideally you should end up with fishery groupings that produce length frequencies with double normal distributions. The workshop also suggested to explore analyses of specific fleets (i.e., Japan).

With regard to the CPUE time series analyses there were constructive suggestions to improve this type of clustering analysis whereby groups are clustered contiguously. It was suggested that spatial patterns in the areas where fleets fish that go into the analysis might also be influencing the clustering rather than underlying population dynamics.

Residency levels in region 8 were discussed, and it was noted that SEAPODYM doesn't predict high residency in region 8.

Information on analysis done by Japan using Japanese longline CPUE and size compositions was shown to the workshop and provided further support for removing the longitudinal split at 170°E, but maintaining separation of northern, equatorial and southern regions.

Overall, there was support for considering simplified spatial stratification for the 2023 yellowfin assessment, that could involve removal of region 9, merging of the two southern (5 and 6) and northern (1 and 2) regions, and exploration of several options involving merging the equatorial regions (3, 4, and 8) (Fig 1). The feasibility of making these changes in terms of developing the corresponding tagging data files or changing the tag mixing rate assumptions or specifying SEAPODYM movement as an assumption to evaluate the spatial impact, would need to be considered. Removing the tag data entirely was viewed as sub-optimal. The additional workload these explorations imply was also noted and the amount of exploration of alternative spatial stratifications will depend on time constraints.

There was general discussion on how to broach a change in spatial structure for the assessment, particularly within the context of the SC. One perspective was that if the alternative structures are seen

as viable alternative hypotheses but have different management advice, they might be best included as an axis in the structural uncertainty grid. If they don't differ, make a case for choosing one. However, it was also pointed out that alternative structures might result in quite different dynamics and responses to other parameter uncertainties. This might require consideration as there could be different parameter/structural uncertainties that apply to the alternative spatial structures, requiring a hierarchical tree type approach rather than the orthogonal grid for characterising management uncertainty.

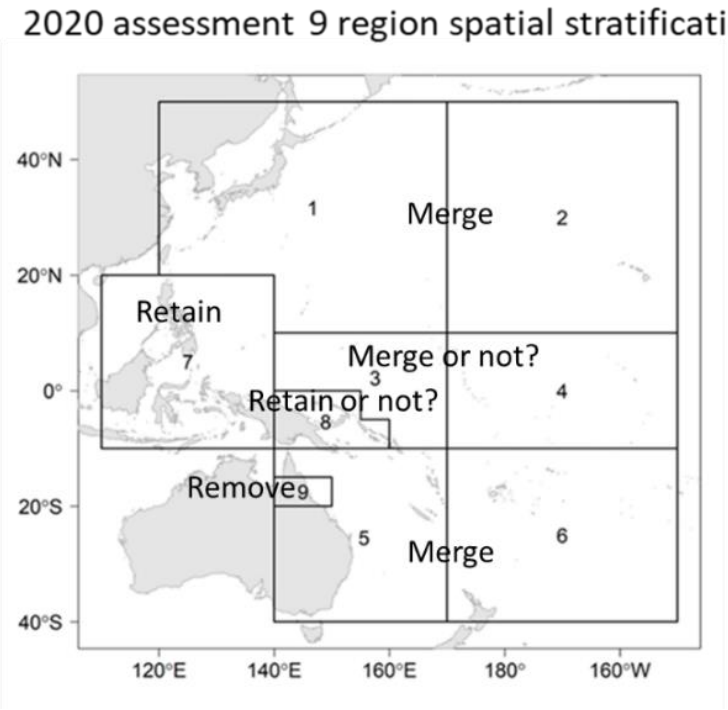

Figure 1. Map showing the 9 region spatial stratification applied in the 2020 yellowfin assessment and the options considered for simplification based on the presentations to inform a conceptual model of population and fishery structure.

Data inputs

Data inputs and related analysis were presented (with group discussion) including fishery data/size composition (Thom Teears, SPC), implications of reduced observer coverage (particularly in relation to purse seine species composition estimates) and size composition re-weighting (Tom Peatman, consultant to SPC).

Questions arose on the patterns of data available by fishery and flag, particularly in the most recent years where data collection appears to have been reduced by the COVID-19 pandemic and related low observer coverage or port sampling. The workshop noted that some data are currently being loaded as part of recent data updates. Patchiness, temporal variation and bi/multi-modality of size composition data were noted and that these distributions present difficulties for the models to fit. Restructuring of fisheries and/or data filtering processes prior to incorporation within the assessment may improve some of these issues. Remaining patterns in size data might be better dealt with through selectivity time blocks.

With regards to spatio-temporal weighting of size composition data, the potential to implement stronger data filtering rules was noted to examine whether it removed apparent noise in the size data due to low (assumed unrepresentative) sample sizes. The workshop noted the potential to standardise the size data, including factors such as flag to identify whether this identified different mean sizes, which might be driven by market pressures, management interventions (e.g., FAD closure period), set misidentification or other operational or sampling factors. It was also suggested that it is worth considering length versus weight composition data for the longline fisheries. Previous assessments have focused on weight compositions, mostly available from the Japanese fleet, but more recently the length composition data has become more comprehensive and with a better fleet coverage in relation to catches since the late 1990s.

Noting the issues arising from COVID-19 restrictions on observer coverage, the potential impacts on the assessments were discussed. Species composition estimates for the purse seine catch derived from observer estimates were compared to cannery data for the US fleet, noting that at least for this comparison, the COVID-19 impacted observers-based estimates remained relatively consistent with the cannery estimates. Incorporation of this uncertainty within the assessment, for example by bootstrapping estimates, could be considered, although this might not be possible for all fisheries. The workshop noted the estimates from cannery data that highlighted that the recent higher percentages of bigeye in the purse seine catch from US cannery data would not be representative of the longer-term composition, since they are influenced by US vessels fishing further east in recent years under the La Nina conditions.

Biology

Conditional age at length data and growth variation were presented by Arni Magnusson (SPC-OFP). He noted that the available age at length information showed relatively little variation in the growth rates between the different regions sampled in the WCPO for the age range that comprises the bulk of the population, and also little difference between sexes. However, the otolith conditional age at length samples were collected between 2009-2018, with some years having very few samples. They were also mostly collected from the region between 10°N and 20°S.

The workshop noted that the peer review panel raised the issue of growth uncertainty, but noting the general similarities between growth estimates, recommended the conditional age at length approach as used in the diagnostic model of the 2020 assessment be used, as there is potential for the age at length data to be unrepresentative of the entire population and model time period. This should allow the model to better estimate size compositions. However, uncertainty in the growth curve may still be considered as an axis within the uncertainty grid.

With regards to natural mortality (M) it was noted that the peer review panel indicated that the natural mortality with age formulation used in the 2020 assessment should be continued for the 2023 assessment, and the workshop noted the likely influence of uncertainty in M on model outputs. As noted earlier, other values of base or asymptotic M are available (i.e., Hoyle et al. 2023 paper) for consideration within the assessment, while alternative functional forms of M at age could be considered based on the recent CAPAM discussions, i.e., Lorenzen. Suggestion was made to explore alternative base M values with a likelihood profile on M.

Day 1 was concluded with a summary by Arni of the significant work done to improve the tools and processes for increasing the efficiency and repeatability of the SPC stock assessments. These include improving user friendliness of aspects of MFCL outputs, additional features in FLR4MFCL, development of R Shiny Apps. for exploring/comparing model results and diagnostics, more efficient approaches for submitting jobs to CONDOR, scripts for increased efficiency and repeatability of R related work, and standardised directory structures and R scripts.

DAY 2 – External presentations, 2023 WCPO bigeye assessment, CPUE analysis and model developments (yellowfin and bigeye)

External and WCPFC project presentations

Day 2 began with three stand-alone presentations on project work with relevance to the stock assessments. The first presentation was by Allen Andrews (University of Hawaii/Upssala University) on the final outcomes of WCPFC project 105: bomb radiocarbon age validation for yellowfin and bigeye tuna in the WCPO, which indicated that annual increment formation interpretation was now validated for yellowfin and bigeye tuna up to 13 years of age. The maximum age of these species in the WCPO remains unknown. The second presentation from Kei Okamoto of the Japan Fisheries Research and Education Agency (FRA) covered the study of growth of individual yellowfin tuna reared in sea cages at Amami oshima, highlighting the seasonality of growth and potential implications for size at age inferences from traditional growth curves. The third presentation from Jed Macdonald (SPC) and Simon Hoyle (NIWA) provided an overview of WCPFC project 90: Better data on fish lengths and weights for scientific analysis. The project provides updated length-weight and weight-weight conversion factors for the 2023 assessments based on an enhanced and improved data set. The presentation also provided evidence for spatial variation in these relationships.

In relation to the radiocarbon age validation work, the need to better understand maximum ages for all WCPO tuna species was noted as this helps to define natural mortality. The potential to apply the bomb radiocarbon approach to larger skipjack caught in the eastern part of the region (where they are taken by longliners) was also mentioned as worth pursuing to assess maximum age of skipjack which is a major uncertainty in the knowledge of skipjack life history.

It was asked how a seasonal growth pattern might be dealt with in MFCL, but it was noted that seasonal growth could not be easily modelled within MFCL. One option may be to treat quarters as years with deviates for each age to account for seasonal growth, but this would have to be fixed across the entire model time period. If seasonal growth is primarily temperature driven it was suggested that the seasonal growth variation may also be less evident in the lower latitude locations, where the bulk of the stock occurs, or away from land where ocean temperature is less seasonally variable. The workshop was interested in comparing the observed seasonal growth in the Amami sea cages with more equatorial open ocean conditions.

The length-weight, weight-weight conversion factors indicated the need to improve the spatial and size coverage of the data available. General Santos has been a focus of project 90 because of access to whole yellowfin and bigeye of very small and large sizes, most importantly the gilled and gutted to whole weight conversions can only be obtained reliably in General Santos. However, large bigeye are limited in Philippines and only from a small area of the WCPO, so the work is now exploring the opportunity to collect data in Samoa (Apia), with new SPC staff in that region. Recently SPC has obtained bigeye length-weight data from Chinese Taipei which has greatly increased the spatial coverage. The expanded data set is being used to explore spatial variations in length – weight relationship and consider what might be driving the observed variation. It was suggested that a way to explore the influence of spatial variation is to convert weights to lengths outside the model using spatially explicit conversions.

The issue of variation in how different flags dressed their tuna was raised (i.e., tail on, tail off etc.), and that conversions of gilled/gutted to whole weight might actually vary among fleets. It was noted that there currently isn't enough data from different flags to understand the extent or implications of this variation.

2023 WCPO bigeye assessment

Stock assessment overview presentation

Jemery Day (OFP-SPC) provided an overview of the 2020 WCPO bigeye assessment summarising the approach used, model settings and changes from the previous (2017) assessment, uncertainties and key outputs. He noted some recommendations and concerns from that assessment including: challenges with growth and fitting to the smallest individuals in region 7 (Indonesia/Vietnam/Philippines), limited data on reproductive biology, the need for improved data for conversion factors, especially weight-weight, CPUE standardisations could better account for potential changes in catchability and influences of oceanography on catch-rates, and that the model is likely over-parametrized and a reduction in complexity was warranted.

Conceptual model development to inform spatial stratification

Similar to for the yellowfin assessment a series of presentations were provided to explore biological, genetic, otolith, and tagging information for bigeye (Paul Hamer, SPC), SEAPODYM predictions of distributions of bigeye tuna life-stages in the Pacific (Inna Senina, SPC), regression tree analysis of fishery size composition data (Jo Potts, SPC), and fishery CPUE time series analysis (Jed Macdonald) to provide a conceptual basis for a revised spatial stratification of the bigeye assessment in the WCPO. The presentations noted the progression of spatial stratification for the bigeye assessment and how it has remained consistent with the WCPO yellowfin assessment since 2005, changing from a 6 region to a 9 region spatial stratification after the bigeye assessment peer review in 2011.

The biological information for bigeye also drew on the review of Moore et al., (2020), and recent publications of digitized summaries of the historic Japanese larval surveys originally reported in Nishikawa et al., (1985), including by Buenafe et al., (2022); Reglero et al., (2014) and recently Ijima and Jusup, (2023), the later who actually have access to the original data.

Overall, the biological information suggested a continuous distribution across the Pacific, with a similar isolation by distance likely from the WCPO to the EPO, perhaps mostly evident from the higher growth and larger sizes of bigeye observed in the EPO, rather than genetics which shows no clear variation. Otolith chemistry suffered the same issue as for yellowfin (juvenile fish and limited spatial sampling coverage). Tagging data suggest a lot of exchange between the EPO and WCPO, especially in the area of the boundary between the two convention areas, so there is no clear stock separation between the EPO and WCPO. Similar to yellowfin, tag-recapture movements mostly involve longitudinal rather the latitudinal movements. Larval distribution was patchier than for yellowfin, with higher density areas in the western WCPO around Philippines, PNG and Chinese Taipei and a few areas in the central equatorial region of the Pacific. There is also a known spawning area that appears to have high spawning site fidelity in the Coral Sea off northern Australia. It was noted that historical work by Japanese scientists on spatial patterns of bigeye spawning intensity did not fit well with the observed larval distributions, and suggested greater levels of spawning in the central and eastern equatorial Pacific than would be expected from the larval abundance data. It was further noted that larval sampling by the historic Japanese program was lower in the eastern compared to western Pacific which may explain the differences, but also that egg/larval survival might be higher in the western Pacific.

The SEAPODYM predictions of life history distribution patterns indicated a more or less continuous distribution of adult life stage across the equatorial region, also in the area of the Kuroshio Current extension from Japan to Hawaii, and the region north of New Zealand. Juveniles are predicted to be more abundant in the Kuroshio Current region and more to the south of the equator, compared to adults.

The regression tree analysis of bigeye length composition data for the longline fishery showed notably greater variation explained than for the yellowfin, with differences between the equatorial and northern and southern regions, and the western equatorial region (Indonesia/Philippines/PNG region), with typically larger fish in the equatorial region. The results for the CPUE analysis were consistent with differences between the northern and southern regions. Overall, these analyses did not lend strong support for longitudinal stock or fishery separations, but there are indications of differences between the equatorial and northern and southern regions, and that at least the dynamics might vary latitudinally.

The workshop discussed aspects of the SEAPODYM predictions, including the prediction of the lower densities of smaller bigeye in the equatorial region but higher densities away from the equator, and high densities of larger bigeye in the EPO that don't closely correspond with the areas of highest longline catches and CPUE. It was noted that the abundances predicted by SEAPODYM still explain the catches, and that bigeye are not specifically targeted by purse seines in the equatorial region, where they are caught almost solely on drifting FADs. It might also be that in the areas of the EPO where SEAPODYM predicts that bigeye are abundant that they are not available to the gear due to oceanographic conditions. There was a view from some workshop participants that the application of the SEAPODYM predictions as information for stock assessments should be done with due caution.

As to tagging data it was noted that there is considerable movement from the central to the eastern Pacific, more so than suggested in the presentations. Also, it was suggested that bigeye are expected to grow larger in the equatorial regions than more northerly or southerly areas, a fixed growth curve could

be contributing to the lack of fit in the previous assessment, but there are very few age/length samples from the northern regions to explore this, this should be improved by collecting more age/length samples from northern regions (i.e., off Japan, Hawaii). Future length-structured models (if developed) could account for variation in growth better than the current age-structured models.

The workshop discussed options for simpler spatial stratification: retaining region 7 (Indonesia/Philippines/Vietnam) supported, consider merging regions 1 and 2 and 5 and 6, but perhaps maintain regions 3 and 8, removing region 9, subject to appropriately dealing with tagging data. There was some discussion on how to use the information on juveniles versus adults for informing spatial stratification, perhaps focussing on adults to guide decisions as they best relate to the target fishery and management. Running a single region fleets as areas model as a sensitivity, or starting from such a model and building complexity from the ground up was suggested. The need to focus on the tractable issues within the time available was noted.

Overall, the workshop discussions on simpler spatial stratification for bigeye resulted in support for simpler structures that were generally consistent with those for yellowfin.

Data inputs

Presentations were provided by Thom Teears (OFP-SPC), Tom Peatman (SPC consultant), Jemery Day (OFP-SPC) and Joe Scutt Phillips (OFP-SPC) on different elements of the assessment data inputs, including size frequency, conditional age at length and tagging data.

Summaries of the raw size composition data were considered briefly by the group, noting similar issues of low sample numbers for the recent covid period, data patchiness historically, and multi-modal data. The size data are complicated to review, and a dashboard type app. would be useful to better explore the sources of variability in the data over time. A thorough review of size composition data for the assessments was suggested. Given the sheer amount of data going back many decades and across fleets and different sampling programmes this is a large task not feasible during an assessment and requires a standalone project and dedicated resources. Size composition data continues to not receive the attention it deserves due to the time constraints on assessment work, continuing to apply size data that has not thoroughly been reviewed and triaged to improve representativeness is an ongoing concern of the assessment scientists. The complex and variable nature of the size compositions over time for the defined fisheries means it is problematic to have a selectivity function that can fit the quarterly data well across most time periods. The review of the size composition data may help to remove unrepresentative variation in size compositions that are difficult to model.

There was discussion on the issue of the reduction in weight composition data available in the recent decade as the longline bigeye catches become more dominated by non-Japanese fleets that provide composition data as lengths (same concern for yellowfin). The length composition data have more coverage in recent years and will be expected to do so in the future, whereas the weight data is more representative historically when the catches were dominated by Japan. Currently weight compositions are used in the assessment for the longline fisheries, although the previous assessment conducted a sensitivity where length-frequency data were used instead of weights for the longline fisheries in regions

4, 5, and 6 beginning in the mid-2000's (this had very little impact). It is not suitable to use weight and length data at the same time for a fishery, even if it is available, as it can either introduce conflict (if data were in disagreement) or over-weight the model fit (if they were in agreement). It was also noted that sampling processes vary between length and weights, and this may have implications for how the two data types represent the catch size compositions. It was noted that it is important that the selectivity functions are well defined so the fish are extracted by the model at the correct sizes, so whatever approach is chosen, unrepresentative data should be removed for estimating selectivity. This further emphasizes the need to review the size composition data in detail. Variability in size composition can be informative on recruitment and you do not want to lose this information by removing too much of the data without a good reason, so data triaging should be carefully considered.

The tagging data summary

The tagging data summary from Joe Scutt Phillip (OFP-SPC) covered yellowfin and bigeye. For yellowfin it was noted that since the last assessment there has been 4,335 new releases with 162 validated recaptures and 56 of these at liberty for over 6 months. The recent releases have mostly been in the equatorial central Pacific south of Hawaii. A diverse size range has been tagged in the recent releases, from 40 to over 100 cm FL. For bigeye, there have been 10,568 new releases since the last assessment with 859 validated recaptures, including 284 at liberty for over 6 months. The recent releases have also mostly been in the equatorial central Pacific south of Hawaii and also mostly ranged in size from about 40 to 100 cm FL, but with some larger mature fish between 100-150 cm FL. The tag-recapture data for yellowfin and bigeye were summarised in terms of bulk transfer rates between model regions, and this was suggested as potential information to inform transfer rates for the assessment as either fixed parameters or priors. This was suggested as an interesting sensitivity analysis but must consider that the data are not representative of all the age classes. The use of electronic tag information to inform movement was also discussed, which could be the subject of an external analysis given the limited capacity of MULTIFAN-CL to incorporate such data.

Thom Teears (OFP-SPC) presented on the treatment of the tagging data for the 2023 assessments, noting it will be consistent with the 2020 assessment, including using actual days at liberty for mixing period assumptions, but with the implementation of the tagger effects correction model that was applied in the recent skipjack assessment and separate tagger effects models for the Western and Central Pacific cruises to account for different vessel set ups, crews etc.. It was noted that a new feature in MFCL can account for tag shedding internally, so this might be useful to remove the need to do the external analysis.

The change in the tagger effects models was suggested as a good change. Low numbers of tag seeding experiments have been done since 2015, and the topic of reporting rates was discussed in relation to the indication of reduced reporting rates from tag seedings since 2015. It was suggested that this may be accounted for by setting up time varying reporting rates, perhaps by using different recapture groups, but there was some uncertainty on how to do this. It was noted the recent low numbers of seeding experiments mean there is also considerable uncertainty as to how reliable the reduction in recent tag seeding reporting rates actually is.

CPUE analyses

Thom Teears (OFP-SPC) provide a presentation on the progress and further plans for CPUE analyses for yellowfin and bigeye. The analyses so far have attempted to initially replicate the indices from the previous assessment produced in VAST, and then repeat the analysis in sdmTMB. It has been decided to continue with sdmTMB for the CPUE analysis as it much faster than VAST, is easier to use and get support on issues, and would appear to be more replicable for future analysts. There are some minor differences in the indices calculated using the sdmTMB compared to previous analysis in VAST that may relate to data updates to the historical data since the last assessment, but it was also noted that the VAST and sdmTMB meshes look different, so running a more similar mesh configuration may improve the matching. The switch to sdmTMB was supported by the workshop.

The CPUE analysis has since progressed to consider some of the suggestions from the peer review to investigate the sensitivity of the relative abundance estimates among regions to alternative CPUE spatial configurations. This is in response to concerns that the CPUE indices are overestimating abundance in the northern and southern regions compared to the equatorial region. As the model closely fits the regional abundance variation (scales) indicated by the CPUE indices, the stock status is being overly influenced by unreasonably high biomass estimates away from the main equatorial fishing areas, thus buffering/biasing the overall stock status. The analyses have so far explored:

- removing grid cells with very low numbers of samples (<500 observations) under the assumption these do not represent viable habitats and should not be used to estimate regional abundances, this has the effect of reducing the areas across which the analysis aggregates the abundance indices for the northern and southern model regions,
- running separate sdmTMB models for the northern, equatorial and southern regions, to remove the spatial correlation influence of the equatorial region on the northern and southern regions.

The results of these analyses were presented and while the analysis that segregated the northern, equatorial and southern regions did show some differences in the CPUE trends, more notably for the southern region, there was little influence on the relative abundance among regions, even with the removal of the poorly sampled areas.

Further analysis will explore; increased number of knots, bilinear interpolation, adding seasonal hierarchical random effects, single fleet indices to compare with multi-fleets, and use biological envelopes with environment variables added to the models to further constrain the viable habitat area.

The workshop recommended providing information on how many CPUE cells are being imputed within the model (assumed to account for lack of data). It was also noted that the global models are trying to do a lot in terms of estimating covariate effects, so perhaps the global model results could be considered to develop regional scaling, and sub-regional models used to identify the abundance trends. This might help with getting more reliable parameter estimates at smaller spatial scales rather than trying to estimate everything at the scale of the entire WCPO.

The potential inclusion of vessel ID within the model was discussed, with a recommendation to analyse a shorter time series of data where that information was available and compare with the results from the longer time series. In turn, using data from a specific area where the Japanese fleet has good data coverage to compare with the resulting model with other flags in a single versus multi fleet index comparison was noted.

Information was presented to the workshop on the time periods for which data on gear configurations and materials are available for the Japanese fishery. While hooks between floats is available from the 1960s, other gear configuration data is not widely available until the late 1990s - 2000s, main line and branch line materials are available from 1994.

The workshop discussed the issue of effort creep, given the potential impact on the longline abundance indices. Estimates developed in other oceans were noted, while the potential to include alternative effort creep levels within the uncertainty grid was raised. Following the advice of the yellowfin peer review, it would be a good time to propose a project to assess available information to better inform WCPO-specific discussions on potential effort creep levels and temporal dynamics of these influences. The potential to collaborate on this issue with other tRFMOs and especially the Distant Water Fishing Nations was noted. SPC could consider providing a project TOR to be considered by SC19.

Biology

Jemery Day (OFP-SPC) provided background on biological aspects for the bigeye tuna assessment. For conditional age at length data, it was noted that there are less age/length samples for bigeye than for yellowfin, with most collected between 2013-2018 and all from below about 10°N. The comparison of growth for sub-areas with the WCPO, and between males and females, for the ages that comprise the bulk of the population and catches showed no notable differences (but acknowledging that the north region has no samples to compare). While the previous assessment experienced issues with estimating growth internally, the peer review recommends this approach with either VB or Richards. Another attempt will be made on internal growth estimation. Preliminary growth curves were shown. It was noted that the approach to natural mortality (M) will be consistent with yellowfin only that the base M for bigeye was lower (i.e., bigeye M was about half yellowfin in previous models, but they seem to live to similar ages, which appears unusual). M for bigeye will be reviewed and attempt to estimate internally. The Lorenzen M at age formulation will also be considered. Reproductive potential will be formulated as per the previous assessment as no new information is available.

Model development - yellowfin

Arni Magnusson provided a summary of the initial focus areas for the development of the 2023 yellowfin models, including implementing the catch condition method and further refinements to this model to achieve a positive definite Hessian solution. Arni showed the Shiny App. that can be used to readily explore and compare model outcomes and diagnostics. He provided an overview of the concerns regarding regional scaling, and the reasoning behind the suggestion that biomass being attributed the northern and southern regions is disproportionately high compared to the equatorial region. He also spoke about concerns regarding recruitment estimates for each model region, noting how over successive assessment recruitment distributions among regions have changed, with different regions having no recruitment

attributed, including regions that would be expected to have significant recruitments. This indicates that the high complexity of the model makes the statistical estimation unstable, and there is insufficient information in the data to inform spatial recruitment patterns. Simplification of the model spatial stratification was discussed as an important avenue to explore to help resolve these instabilities, noting the simpler 4 and 5 region structures. He listed a number of model runs he was planning to explore.

DAY 3 – MFCL, MSE/harvest strategies, 2022 skipjack assessment follow-up work, TARP – tuna assessment research plan

Model development – bigeye

The model development discussion continued on day 3, with Jemery Day discussing the bigeye model developments. He provided an initial summary of the yellowfin peer review recommendations/outcomes, highlighting aspects that he expected would be a focus of his work on the bigeye assessment. He noted that upon implementing a catch conditioned model and not estimating tag reporting rates (RRs) for release groups with zero or only one recapture (this removed tag RRs on bounds) it was possible to obtain a positive definite Hessian solution. Several model development steps have been conducted and preliminary results where shown, including; updating the MFCL executable, implementing catch conditioning, not estimating the reporting rates for zero/one recaptures, estimating growth internally with VB and Richards. Ideas for the next model explorations were described.

The workshop commented that when catch conditioning was implemented for skipjack for the first time last year, to run effort projections an additional step was required to conduct a fishing mortality – effort regression, this resulted in a change to the management quantities, which was concerning. It was responded that this is being looked into.

The issue of recruitment distribution was raised to ask if the unusual recruitment estimates observed for yellowfin are a problem for bigeye. It was noted that the model has flexibility to trade off recruitment with movement as a there is not a lot of tagging data to inform movement, something that will need to be considered closely.

The workshop noted the CAPAM best practice recommendation to analyse the tagging data outside the model to provide information on mixing rates as external inputs. This was discussed further in relation to perceived inadequacy of simulated mixing periods using the approaches such as SEAPODYM/Ikamoana, and that the actual tag data should be analysed. It was responded that SEAPODYM is not a simulation as such but is an estimation fitted to real data including tagging, and that the opportunistic and limited tag data coverage for yellowfin and bigeye makes the external mixing rate analysis problematic, some form of integrated modelling approach is probably preferred.

The workshop noted it will revisit discussions of model development, uncertainties, spatial stratification options and diagnostics on Friday.

Multifan -CL

Nick Davies (Takina - MFCL development and support consultant to SPC) provided a summary of the recent developments and workplan for MFCL including exploration of additional features suggested by the peer review.

The most notable development work included implementing and fully testing the stochastic projections for the catch conditioned (CCond) model, revision to the orthogonal polynomial recruitment penalty function, and attempting to complete/implement a partially developed length-based selectivity feature. The stochastic projections and the orthogonal polynomial recruitment penalty function were achieved successfully. The length-based selectivity feature that was partially developed in previous versions but never fully tested or implemented in any assessments, required significantly more effort. An in-depth investigation found that there were three anomalies with the feature that would require significant additional work, and these would require refining of numerical procedures as well as code. Given the time required to work through this (and that Dave Fournier has retired), it was decided it was not feasible for the current assessments, and focus has shifted to other peer review requests that were considered feasible for the current assessments. A range of other corrections and enhancements have been made and included in release version 2.1.0.0 (since updated to version 2.0.8.2): CCond model initial equilibrium population calculation for fishing impact analysis, CCond CPUE concentrated likelihood formulation, added depletion ratio and log-MSY variables to the standard deviation report, initial equilibrium population calculation based on M has the multiplier disabled by default for the fishing impact analysis, enabled recruitment compensation option for the CCond model fishing impact analysis, initial equilibrium population calculation for the fishing impact analysis during standard deviation calculations, and removed iterative incrementing of maturity-at-length vector at model re-start.

Aside from this development work, Nick has spent much time and effort to train and assists new staff in MFCL, supporting the work on the 2022 skipjack assessment, supporting the peer review workshop and the development of initial assessment models to support the 2023 assessments. The immediate workplan for 2023 was noted and will be summarised in an IP to the SC19.

Next generation tuna assessment model

Following the MFCL update Nick Davies provided a perspective on where to next for MFCL and software developments for tuna assessments. The recent retirement of the lead developer and innovator of MFCL, Dave Fournier, means that any significant ongoing development of MFCL will be curtailed, and the sole person now taking care of MFCL has limited time for development work with their other stock assessment/MFCL training support duties. While MFCL remains fit for purpose, Nick was keen that we look forward and plan for the future needs of tuna assessments. Nick provided his perspectives on the direction he thinks tuna assessment software and approaches should be moving towards. He provided some historical perspective, including that MFCL has been applied and developed for around 30 years, but without Dave it is clear that future innovation will be limited, and consideration of alternative options is important. These may include using other software that has dedicated development teams or developing a new software package and encourage a broader user and developer base. This situation is

not unique to MFCL and was addressed at CAPAM 2019, that focussed on the topic of the 'creation of the next generation of general stock assessment models'.

Nick suggested that when considering why a model package should be replaced a key reason is usually the underlying model structure. Most models, including MFCL and Stock Synthesis that are applied to tuna assessments, are age structure – that is growth in the population numbers is in respect of age. But most of the data collected to inform model estimations is in terms or size and many of the processes are length based, this causes difficulties for age-structured models because growth can be expected to vary spatially and temporally for tunas in response to environmental variation and other factors. It was noted that scenarios where growth rates vary require a length-structured model.

It was noted software such as GADGET and CASAL2 have length-structured dynamics and should be explored further along with L-SCALA developed for skipjack in the eastern Pacific. GADGET3 is being moved to TMB that might improve the computation efficiency, but neither GADGET or CASAL2 include random effects estimation yet. It was noted that many of the pieces required to build a new age-length structured software package exit in other existing packages including MFCL. NOAA's next generation 'Fisheries Integrated Modelling System' (FIMs), project, was mentioned noting it will be based on TMB but that there was no clear indication it would incorporate length-based or combined age-length structure. The initial work is focussed on a standard age-structured model. Features that are available from MFCL development already include: a simple age-length-structured model tested and published, a length-structured growth model that is fully differentiable and tested, and a range of other features such as spatial, movement, tagged population, etc. It was noted that it's clear that TMB is the code base necessary for any new software.

Further points were discussed on why age-length structure models are advantageous, including conserving historical length-specific process within cohorts, allowing for both age- and length-specific effects in growth processes, and will address the key requirement of modelling spatially-varying growth and fish movement. An age-length structure can directly provide age and length predictions specific to observations, such as: conditional age-length data, and tagging in respect of lengths at release and recapture, and recapture length-increment observations. An additional reason for tracking both age- and size-dynamics is that for some stocks tagging data are the key source of information on movement, growth and fishing mortality, but it is necessary to allocate tags (for which length-at-tagging but not age-at-tagging are usually known) to age, often based on a deterministic growth curve. In contrast, the assignment of age to tagged animals is straightforward if the model is age and length-structured.

An example of an age-length structured model (CALEN) applied to the Hauraki Gulf-Bay of Plenty snapper population was presented explaining the age-length structured population state matrix versus the standard age-structure population state vector. It was also noted that the key requirement for an age-length structured model is a length-based growth function. An example was provided from the CALEN model, but for TMB a length-based growth function would need to be fully differentiable. Another example of a length-based fully differentiable growth function created from an MFCL sub-project was presented. This was tested as highly successful and is a ready-made length-based fully differentiable growth function.

Nick presented as possible strategy for developing a new software in TMB that involved coding a simple age-length-structured model in TMB, import the MFCL developed length-based growth function, simulation test this using a simple length-structured model with assumed growth rate (e.g., L-SCALA) and explore further developments for: environmental or annual covariates; age/cohort dependent effects. Then importing the desirable other features from MFCL into the TMB model; tagged populations, spatial processes (movement), initial equilibrium population conditions, selectivities (splines), size composition, tagging, etc. likelihoods, fishing impact analysis, Hessian diagnostics.

The workshop was asked to consider the ideas of:

1. Conduct a simulation scoping study to evaluate the degree of bias incurred in management advice due to length-based processes in time and space not being accurately modelled by age-structured models.
2. Evaluate utility of existing length-based models (GADGET, CASAL2, spatial L-SCALA).
3. Collaborative opportunities to work with other agencies on developing new length-age structured TMB models.

The workshop noted that there were strong benefits to ensuring any software was developed collaboratively and used by a wider user base, including other tRFMOs, to help ensure that development and support can continue when individuals move on, and that the software has longevity. It was also discussed that even age structured models, notably MFCL are complex enough and difficult to learn, this creates problems when there is high staff turnover in organisation like SPC, training takes a lot of a new staff's time, and there also needs to be available ongoing people to support/provide training. Any new software should be developed with the end user in my mind and should not be so difficult to use that people ultimately avoid it in favour of simpler options.

A steering committee of scientists and developers was considered worthwhile, and it was noted that development of an age-length structured model would be complex and would need specialised software developers rather than scientists to build it to ensure it was coded in an efficient and refined way. The work required to build a new software was noted as being immense and would likely be a 5+ year process.

Existing potential frameworks and their continued development were noted as important to consider including CASAL2. Developments undertaken by Dragonfly were also noted, in the context of the complexity of the age-length structured model approach and difficulties of interpretation and diagnostics, emphasising the need for specialist developers and ensuring the appropriate support tools are developed.

The workshop supported the staged approach of a scoping study that could include developing a Pacific-wide simulation model/operating model to allow the impact of capturing the additional structural model complexity on management parameters to be tested, before embarking on any serious new software development. A small informal model development working group at SC was also supported as a good idea to keep this conversation going and further develop a proposal for a scoping study.

Management Strategy Evaluation

Rob Scott (OFP-SCP) provided a presentation to update the workshop on progress in the development of MSE modelling and technical support work for the WCPFC harvest strategy workplan. Noting that an interim Management Procedure (MP) for skipjack was adopted at WCPFC19, he provided an overview on the skipjack MP and the process of adopting the HCR within the MP. The operational schedule of the MP was described, and pointing out that this year would be the first year to run the MP and provide a recommendation on the catch or effort levels. The MP would be run in the year immediately after the stock assessment, with the stock assessment being the central part of the monitoring strategy. Planned papers for SC19 were outlined.

Nan Yao (OFP-SPC) provided a presentation on an exploration of alternative CPUE indices that could be used to capture additional uncertainty within the MSE operating model grid. The exploration was a response to a request from the SMD in 2022. The study explored several plausible alternatives, but they did not result in appreciable changes to the standardised CPUE and/or model estimations were insensitive to the alternatives. It was therefore suggested from this analysis that the inclusion of the alternative CPUE formulations into the OM grid was not justified. Further work might occur on this aspect.

Rob Scott then presented on the development of the estimation model for the albacore MSE and MP. A key issue that has become problematic for the MSE work and south Pacific albacore projections for other management evaluations is the estimated recent dip in stock status that is due to estimated recent exceptionally low recruitment. The drop in recruitment carries through into the early projection period causing a rapid dip in the projected stock status until more typical simulated recruitments start to take effect in the projection period. The predicted drop in stock status is not clearly evident in recent industry perceptions or nominal CPUE, so a thorough exploration to understand the cause of 'the big dip' is underway. If this dip is applied in the context of MSE it will require a rapid reduction in catches or effort which may not be feasible to implement. Retrospective patterns were also noted and how these differ between movement hypothesis (model estimated v SEAPODYM) and further contribute to complication of interpreting the cause of the recent dip in recruitment. With regards the candidate MP, several estimation methods have been tried based on surplus production methods (JABBA and SPICT) that provide similar estimations but do not predict the dip at the end of the time series. An age structured production model (ASPM) has also been tried, implemented in MFCL, and this approach was fairly consistent with MFCL in terms of the variation and trends in stock status but tended to overestimate the stock status. Overall, the ASPM shows promise, but needs more work. Data simulation, especially CPUE is noted as very important for the MSE, currently the simulated CPUE is producing the high frequency variations but is perhaps too consistently centred on the expectations from MFCL operating models, and perhaps too well behaved to be realistically simulating uncertainty in projected CPUE, so still some work to do on this aspect.

The workshop suggested that as the big dip occurred in the terminal years for all the retrospective models there may be an issue with the model setting rather than the data. It was also noted that the terminal dip in depletion was not present in the 2018 assessment (but a drop in recruitment was estimated in the terminal years), and the last year of that assessment was 2016, but the retrospective model from the 2021 assessment that ended in 2016 did show the depletion dip, so it would be worth to compare those models.

It was noted that projections from the 2018 assessment model also predicted a dip in stock status (assumedly due to that model estimating a recruitment drop in the terminal tears), so the 2021 assessments prediction of a recent dip in depletion is not inconsistent with the short-term projections of the 2018 assessment before the more typical historical recruitments take effect. Further comments suggested that the issue may be related to model settings, structural changes or other conflicts, as it doesn't seem to be driven by recent data updates.

The workshop noted the implications of this behaviour on candidate management procedures and their near-term outputs. It was important to understand whether the cause is an artifact and could be misleading or is a real feature of the data that the model is appropriately fitting to. The next stock assessment was noted to be in 2024 when the albacore MP is scheduled to be developed, so if these issues remain unresolved it will be quite problematic.

Further comments were made on the likelihood of occurrence of the recruitment drop in the context or historic estimates. It was noted that the albacore indicators paper to SC19 will provide indications of the most recent CPUE trends that can be compared to the model predictions, but will need to be wary of targeting shifts when interpreting recent CPUE. It was also mentioned that recent US fleet CPUE was the best in a long time. It was noted that in the stepwise for the 2021 the dip came in when the new regional structure came in for the WCPO (before the EPO or any new date were introduced), so worth exploring what was happening in more detail at that step.

Detailed investigation of the cause of the dip in recruitment and stock status is a current focus and hopefully this can be resolved soon, and any implications of this issue for the MSE related work understood and communicated.

It is noted that for management projections of fixed catch/effort strategies for TRP evaluations the stock status stabilises at equilibrium levels under long-term average recruitment by the end of the projection period and it is these result that get used to compare the outcomes of the fixed management strategies.

Finlay Scott provided an overview and update of the work on the mixed fishery MSE approach. The presentation discussed the rational for the approach relating to the proportions of the catches of the four tuna stocks being taken by the fishery groups: tropical longline, southern longline, pole and line, purse seine and other. The mixed fishery framework has single stock MPs for skipjack, bigeye and south Pacific albacore, but no MP is required for yellowfin which is indirectly managed through the other MPs. Progress in the development and testing of the modelling framework was outlined, noting that operating models for yellowfin and bigeye are dependent on the 2023 assessments, so completing the full set of OMs for the mixed fishery framework to enable full testing will not be possible until the new assessments are accepted by SC19.

The workshop noted that the framework is hierarchical with the skipjack MP having control of a large proportion of the bigeye and yellowfin catches. This could be a problem for bigeye for example, as declines in bigeye would mean adjusting the longline catches but there may be no adjustment to bigeye catches by the purse seine sector if the skipjack stock is not also declining. This was noted as a risk that can't be fully appreciated until scenarios are tested with the complete framework. It was also noted that if this

was to occur there are other options to adjust bigeye catch by purse seine (which really only catches bigeye in FAD associated sets). The skipjack MP manages total purse effort, not the proportion of effort allocated to free school and associated sets, so this proportion could be adjusted to limit purse seine bigeye catches using levers such as the FAD closure.

Follow-up work on the 2022 skipjack assessment

Claudio Castillo-Jordan (OFP-SPC) provided a presentation on ongoing work in response to recommendations from the SC18 of the 2022 skipjack assessment. He described some history of the skipjack assessment noting the changes from assessment to assessment in the people doing the assessment and the spatial configurations (8 regions 2000-2008, 3 regions 2010 and 2011, 5 regions 2014 and 2016, 8 regions 2019 and 2022). He noted the key uncertainties in the assessment, specifically discussing growth and tag mixing and how these were dealt with using new approaches in the 2022 assessment. He noted the other uncertainties subject to sensitivities, tagger effects, effort creep and movement and the results of the assessment. He then focussed on follow-up work to address some of the issues raised by the SC18 assessment, in particular:

- Indices of abundance: develop better informed effort creep scenarios – pole-and-line focus, other effort metrics (i.e., further explore 'travel distance between fishing events').
- Detailed exploration of evidence and mechanisms for the increased recruitment predictions/trend of the assessment.
- Further model simplification to remove negative eigenvalues (achieve a positive definite Hessian), inclusion of new diagnostics.
- Exploring software alternatives (i.e., SS3 for skipjack).

Paul Hamer (OFP-SPC) discussed effort creep and recruitment trends, noting for the skipjack assessment, the issue of effort creep is most important in relation to the Japanese pole and line fishery CPUE abundance indices. The recent work by Matsubara et al., (2022) was discussed to describe the operational/technological changes that have occurred in Japan's pole and line fleet and the timing of the changes/uptake, and associated trends in standardised CPUE. It was noted that the CPUE metric of catch/day may not be very sensitive to changes in abundance or efficiency. The issue of effort creep requires more dedicated work, and this will be supported by WCPFC *Project 115: Exploring Evidence and Mechanisms for a Long-term increasing Trend in Recruitment of Skipjack Tuna in the Equatorial Pacific and the Development and Modelling of Defensible Effort Creep Scenarios*. This project has a small budget to facilitate collaboration between Japan FRA and SPC on the pole and line effort creep work. It was also noted that effort creep is perhaps not the best terminology as nominal effort is not increasing, the issue is really about how improved operational effectiveness is influencing catch rates over time, perhaps technology or effectiveness creep is a better term. The conceptual approach to exploring this issue in Project 115 was discussed, starting with a phase of more qualitative and empirical investigation, including consultation with/survey of industry to better document how operational efficiency has changed overtime and the perceptions for how the changes have affected catch rates. The idea is to build a narrative and evidence base for the influence of operation changes, including technology, on the catch rates, with identification of the key factors to analysis with statistical methods that would occur in phase

2. If some plausible scenarios of creep can be identified these would then be tested by MFCL models for their impact on the stock assessment predictions.

The workshop highlighted that a consideration of the recording of catch and effort and its impact on the abundance metric was warranted. Effort in days may not reflect true effort, while the catch level may be affected by the number of schools visited in a day and the proportion of the school taken in a fishing event. Yoshinori Aoki (FRA-Japan) noted that the number of schools fished were recorded during tagging research cruises, which might provide a source of additional information. Information is also gained during the tagging cruise through talking with the industry; recent conversations highlighted the uptake of new sonar equipment that improved performance and range. These factors would be part of the proposed industry questionnaire to be used as part of Project 115. The potential to talk with long-serving skippers to gain an understanding of operational changes was highlighted as being very useful. The need to ensure the conversations focussed specifically on skipjack was noted, given vessels could fish on both that stock and North Pacific albacore. Others noted that the Project 115 was encouraging, but irrespective, some approach to considering effort creep uncertainty in assessment is important. SPC noted that incorporating the results of the study by Matsubara et al. (1% per year effort adjustment) within the skipjack assessment as a sensitivity had little impact on management quantities. The workshop noted that there might be confounding effects with the tagging programme, that might constrain the scaling of population size at either end of the time series and negate the influence of the effort adjustments, which could be further investigated, although this was discounted by other comments. Finally, discussions focussed around whether effort creep was a long-term trend, or occurred in step changes, as suggested by the available information from the Japanese fishery. Conversations at SC5 in Vanuatu, which noted rapid uptake of technology within the fleet (with a subsequent short period of 'learning'), supported the step/regime change scenario. Noting it is a difficult problem to solve, the likelihood that fishing power was increasing, and the limited explanatory factors that were available for modelling, suggests any empirical value may be underestimated.

As to the investigation or recruitment trends it was noted that this aspect of Project 115 had not been considered in much depth as yet, and it was unlikely that this would not be a focus until later in 2023, with the initial focus being on the effort creep aspect and the phase 1 surveys of the Japanese pole and line industry. It was noted that some models have recently been run to explore the sensitivity of estimated recruitments and management quantities to the CPUE indices. Excluding either the recent purse seine indices or the longer-term pole and line indices had little impact of the increasing recruitment trends. Removal of the recent purse seine indices for the equatorial region did result in a more stable recruitment trend in recent years, higher spawning biomass and more positive stock status (depletion) compared to the declining trend when it was applied. Interestingly, removal of the pole and line index had very little impact on recruitment, spawning biomass or depletion, so it seems the model is more sensitive to the abundance indices applied to the equatorial regions.

In relation to model convergence and Hessian status, progress in the area was noted, in that modifications to the 2022 diagnostic model had achieved a positive definite Hessian (PDH) solution. The model with the PDH estimated spawning biomass and depletion to be very similar to the original diagnostic model.

Claudio provided a summary of his work exploring the application of Stock Synthesis 3 (SS3). He noted difficulties he was having with applying SS3 to the skipjack assessment data, especially the tagging data, which is a key data source for the skipjack MFCL assessment. SS3 requires tagging data to be input by age as opposed to length as for MFCL. More work is required to get a better understanding of whether and how to best apply SS3 to the skipjack data.

The workshop raised the potential for a review of the skipjack stock assessment to be discussed at SC19. Noting workloads and the fact that outcomes of the recent yellowfin review had implications across all the key tuna stock assessments another peer review would really stretch the assessment team, give the current workload. This is of course a decision for SC, but SPC noted there were advantages in allowing those improvements to be embedded within the next round of regional assessments before another review was undertaken.

Tuna Assessment Research Plan

Graham Pilling (OFP-SPC) provided an overview of the Tuna Assessment Research Plan (TARP). The TARP has been put forward as an IP (information paper) each SC since SC16 in 2020. The TARP was created to provide a document to house important recommendations for future work noted from stock assessments and other discussions (i.e., PAW) as well as other strategic research and data improvements needs. It is intended to be updated annually to note research work that has been done or instigated and add new needs. This SC the paper will be discussed as a working paper in the plenary for the first time, so it was good to get some perspectives and endorsement from the PAW.

The aims of the TARP were noted:

- Capture key R&D recommendations arising from assessments.
- Allow SC prioritization before subsequent assessments.
- Capture relevant research by all WCPFC members.
- Allow SC and the Commission to prioritise the research budget.
- Identify funding gaps to be filled by proposals external to WCPFC.
- Enables SC to review activities and progress over time.
- Track how ongoing SC projects support improved advice.
- Foster collaboration

The current focus on activities and projects relevant to key tuna assessments was discussed, along with the approach to keeping the TARP relevant and the current structure of the TARP, which is a lot more focussed on documenting the research needs than other plans such as the Shark Research Plan that has a much broader scope in terms of also summarising data and other information, recommending assessment schedules etc.. Example sections of the main TARP summary table were presented.

Finally, it was put to the workshop to discuss:

- There is a lot of information is already captured (and to be updated).
- Is this the best way to identify research in other groups (cf SRP)?
- Is an SC19 presentation needed? And/or potential to discuss content in a Small Working Group?

- Consider how to implement TARP within the SC framework.
- Provide any inputs on updated activities and ideas on process prior to SC paper deadline.
- Are table format adequate?
- Any fields or information to add?
- Anything missing from the paper?

It was also noted that SPC will provide a background paper to inform SC discussions on options to address time challenges in the review of WCPFC stock assessment inputs prior the delivery at SC.

The workshop agreed on presenting the TARP to SC, and that this was an opportunity for the PAW to have input. A SWG at SC19 was supported as worthwhile, and that group could also include consideration of the future assessment software platforms for tuna assessment.

It was noted that it has typically been that SC provides the research/project priorities and the Commission just provides the fund available to cover the priorities, and depending on available funds lower priority projects may not get funded. Graham indicated that he would make it clearer that the research priorities are recommended by the SC, but also the TARP is not just about short-term research projects for upcoming assessments but also to allow focus on allocating research budget in the longer term and to ongoing research needs, thereby helping the Commission members to better forward plan their budget allocations. It was also suggested by the PAW that given the amount of work listed already, additional science budget should be requested to support the work on what is a highly valuable fishery.

It was also noted that as research needs can be picked up by other RFMOs or organisations, the leads and collaborations can be diverse. Further, some projects emerge at SC, so a challenge is to consider resourcing the immediate and longer-term needs during the short time available at SC meetings. Voting on projects at SC has been a bit rushed and not well considered, so having a SWG at SC to discuss the research needs and make more informed decisions on priorities would be useful.

DAY 4 – SEAPODYM, Silky shark assessment, Project 113 – review of assessment uncertainty characterisation, discussion of assessment models/spatial structure/uncertainties and diagnostics

SEAPODYM

Inna Senina (OFP-SPC) provided a detailed presentation on SEAPODYM (Spatial Ecosystem and Population Dynamics Model). She began with an historical perspective on the development of SEAPODYM, from the initial work showing how skipjack distribution is influenced by the ENSO cycle, through to the current day where SEAPODYM is now open-source code. She explained:

- the modelled processes of reproduction and survival that drive the drive movement to predict population density by time, age and space.

- the environmental forcing and model dimensionality, including how the ocean vertical profile is condensed to three layers and the space dimension is simplified to two dimensions. So, there are dimensions: age (larvae, juveniles, young and adult) x time x 2d-space.
- underlying continuous equations
- functions for spawning habitat and reproduction
- feeding habitat and environment restrictions on accessibility with age
- natural and fishing mortality
- directional movement (advection) in the Eulerian model approach
- non-directional movement (diffusion) in the Eulerian model approach

An example of the approach was provide based on a data set from satellite tagged turtles. Then it was explained how the quantitative SEAPODYM model is developed, noting that dynamic processes can be explained by up to 30 parameters:

- 2 reproduction/Beverton-Holt function,
- 5 parameters of natural mortality function,
- 5 parameters in spawning habitat,
- 12 parameters in feeding habitat,
- 4 parameters to define movement rates
- 2 for seasonal spawning migrations

Fisheries parameters (catchability and selectivity) for each fishery are estimated given the species spatial distributions at age and time.

It was emphasised that the model parameters are estimated by fitting to observed data.

The methods for generating model predictions for population density at age, time and space, and the catch and length frequency of the catches were explained.

Preparation of data inputs to inform model parameters were discussed: fishery catch and effort data, size composition, tagging data. The likelihoods for catch/CPUE, size composition and tagging data were then discussed, followed by how the different data inform the model dynamics.

An example was provided to demonstrate how the inclusion of tagging data improved the skipjack movement modelling.

The model outputs were shown and explained including the ability to present predictions for both fished and unfished situations. The model conditioned on data for south Pacific albacore was also shown to successfully predict albacore distributions for the Atlantic stocks.

Some conclusions and outline of the future workplan were provided.

The workshop was very appreciative of the in-depth presentation on SEAPODYM.

The workshop asked about the differentiation between abundance and catchability, and it was noted that because of the environmental variation, like mixed layer depth, being accounted for in prediction of density there is no confusion between catchability and density, catchability for a fishery can be modelled as constant. The differences observed between predictions by SEAPODYM and CPUE observations were discussed noting that SEAPODYM will make prediction of fish abundance due to the environmental forcing even where there are no catches. In the region between Hawaii and Japan for example there are some high catches of juvenile bigeye south of Hawaii by the fishery targeting swordfish, but because the conditions in that region are similar to those off Japan in Kuroshio Current extension, SEAPODYM will predict similar densities across that whole region.

The workshop asked how the young are introduced into the model and it was explained that recruitment of larvae and juveniles is modelled by the combination of the spawning habitat function and the Beverton and Holt stock recruitment function. It was noted however that observational data on the early life history stages is lacking. This was followed with a question on how static oceanography variables/climatology's are treated in the continuous model. The model uses coupled model outputs, ocean circulation and biochemical models to provide the dynamic fields (e.g., O_2).

The provision of additional plots, data sets and diagnostics, including the likelihood figures for diffusion and advection rates, was welcomed. However, it was suggested that these are difficult to interpret and that it would be good to provide likelihood profiles for other parameter or quantities that are more familiar for population dynamics models. There was further discussion on the patterns of juvenile distributions predicted by SEAPODYM, particularly for the regions 4 and 6 extending into the EPO. It was noted the patterns of longline CPUE do provide a signal to the model for these patterns, but also there can be differences due to that fact that the model predicted habitat suitability can still indicate higher densities. It was clarified that the CPUE used is nominal as you don't want to remove the environmental effects from the CPUE as often occurs in the standardisation process. It would be useful to explore the sensitivity of predictions to removal/inclusion of different data sources to better understand which data are influencing the predictions.

It was also noted that the differences in the predictions for juvenile dynamics (i.e., recruitment) in equatorial region 4 between SEAPODYM and MFCL relate to how SEAPODYM uses environmental forcing of the ENSO to move fish throughout the equatorial regions (i.e. MFCL model regions 3 and 4), but MFCL can't respond to the environmental influences and instead predicts recruitment variation to explain the variation in the fishery data for region 4.

The workshop asked about the northern and southern hemisphere populations of albacore and whether there is an environmental barrier that keeps the populations separated. It was replied that albacore have strict dissolved oxygen (DO) limitations, the low DO in the equatorial region (upper mesopelagic layer) provides a barrier to mixing between the northern and southern hemisphere stocks. But no similar boundary was present in the Atlantic where they do mix north and south.

The workshop asked about plans for further characterising uncertainty in SEAPODYM predictions, noting the different types of uncertainty related to the environmental forcing products used in the model,

estimation uncertainty of parameters and process uncertainty. It was responded that currently uncertainty is only characterised from the variance/covariance matrix derived from the Hessian, so it only considers the uncertainty around the converged parameter estimates. The future plan is to start to incorporate sensitivities to the forcing variables/products, and the fishery data and develop an ensemble type approached to characterising the prediction uncertainty.

There was interest by the workshop in the approach to ongoing development of SEAPODYM noting it was now open-source software. It was responded that the development team at SPC is very keen on collaboration and involvement of others, it is hoped by showing how the quantitative aspects of SEAPODYM are being enhanced, more people will be attracted to use it and collaborate.

The workshop noted that SEAPDODYM was a highly sophisticated and useful estimation model fitted to data, and like all models it we be better/more reliable at some things than others. It would be good to have more information to understand the strengths and weaknesses, emphasising again the need for the diagnostics, sensitivities to data inputs and provision of more spatial and temporal diagnostics to better understand where misfits are most pronounced. It was noted that there is a toolkit in R to diagnose and validate SEAPODYM models and inform model selection. Another presentation could be provided to explain the model validation and selection process as it was not the focus of the presentation at this workshop.

Silky shark assessment

Kath Large (Dragonfly), Steve Brouwer (Saggitus), and Phil Neubauer (Dragonfly), provided presentations on the background to the silky shark assessment, characterisation of fisheries data and planning for the assessment work. They also introduced a new member to the Dragonfly team, Kyuhan Kim, who will be working on the silky shark assessment.

The background presentation from Kath Large noted the importance of the increased time available for the assessment, with delivery of the assessment to SC20 in 2024. More time will be available to work on the challenging data inputs, CPUE and catch reconstruction aspects. She provided a summary of the previous silky shark assessments:

- Rice and Harley 2012 (updated in 2013, WCPO, Stock Synthesis) suggested that the silky shark stock was likely overfished, and that overfishing was likely occurring against MSY based reference levels, although there was high uncertainty of the model predictions across the uncertainty grid.
- Clarke et al. (2018), used Stock Synthesis, initially attempting a Pacific wide model but could not fit western and eastern Pacific CPUE indices so reverted back to a WCPO only model. The assessment results had high uncertainty and were considered 'indicative' that the fishing mortality was above $F_{MSY.}$

This will be the third attempt a WCPO silky shark assessment. Kath discussed the differences between the two previous assessments, particularly in relation to uncertainty characterisation, catch reconstruction and CPUE indices. She discussed information on movement and possible stock structure and summarised information on the biology.

Steve Brouwer provided a detailed summary of the fishery dependent and observer data available to support the assessment. He noted fishery data shows that silky shark are widespread through the Pacific, however, since 2015 most fleets are releasing most silky sharks (cut free), and a high proportion of releases are observed as alive and healthy at release. This has impacted some of the length data where large fish are not measured if cut-free. However, some length data are available but not for all fleets. Silky sharks are landed in both shallow and deep sets but are most frequently caught in the shallow hooks and comprise a higher proportion of the catch in shallow sets. CPUE data is available from both observer and logbook reporting, although coverage is poor prior to the late 2000s. Gear attributes (e.g., HBF) are more likely to be informative than specified shark targeting information as targeting is rarely/poorly reported. It will be important to consider the introduction of policies regarding shark retention as these do complicate the analysis and interpretation of CPUE standardisations and the use of CPUE for catch reconstructions. In recent years most flags have required sharks to be released.

Phil Neubauer provided a presentation on the teams workplan for the assessment noting the two-stage approach.

The first stage would include:

- Data characterisation (Steve Brouwer - under way)
- Biological data - review and compilation (under way)
- Catch reconstruction
- Length compositions
- CPUE
- Suggesting assessment options to report to SC19 in 2023.

The second stage, if endorsed by SC19, would continue the assessment into 2024 to be delivered to SC20:

- Revisiting inputs based on SC19 feedback.
- Complete the assessment(s) based on SC19 and PAW 2024 agreed inputs.

Phil provided some insight into the plans for catch reconstruction, applying a similar approach to that developed for the recent mako shark assessment, and length composition data (which comes from observers where representativeness depends on the observer coverage) for which a new approach of length frequency standardisation to account for the often sparse and possibly biased (more so in recent years?) length data will be tried. Dragonfly have tried this method on other assessments and are hopeful that it can help remove some of the non-representativeness in the length composition data. He noted the importance of CPUE but emphasised the challenges of CPUE analysis for sharks, including that logbook reporting likely underestimates shark captures, a lack of species-specific reporting until the 2000s, recent trends of cutting sharks free and associated decreases in reporting. Not sure of the solutions to the issue of sharks being cut free. They may explore alternative approaches to CPUE analysis compared to previous assessments.

The USA perspective was provided noting that sharks have been cut free for a long time now, even before the non-retention policy, but that their statistician already does the bycatch estimates so this could save

effort with the observer data and catch time series can be provided for the deep set, shallow set, and American Samoa fisheries, the release mortalities can be applied to these. Prior to the finning ban in 2000 there are more length data, but since then there will be low data. As to length compositions when they deck the sharks they take a fairly accurate fork length but now when they cut them off they just guestimate the length (plus or minus a foot) which is fairly useless. CPUE is mostly derived north of Hawaii where few silky sharks are taken, so not much use, American Samoa might have more information but not many vessels and low observers (20%). Dragonfly would be keen to get the data mentioned to compare to their broader analysis.

The workshop asked about the approach to account for post-release mortality in relation to condition at release. It was replied that the post-release mortality rates from the tag studies were applied to numbers released in that were classified as in good condition, but releases in less than good condition were assumed to die. A follow-up question asked how the size composition will be dealt with for the interactions when some fish are released live while others die. The analysts weren't sure yet of the implications of this or what they might need to do to account for it.

It was noted that the IATTC port sampling programme has measured a lot of silky sharks being landed. They are now exploring CKMR and wonder if there are collaboration opportunities on this approach with WCPFC. SPC responded that they can see that this might be the necessary approach if we really want to get a fix on the population size given the data limitations, but up to the SC and WCPFC Commission to make decisions around the value of pursuing this and allocate funding for the work. There would need to be a scoping study first to understand the sampling requirements; but it was worth putting it on the table. The issue of the stock recruitment relationship (SRR) for shark assessments was raised and that this is a problematic area that needs some more research. It was noted in the previous assessment the survivor-based SRR was used following the north Pacific blue shark assessment, but this led to interesting results – particularly for the dynamic reference points. It was agreed that more work is required for low fecund species on the appropriate approach for the stock recruitment relationship.

The workshop noted that for CKMR it is critical that you have good information on fecundity at age/length, and maturity and age/length for the kin probabilities, while the SRR is important, this life history information is more of a priority.

WCPFC Project 109 is designed to train observers in biological sampling techniques for elasmobranchs but delays due to COVID mean that it would likely not be until 2026/27 that samples might start to be collected in numbers. It was worth adding into the shark research plan a consideration of CKMR scoping and genetic sampling.

Project 113 – review of assessment uncertainty characterisation

Phil Neubauer provided a presentation on the plan for WCPFC project 113: *Further development of ensemble model approaches for presenting stock assessment uncertainty*. Dragonfly has been contracted to conduct this project by the WCPFC.

The objectives and expected outcomes as specified in the project terms of reference were outlined.

General review:

- Review and summarize the various approaches used for characterising uncertainty in WCPFC stock assessments for tuna, billfish and sharks over the last 5 years.
- Describe how uncertainty was communicated in the context of management risks and its influence on decision-making processes used by the WCPFC.
- Comment on the suitability of the recent approaches to characterising uncertainty for the management systems, including the harvest strategy approach.

Specific review:

- Critically review the ensemble approach that was applied for the 2021 southwest Pacific Ocean swordfish assessment (SC17-SA-WP-04) to capture both 'structural' and 'estimation' uncertainty.
- Conduct a similar review of the approaches used in SC18-SA-WP-03 (Report on WCPFC Project 107b: Improved stock assessment and structural uncertainty grid for Southwest Pacific blue shark).
- Considering the above reviews, provide recommendations for model ensemble construction, model retention, and weighting of models included within ensembles in the context of the WCPFC tuna, billfish and shark assessments.

Expected outcomes:

1. Provide a basis for stock assessment teams to better consider and apply alternative approaches for characterising stock assessment uncertainty (including model selection and weighting) across the WCPFC tuna, billfish and shark assessments;
2. Provide guidance to the SC on the approaches for capturing assessment uncertainty in the provision of management advice; and
3. Ultimately provide managers and stakeholders with a better understanding of the implications of alternative approaches to characterising uncertainty for their perceptions of risk.

It was also noted that the review comes on the back of two recent CAPAM workshops (Model Ensembles and Weighting and Tuna Stock Assessments Best Practices) discussing this topic, and those workshops provide some valuable additional background. Take away messages from those workshops included:

- We do not sufficiently understand the properties of various diagnostics to derive model weights automatically: universally applicable best practice for model weighting can't be identified.
- Model weighting should be agreed a priori based on axes (inputs), not outcomes, for the set of plausible models.

An example of a type of topology for characterising uncertainty was discussed showing how structural and parameter uncertainty might be integrated to represent overall uncertainty.

The outline of the work followed:

1. Initial discussions with assessments teams (SPC, ISC) - ASAP
2. Review tuna, billfish, and shark assessments (under way)
3. Review 2021 SWPO swordfish & 2022 SWPO blue shark model grid approaches (May)
4. Summarise these reviews and provide recommendations and guidelines to consider in reporting assessment uncertainty. (June)
5. Review outcome and suggestions with assessment teams (June/July)
6. Present outcomes at SC19

The progress of the review stage was displayed in spreadsheet to demonstrate the approach to reviewing the previous assessments.

The review will hope to identify ways to help make the treatment of uncertainties in management advice more transparent/standardised.

The workshop discussed the issue of weighting models and how this has typically occurred at the SC where delegates get to discuss the grid of models presented as the basis for management and have the opportunity to discuss removing particular models/grid axis or assigning weights to particular groups of models/axis combinations. It was noted that this also happens after delegates have seen the results and it is therefore possible to understand the implications of different weighting for the management advice. Further, it was noted that the recent CAPAM Tuna Assessment Best Practices workshop indicated that objective weighting of models in ensembles (i.e., using diagnostics) is still an area of research where it is difficult to define a general best practice, but that weighting models in the context of a regional stakeholder meeting after seeing results was clearly not good practice.

The workshop heard from the current and previous SC theme convenors on this issue of weighting model grids used for management. The previous theme convenor indicated that their preference was not to do the model weighting at the SC, but have SC presented with a grid with equal weightings and have the SC discuss if any gid axis should be removed. The current co-convenor did not express strong opinions but indicated that currently, equal weighting is a better way, but that there would need to be more discussion on this issue with ISC people where model grids are typically not used as a basis for management advice.

The workshop commented that while a priori weighting is desirable, there are circumstances where this is difficult without seeing results as sometimes it depends on plausibility of those results. The workshop also commented that if there is no weighting after you have seen the results, then it becomes critical to have a good scientific processes or forum for determining which options go into the grid, to consider data inputs and diagnostics to guide the recommendations for model grids.

The workshop commented that developing criteria to guide any weighting processes was more the concern. It was also asked whether the project would be considering operating model grids applied in MSE frameworks, that are developed with the objective of testing management procedures against a broader range of plausible alternative scenarios than a stock assessment. While the project will be focussed on the stock assessment aspects there is interest to explore how the approaches for stock assessment might differ from those for MSE operating models.

The workshop commented on a priori weighting versus equal weights, that a priori weighting allows for broader options for model grids, for example for biological assumptions there might be a range of options and some might be less plausible or require more information to support and these can be down weighted to create priors for these assumptions, whereas for some other alternatives, such as input data it can less clear of which are more reliable and equal weights will be a default. Some a priori weighting is probably necessary in situations with considerable uncertainty in biology and even data (i.e., CPUE for sharks), while for small model grids equal weighting might be OK.

Noting that Shiny Apps. are likely to be used to display model results this could provide opportunities to explore implications of including and removing models on management quantities. It would be good if this could be avoided in the design of the Shiny App. Perhaps provide a Shiny App. that does not show the management quantities?

The workshop commented that ultimately not having a grid weighting/axis selection discussion at the SC would be best. There could be a transparent process for selecting grids access, a priori weightings etc. that is agreed, then just trust the scientists to go away and apply this. It was noted that this was the approach suggested at CAPAM.

It was also noted the PAW is an SPC meeting and not an official WCPFC meeting, any decisions on models for management advice and approaches for making these decisions lie with the SC. This review project will hopefully help the SC to make improvements to their current approaches.

What do other RFMOs do in this process? The CAPAM diagnostics and weighting workshops were convened to try to develop more objective methods. For the IATTC there is a risk analysis where they weight models in a subjective process, but this is done by the stock assessment scientists at the tuna Commission by voting using diagnostics and other metrics. This tries to divorce the model weighting discussion from those that have an interest in the results. The CAPAM workshops were aimed at developing more quantitative objective methods using diagnostics etc. but failed to find a clear way to do this. The approach for now would be to use the model diagnostics to ensure you have 'good' models in the grids, either fix models with poor diagnostics or remove them if they can't be fixed, and then apply equal weight to what's left, until better methods can be developed. This approach seems likely for upcoming IATTC assessments.

A final suggestion was for the Scientific Services Provider to decide what uncertainties are in the grid, apply equal weightings, and suggest that the SC has a discussion to explore and develop an approach for future.

Diagnostics

Laura Tremblay Boyer (CSIRO) provided a summary of the results for the workshop survey on diagnostics.

The objective of this was to provide clear guidance from PAW as to the type of diagnostics and/or model outputs that would be desirable to inform SC review of the 2023 assessments.

The survey compiled 43 diagnostics / model output plots spanning data inputs, model development, fit to data, model prediction, model convergence and model information content themes. Of those, 34 have already been presented in at least one SPC assessment report and/or are already included in the Shiny app. It was noted most recommended diagnostics should be provided for the diagnostic model and a subset for grid models.

Fifteen workshop participants completed the survey. The discussion results focussed on the diagnostics that would be new, and the top four of these were:

1. Model convergence: Diagnostic case model showing positive definite hessian (13/15)
2. Model convergence: Jitter analysis (12/15)
3. Model convergence: List of parameters estimated at boundary values are listed (12/15)
4. Tag mixing diagnostics: Tag recovery density plots (8/15)

The following were listed as optional:

1. Age-structured production model (ASPM) diagnostic
2. Size (length and weight): Runs test on residuals
3. CPUE [Residual diagnostics disaggregated by modelled (and un-modelled) covariates]
4. CPUE [Runs test on residuals]
5. * Model convergence: Identification of any correlated parameters (> 0.95 or <-0.95)
6. Runs test on recruitment deviates over time (by region)
7. Tag mixing diagnostics [Custard and TART analyses (cf. Kolody & Hoyle 2015)]
8. Hindcast cross-validation of CPUE
9. Calculation of autocorrelation in recruitment deviates (by region)

Diagnostics that were suggested but not included in the original survey list included:

1. Empirical selectivity on the fleet with asymptotic selectivity assumptions (X 2)
2. One area model instead of a spatial model (using the areas in the spatial model as "fleets").
3. Likelihood profiles on parameters of interest
4. Catch curve diagnostic
5. "Regular" comp data and selectivity diagnostic
6. Development of tag mixing diagnostics for observed tag data is required
7. Size residual bubble plots by fleet and time period.

It was noted that the full set of diagnostics is not practical for the full grid of models and that it might be possible to generate full diagnostics for a subset of models, perhaps including some that represent extremes of stock status estimates.

The PAW was asked - should all grid models be required to have a positive definite Hessian (PDH)?

The discussion around the PDH diagnostic led to discussion around what to do if a PDH could not be achieved for grid models - at least the diagnostic case models. It was noted that previous assessments had been accepted as best available science without PDH solutions based on consideration of the overall

assessments and other diagnostics, but it was also noted that last year it became clear that some delegations would now be seeking PDH solutions for at least the diagnostic case models as mandatory requirements. The recommendation from PAW, and the preference for this diagnostic to be satisfied was noted, however it was reiterated that the PAW, being an SPC meeting, can't set criteria for SC on how to accept or reject assessments or components of grids; those criteria/decisions will be up to the SC to determine.

Clarification was sought on the empirical selectivity diagnostic: this basically takes the observed length frequency and divides it by the model predicted numbers /proportions at length. It estimates what the selectivity curve should look like. This is for length-based selectivity, for age-based selectivity you would have to convert the numbers at age to lengths.

In relation to the other diagnostics, the value of the size residual bubble plots by fleet and time period was noted and the development of tag mixing diagnostics for observed tag data was noted to cover tag recapture density plots and the CUSTARD and TART analyses. SPC scientists were interested in the tag mixing diagnostics but were unclear on how these get used to inform decisions that the assessment team have to make on these assumptions, as at least for yellowfin it can be expected that these analysis won't indicate that tags are well mixed over any reasonable period, so more discussion between the tagging data analyst and the assessment scientists was required on if and or how to apply these analyses.

Discussion of assessment models/spatial structure/uncertainties

The meeting revisited the discussion on the approaches for the bigeye and yellowfin assessments, starting with the consideration of the spatial stratification.

A proposed approach to step from a 9 region to a 4 region model structure was put forward for discussion (Fig. 2). In putting this forward it was noted that there would not be time to do what some workshop members had suggested in starting from a single region model and introducing complexity. The 9 region model would have to be the starting point and complexity reduced from there. Further, exploration of simpler stratification would likely need to be done on the old data as there is not enough time to do all this after the new data are finalised.

There was a productive discussion that considered/suggested a range of options, that included:

- Dealing with tagging data when smaller regions that were originally implemented to cater for tag data.
- Supporting the approach to step through models the reduce spatial complexity incrementally.
- Consider running an areas as fleets model for either the equatorial regions (as a single region) or for the entire model area as a single region area as fleets model to avoid the need for estimating movement.
- Criteria for deciding if a simpler model is to be chosen, and then which one.
- If it's not obvious which spatial stratification is best, consider if different spatial structures are included as a grid axis.

- In terms of choosing the best models, different models may perform better in different areas than others. As a result, it may not be straightforward to choose, and there might be difficulty to exclude the spatial structure uncertainty from the uncertainty characterisation. But these decisions might be best left to the scientists. This needs to be communicated well in terms of reasoning for any decisions.
- Stepping though models without the tag data might save time.
- With this exploration of different regional configurations there are perhaps two options for how movement gets estimated: 1. Leave the tag data in and adjust the mixing periods in a sensible way depending on the spatial configurations, 2. If tag data is not used then apply the SEAPODYM movement coefficients as fixed and not estimate them. Allowing the model to estimate movement in a very unconstrained way may not be sensible.
- A suggestion was made to step from the 9 to 6 to 4 region model to simplify the process and workload. It was noted that if there are major changes in estimates you might have to go back and run the intermediate structures to explore the change.
- The exploration of alternative structures will be easier if the same data structure is used, fishery definitions, size data compilation, tag data file (might even help to keep tag files as is, for example to allow separate mixing periods for different releases within regions, retain the retain 8 tag structure). But won't be able to preserve the data structure for CPUE and these will have to be recalculated for each structure.
- Will need to think carefully about how to present the results, it is not the desired outcome to have two spatial structures to choose from but rather come up with a simpler model structure that works better. So, we'd want to avoid a situation where we present results for models with different structures and then SC chooses one.
- Finally, it was noted that this exploration was prompted from the peer review that noted that the 9 region model was not well estimated and unstable, and the data was not informative enough for the model spatial complexity. So we are trying to fix the issues with the 9 region model, and simplification was seen as an avenue for this. It is well accepted that a simpler model that performs as well or better than a more complex model is a better option (parsimony). There are various diagnostics that can help to inform the decision on a recommended structure.

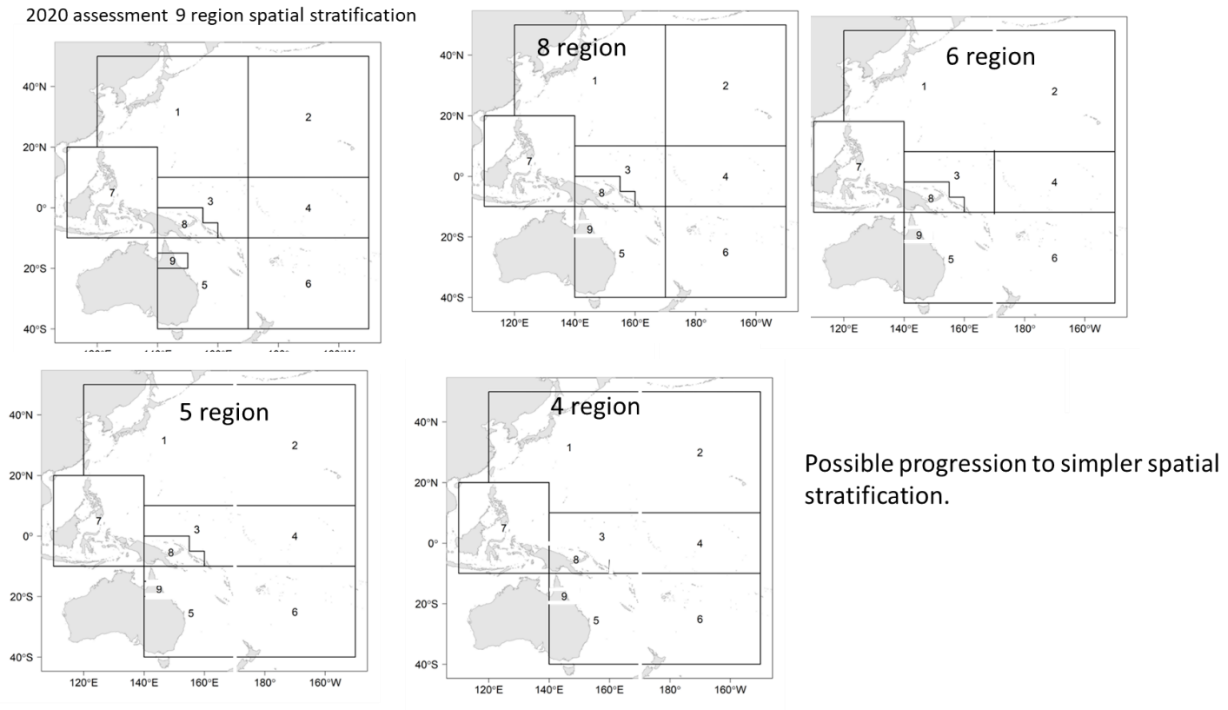

Figure 2. Possible stepwise approach to explore reducing spatial complexity.

The discussion moved on to planning for the model stepwise (bridging) analysis, sensitivities and important uncertainties. These were outlined in a table discussed by Jemery Day, noted as 'preliminary' options.

Likely sensitivity analyses for both yellowfin and bigeye

- Spatial stratification 9, 8, 6, 5, 4 regions.
- Alternative base M (Hoyle et al. 2023) and M estimated in model, Lorenzen form for M at age.
- Size comp pre-treatment/filtering approaches.
- Growth: estimating Richards and VB growth curves internally.
- Data weighting: Explore whether empirical data weights can be used, comparing residuals to statistical distributions. Lognormal CV for CPUE. If Dirichlet multinomial is not used, may consider iterative reweighting for length and Punt method for conditional age-at-length data.
- Selectivity: Explore merging/splitting of fisheries, as well as sharing selectivity between similar fisheries.
- CPUE models:
    - Apply CPUE developed from regions modelled independently and scaled by the global model. Does this form of scaling give reasonable MFCL outputs and fits to the data? This may provide more accurate estimates of region-specific parameters in the CPUE standardization and, possibly, less biased CPUE indices, but regional scaling needs to be maintained.

    - In the previous assessments, MFCL did not fit the seasonal pattern in CPUE. Apply CPUE developed from seasonal model to understand if accounting for seasonality in the CPUE standardization improves the MFCL fits to the CPUE. Since catchability is not estimated in the catch-conditioned model, there is no confounding between seasonal catchability and a seasonal CPUE standardization model.
    - Apply variance-covariance matrix for CPUE in MFCL. Can ungroup selectivity but catchability needs to be grouped for regional scaling.
- Tagging: Consider dropping tags from region 9, or not using them for movement estimates. Explore tag mixing options, especially for spatial structures with fewer regions.
- Technical/effort creep in LL index fisheries, perhaps 1% per year as a sensitivity model run.
- Steepness: standard range for values of *h* included, this is in the uncertainty grid either way.
- Movement: Externally estimated movement coefficients (SEAPODYM).

Discussion on the candidate uncertainties for inclusion in the model grid used for management advice, indicted the following, with those in 'bold' expected to be included:

- **Steepness** (as per previous assessments)
- Growth – possibly
- **Tag mixing**
- **Movement** (estimated, SEAPODYM age specific)
- Data weighting
- Spatial structure – possibly
- **Natural mortality**

The workshop finally noted recommendations for research projects that came out of discussions:

- Investigation of effort/efficiency creep in the longline fishery and review of the approaches to analysis and use of longline CPUE as abundance indices.
- Review of the size composition data used in WCPFC tuna assessments to generate 'best representative' data sets.
- Developing a plan for enhancing otolith collections and routine age/length data collection.
- Beyond MFCL: Scoping study for a next generation tuna assessment modelling software.
- Close kin mark recapture scoping study for a key shark species to be determined by the shark IWG.
- Tuna life history information: identification of key gaps and research needs/approaches to fill these.
- Key shark life history information: identification of key gaps and research needs/approaches to fill these.

Meeting closed

The meeting was closed by the chair, expressing that the hybrid meeting was a success and provided great value to the teams charged with doing the assessment work. The hybrid approach would be followed for

next year's PAW. The chair thanked the presenters, the OFP team, the Dragonfly/shark assessment team, Karen Kaspan (SPC) for meeting organisation support, the online attendees who stuck it out and in different time zones, and those that travelled to join in person. Graham Pilling finally thanked the chair and the attendees for their participation and valuable input.

Post-meeting follow-up

Following the meeting, the assessment team further considered the stepwise analysis approach, taking into consideration timelines and logistics around preparation of data inputs for alternative structural configurations. The decision was made to follow a stepwise process whereby the 9 region is model is maintained through all steps up until and including the full data updates and completed CPUE analysis, and then explore the implications of alternative spatial stratifications. Some steps involve sensitivities that guide decisions on options to apply to the diagnostic case or to include as grid axes, and these are explored as part of the relevant steps. This is to avoid the need to explore implications/interactions of alternative spatial stratifications with the many other stepwise changes, as this was considered inefficient and unduly time consuming. The models with different spatial stratifications will necessarily require some different assumptions when it comes to tag mixing, and the applications of alternative information such as SEAPODYM, and other age specific or fixed movement assumptions. Exploring these aspects on alternative spatial stratifications once all other steps had been completed was viewed as more tractable and less prone to becoming caught up in the stepwise development, i.e., get a new 9 region diagnostic model in the bag before exploring/comparing the alternative spatial configurations. Finally, time constraints will be limiting to do a full stepwise approach to simplifying spatial stratification and we now aim to focus on the 5 region structure as an alternative, if time permits we may explore other options such as an equatorial only model.

References

Buenafe, K.C.V., Everett, J.D., Dunn, D.C., Mercer, J., Suthers, I.M., Schilling, H.T., Hinchliffe, C., Dabalà, A., Richardson, A.J., 2022. A global, historical database of tuna, billfish, and saury larval distributions. Scientific Data 9:423, 1-9. https://doi.org/10.1038/s41597-022-01528-7

Hoyle, S.D., Williams, A.J., Minte-Vera, C.V., Maunder, M.N. 2021. Approaches for estimating natural mortality in tuna stock assessments: application to Indian Ocean yellowfin tuna. Fisheries Research 257, 106498. https://doi.org/10.13140/RG.2.2.29735.42400

Ijima, H., Jusup, M., 2023. Tuna and billfish larval distributions in a warming ocean. http://arxiv.org/abs/2304.0944

Matsubara, N., Aoki, Y., Tsuda, Y., 2022. Historical development of fishing devices in Japanese pole-and-line fishery. WCPFC-SC18-2022/SA-IP-16.

Moore, B.R., Bell, J.D., Evans, K., Farley, J., Grewe, P.M., Hampton, J., Marie, A.D., Minte-Vera, C., Nicol, S., Pilling, G.M., Scutt Phillips, J., Tremblay-Boyer, L., Williams, A.J., Smith, N., 2020. Defining the stock structures of key commercial tunas in the Pacific Ocean I: Current knowledge and main uncertainties. Fisheries Research 230, 105525. https://doi.org/10.1016/j.fishres.2020.105525

Nishikawa, Y., Honma, M., Ueyanagi, S., Kikawa, S., 1985. Average distribution of larvae of oceanic species of scombroid fishes, 1956–1981. Far Seas Fish Res Lab, 99 p.

Reglero, P., Tittensor, D., Álvarez-Berastegui, D., Aparicio-González, A., Worm, B., 2014. Worldwide distributions of tuna larvae: revisiting hypotheses on environmental requirements for spawning habitats. Marine Ecology Progress Series. 501, 207–224. https://doi.org/10.3354/meps10666

Appendices

APPENDIX 1: Agenda

Note the agenda below covers the meeting content however to order of topic was modified during the meeting to adapt to timing and flow if topics.

2023 SPC Pre-assessment Workshop Agenda (original version) 25th- 28th April, Noumea and hybrid.

Times are New Caledonia

Chair: Paul Hamer, pauh@spc.int
Alternates: Graham Pilling, grahamp@spc.int Claudio Castillo Jordan, claudioc@spc.int

Tuesday 25th April (Mon 24th US)	**DAY 1: 2023 yellowfin tuna assessment**	**Presenter initials and presentation number**
09:00 – 09:15	**Introduction** • Reminder of TOR and objectives for the SPC preparatory workshop • Agenda and meeting format/procedures • Any other introductory comments	PH
09:15 – 10:00 *Session 1* *(45 mins)*	**Background and peer review** • Previous yellowfin tuna assessment summary (10 mins) • Peer review summary – key recommendations and related focus areas for 2023 assessments (15 mins) Discussion (15 mins)	• AM **(P1)** • AM/PH **(P2)**
10:00 – 10.30	BREAK	
10.30-12.00 *Session 2* *(90 mins)*	**Conceptual models of biology and spatial structure** • Consider information and options for spatial structure and fishery definitions (60 mins) • Discussion (20 mins)	• PH (background and biology) **(P3)** • IS (SEAPODYM) **(P4)** • JP (Size composition) **(P5)**

		• JM (CPUE patterns) **(P6)** (15 mins each)
12.00-13.00 *Session 3* *(70 mins)*	**Data inputs** • Catch and effort, and raw size data (15 mins) • Implications of reduce observer coverage (5 mins) • Size composition data treatment - spatial/temporal weighting (15 mins) • Basis for CPUE indices (10 mins) • Conditional age at length data summary (10 mins) • Tagging data summary, tagger effects, tag seeding/reporting rates (10 mins)	• TT **(P7)** • TP **(P8)** • TT (**P9**) • TT/AM (**P10**) • JS (**P11**)
13.10-14.00	Lunch BREAK (50 mins)	
14:00 *Session 4* (30 *mins)*	**Biology** • Growth • Natural mortality • Maturity/reproductive biology • Movement • Length-weight conversions (20 mins) Discussion (10 mins)	• AM (**P12**)
14.30-15.40 *Session 5* *(70 mins)*	**CPUE analysis (covers yellowfin and bigeye)** • Background • From VAST to sdmTMB • CPUE results so far, peer review and related analyses Discussion (20 mins)	• TT **(P13)**
15.40-17.00 *Session 6* *(80 mins)*	**Model development** • 2023 diagnostic model development • Preliminary model results • Proposed next stepwise models, other sensitivities • Data weighting • Diagnostics • Uncertainties to consider, uncertainty characterization for management advice	• AM/JH/ND and all **(P14)**

	Discussion (20 mins)	
17.00	Discussion and wrap up for the day (as needed)	All/PH
Wednesday 26th April (Tuesday 25th US)	**DAY 2: 2023 bigeye tuna assessment and external presentations**	
09:00 –10.00 *Session 7* *(60 mins)*	**External presentations: age/growth, conversion factors** ● Bomb radiocarbon age validation work (15 mins + 5 mins) ● Yellowfin sea cage growth experiment (15 mins + 5 mins) ● Project 90: conversion factors (15 mins + 5 mins)	● AA **(P15)** ● KO **(P16)** ● JM **(P17)**
10.00-10.30 *Session 8* *(30 mins)*	● Previous bigeye tuna assessment summary (15 mins) ● Implication from yellowfin peer review (10 mins)	● JD **(P18)**
10:30 – 11.00	BREAK (30 mins)	
11.00-12.00 *Session 9* *(60 mins)*	**Conceptual models of biology and spatial structure** ● Consider information and options for spatial structure and fishery definitions (40 mins) ● Discussion (20 mins)	● PH (background and biology) **(P19)** ● IS (SEAPODYM) **(P20)** ● JP (Size composition) **(P21)** ● JM (CPUE patterns) **(P22)** (10 mins each, focus on results)
Session 9 *12.00 -13.00* *(60 mins)*	**Data inputs** ● Catch and effort, and raw size data (15 mins) ● Implications of reduce observer coverage (5 mins)	● TT **(P23)** ● TP **(P24)** ● TT **(P25)** ● TT/AM **(P26)**

	• Size composition data treatment - spatial/temporal weighting (15 mins) • Basis for CPUE indices (10 mins) • Conditional age at length data summary (10 mins) • Tagging data summary, tagger effects, tag seeding/reporting rates (10 mins)	• JS **(P27)**
13.00-14.00	Lunch BREAK (60 mins)	
Session 10 *14.00 -14.30* *(30 mins)*	**Biology** • Growth • Natural mortality • Maturity/reproductive biology • Movement • Length-weight conversions • Sex specific? (20 mins) Discussion (10 mins)	• JD **(P28)**
14.30-15.10 *Session 11* *(40 mins)* *(reduced as methods covered for yellowfin)*	**CPUE analysis** • CPUE discussion follow-up (30 mins) Discussion (10 mins)	• TT **(P29)**
15.10 -16.20 *Session 12* *(70 mins)*	**Model development** • 2023 diagnostic model development • Preliminary model results • Proposed next stepwise models, other sensitivities • Data weighting • Diagnostics • Uncertainties to consider, uncertainty characterization for management advice Discussion (20 mins)	• JD/JH/ND and all **(P30)**
16.20	Discussion and wrap up day (as needed)	All
Thursday 27th April (Wed 26th US)	**DAY 3: MFCL, MSE, and Skipjack assessment follow-up work**	

09.00-9.15	Overnight thoughts/follow-ups	PH
9.15-10.30 *Session 13* *(75 mins)*	**MFCL** ● Recent developments and future work for Multifan-CL (25 mins) ● Looking forward – future software for tuna assessments – what's next for MFCL (20 mins) Discussion (20 mins)	● ND **(P31)**
10:30 – 11.00	BREAK	
11.00-11.20 *Session 14* *(20 mins)*	**Tuna Management Strategy Evaluation** ● Update on progress and 2023 technical workplan (10 mins) Discussion (10 mins)	● RS/FS **(P32)**
11.20-13.00 *Session 15* *(40 mins)*	● MSE – Skipjack MP (40 mins)	● RS **(P33)**
13.00-14.00	● Lunch BREAK	
14.00-15.00- *Session 16* *(60 mins)*	● MSE – Albacore MSE, OMs and EMs (40 mins) ● MSE – mixed fisheries overview and update (10 mins) Discussion (10 mins)	● RS/FS/NY **(P34)** ● FS **(P35)**
15.00-16.00 *Session 17* *(60 mins)*	● Skipjack stock assessment follow-up work (40 mins) (Discussion 20 mins)	● CCJ/JH/PH **(P36)**
	BBQ on the deck 6pm	
Friday 28th April (Thurs 27th US)	**DAY 4: SEAPODYM, Silky shark assessment, tuna research plan, unfinished discussions.**	

9.30-10 30 pm *Session 18* *(60 mins)*	• SEAPODYM overview and development plan	• IS **(P37)**
10.30 -11.00	Break	
11.00 -12.30 Session 19 (90 min)	• Previous silky shark assessment summary • Available data summary/fishery characterization • Plans for catch reconstruction and potential assessment approach. • Preliminary analysis	• PN **(P38)**
12.30 -13.00 Session 20 (30 min)	• Tuna research plan review	• **GP (P39)**
13.00-14.00	Lunch BREAK	
14.00-15.30 *Session 21*	• Review on approaches for ensemble model development and characterising uncertainty in WCPFC assessments (30 mins) • Diagnostics • Unfinished discussions as needed	• PN (**P40**)
15.00 Wrap up and Follow-up	**WRAP UP-Follow-up** • Note any key recommendations • Meeting draft paper circulated for comments • Comments received • Meeting paper finalized for SC16 submission • Aim for end of May	PH

PH Paul Hamer, AM Arni Magnusson, CCJ Claudio Castillo Jordan, ND Nick Davies, JD Jemery Day, TT Thom Teears, RS Rob Scott, FS Finlay Scott, TP Tom Peatman, KO Kei Okamoto, NY Nan Yao, PN Philipp Neubauer, JSP Joe Scutt Phillips, JH John Hampton, AA Allen Andrews, GP Graham Pilling, IS Inna Senina, JP Joanne Potts.

APPENDIX 2: List of participants

Name	Affiliation	In person (IP) / Online (O)
John Annala	Ministry for Primary Industries, NZ	O
Leyla Knittweis	Ministry for Primary Industries, NZ	IP
Bradley Moore	NIWA, NZ	O
Simon Hoyle	NIWA, NZ	O
Philipp Neubauer	Dragonfly Data Science	IP
Kath Large	Dragonfly Data Science	O
Kyuhan Kim	Dragonfly Data Science	O
Steven Brouwer	Sagittas LTD	O
Nick Davies	Takina LTD, SPC consultant	IP
Rob Campbell	CSIRO, AU	IP
Jessica Farley	CSIRO, AU	O
Naomi Clear	CSIRO, AU	O
Paige Eveson	CSIRO, AU	O
Ashely Williams	CSIRO, AU	O
Laura Tremblay Boyer	CSIRO, AU	IP
James Larcombe	Department Agriculture Water and the Environment, AU	O
Lianos Triantafillos	FFA Secretariat	O
Keith Bigelow	NOAA (Pacific Islands Fisheries Science Centre), US	IP
Jon Brodziak	NOAA (Pacific Islands Fisheries Science Centre), US	O
Michelle Sculley	NOAA (Pacific Islands Fisheries Science Centre), US	O
Nicholas Ducharme-Barth	NOAA (Pacific Islands Fisheries Science Centre), US	IP
Mark Maunder	IATTC	O
Haikun Xu	IATTC	O
Carolina Minte-Vera	IATTC	O
Eric Chang	National Sun Yat-sen University, TW	IP
Yi-Jay Chang	National Sun Yat-sen University, TW	O
Ren-Fen WU	Overseas Fisheries Development Council, TW	IP
Keisuke Satoh	Japan Fisheries Research and Education Agency	IP
Kei Okamoto	Japan Fisheries Research and Education Agency	O
Yuichi TSUDA	Japan Fisheries Research and Education Agency	IP
Yoshinori Aoki	Japan Fisheries Research and Education Agency	IP

Takaaki Hasegawa	Japan Fisheries Research and Education Agency	IP
Hirotaka Ijima	Japan Fisheries Research and Education Agency	O
Francisco 'Curro' Abascal	EU	IP
Zhe Geng	Shanghai Ocean University	O
Jiaqi Wang	Shanghai Ocean University	O
Meng Xia	Shanghai Ocean University	O
Yiqian Shi	Shanghai Ocean University	O
Xiaodong Li	Shanghai Ocean University	O
Yuchen Huang	Shanghai Ocean University	O
Haewon Lee	National Institute of Fisheries Science, KR	O
Haewon Park	National Institute of Fisheries Science, KR	O
Jung Hyun Lim	National Institute of Fisheries Science, KR	O
Anis Aisi	Fisheries, PNG	O
SungKwon Soh	WCPFC Secretariat	IP
Elain Garvilles	WCPFC Secretariat	O
Tim Adams	Representing Kiribati	IP
Allen Andrews	University of Hawaii/Uppsala University	O
Claudio Castillo Jordon	SPC	IP
John Hampton	SPC	IP
Graham Pilling	SPC	IP
Peter Williams	SPC	IP
Sam McKechnie	SPC	IP
Rob Scott	SPC	IP
Finlay Scott	SPC	IP
Nan Yao	SPC	IP
Jemery Day	SPC	IP
Arni Magnusson	SPC	IP
Thom Teears	SPC	IP
Paul Hamer	SPC	IP
Joe Scutt Phillips	SPC	IP
Jed Macdonald	SPC	IP
Steven Hare	SPC	O
Joanne Potts	SPC	IP
Inna Senina	SPC	IP
Patrick Lehodey	SPC	IP
Lucas Bonnin	SPC	IP
Raijeli Natadra	SPC	IP
Moses Mataika	SPC	IP
Tom Peatman	Independent Consultant	IP

APPENDIX 3: Terms of Reference

The Oceanic Fisheries Programme (OFP) of SPC is contracted by WCPFC to undertake stock assessments. The results of these assessments will be presented at the WCPFC Scientific Committee. In preparation for these assessments, OFP is hosting a pre-assessment workshop to discuss key issues related to the assessments. The terms of reference for this workshop are provided below.

Terms of Reference

- Review the most recent completed assessments, in particular, any concerns, suggestions and/or recommendations raised by the Scientific Committee, the Commission, research providers, individual CCMs, or any independent reviews;
- Review preliminary work undertaken by the service provider relating to the stock assessments, including any proposed:
    - revisions to biological parameters
    - revisions to historical data
    - changes to structural assumptions in the model
    - methodological issues, e.g., characterization of uncertainty
    - standardized CPUE analysis
    - incorporation of tagging data or other auxiliary data
- Provides guidance to the OFP on:
    - the suitability of any proposed changes and any suggested additional work
    - a minimum set model runs to be undertaken, in particular the range of key sensitivity analyses
    - desired model diagnostics to be presented.
    - alternative modelling approaches that could be considered

The outcomes of the meeting will be documented in two ways, a report of the meeting and in the assessment working papers themselves. The report of the meeting will be distributed to workshop participants for comment within 10 working days of the meeting and revised and provided to WCPFC Scientific Committee members 30 days after the meeting. It will also be submitted to the next Scientific Committee as a Working Paper. Many of the matters discussed to the workshop will be the subject of meeting papers to the Scientific Committee.

Due to the timing of the meeting, any model runs presented will be based on previous assessment data sets, and therefore no preliminary stock assessment runs will be undertaken. Further, the workshop will occur prior to the submission of data and completion of supporting analyses (e.g., CPUE analyses). Therefore, any major changes to historical data submitted by CMM's, or new data could result in a need to consider alternative model runs or structures not considered previously. In such instances, supporting documentation will be provided to the SC via working papers to allow the SC to determine the merits of any proposed changes.

The consultation will be open to participation by all CCMs and to other experts, by invitation. CCMs will be expected to fund their participation although SIDS and participating territories may seek support from the Commission's Special Requirements Fund or other sources, as appropriate.